Taxonomic and Functional Characterization of Human Gut Microbes Involved in Dietary Plant Lignan Metabolism

Isaac M. Elkon

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2015

Committee:

Johanna W. Lampe

Meredith A.J. Hullar

Program Authorized to Offer Degree:

School of Public Health

University of Washington

Abstract

Taxonomic and Functional Characterization of Human Gut Microbes Involved in
Dietary Plant Lignan Metabolism

Isaac M. Elkon

Chair of the Supervisory Committee:
Research Professor Johanna W. Lampe
Department of Epidemiology

Background: Dietary plant lignans, such as secoisolariciresinol diglucoside (SDG), are metabolized to the enterolignans, enterodiol (END) and enterolactone (ENL), by gut microbes. Evidence suggests that enterolignans may reduce risk of cardiovascular disease and several forms of cancer. Our aim was to characterize the microbial community involved in enterolignan production by using an *in vitro* batch culture system to enrich for lignan-metabolizing organisms.

Methods: Stool samples from eight participants were incubated separately for ~1 week with a mineral salts media containing formate, acetate, glucose, and 6.55 μM SDG. Daily secoisolariciresinol (SECO), END, and ENL concentrations were measured using gas chromatography–mass spectrometry (GCMS). Microbial community in initial stool (Day 1) and *in vitro*-incubated fecal suspensions (final-day) was assessed via Illumina paired-end 16S rRNA gene amplicon and whole-metagenome shotgun sequencing. 16S rRNA gene sequences were taxonomically annotated using an in-house QIIME pipeline. Metagenomic shotgun sequences

were taxonomically annotated using MetaPhlAn and functionally annotated using DIAMOND and HUMAnN. Annotation-based alpha diversity, organism abundance, and functional gene abundance were used to assess differences between Day 1 and final-day microbial community composition.

Results: Stool of all 8 participants converted the majority of SDG to ENL by Day 6 (average $\pm$ standard deviation: 88.4% $\pm$ 6.9). In most incubations, Bacteroidetes:Firmicutes ratio increased; methanogenic, deglycosylating, and demethylating organisms were enriched; and KEGG Orthology gene families (KOs) corresponding to beta-glucosidases were present at high levels. In some incubations, *Eubacterium limnosum*, *Akkermansia muciniphila,* and KOs corresponding to demethylases, dehydroxylases, dehydrogenases, methane metabolism, the Wood-Ljungdahl pathway, and acetogenic bacterial O-demethylation were enriched.

Conclusions: Organisms and functional genes identified or hypothesized to be involved in lignan metabolism were not unanimously enriched, but sample-specific. These findings may help further elucidate organisms, enzymes, and biochemical pathways involved in lignan metabolism. Future studies with larger sample sizes may help to confirm these results.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Chapter 1. BACKGROUND

1.1 INTRODUCTION

Evidence suggests that lignans, one of the many bioactive phytochemicals in the human diet, may reduce the risk of several chronic diseases.[1,2] Below I briefly review lignan sources, metabolism, and health effects, and discuss the crucial role that gut bacteria play in human exposure to lignan metabolites.

1.2 PLANT LIGNANS IN THE DIET

Lignans are polyphenolic compounds composed of two dimerized cinnamic alcohols.[2] They are found in seeds, grains, leaves, fruits, and woody portions of various plants.[1] Humans ingest many lignan-containing plant products, including: flax, sesame, cloudberry, hemp, and blackberry seeds, cereal grains, dates, apricots, prunes, garlic, asparagus, brassica vegetables, beer, wine, coffee, and tea.[3] Lignans found in the human diet include secoisolariciresinol diglucoside (SDG), secoisolariciresinol (SECO), medioresinol (MED), matairesinol (MAT), lariciresinol (LARI), pinoresinol (PINO), syringaresinol (SYR), and sesamin (SES).[4] The degree to which each of these lignans are found in the diet is subject of debate.[5] SDG is the primary lignan found in flaxseed, the richest known dietary source of lignans.[3]

The gut microbiota are crucial to the metabolism of lignans in humans. When lignans are ingested, metabolism and absorption appears to be quite rapid.[6,7] Because of short transit time in the mouth, oral enzymatic digestion is unlikely.[4] Lignans also appear to be resistant to the low pH of the stomach and are not metabolized by gastric enzymes.[8-10] The majority of lignan bioavailability is due to microbial metabolism in the ileum and the large intestine, where great numbers of human microbes reside.[4,10] This is well supported by the timing of lignan metabolite appearance in the blood in relation to transit time to colon,[4] as well as the disappearance of measured lignan metabolites with antibiotic administration.[11]

Upon entering the colon, lignans are metabolized by microbes into the bioactive compounds enterodiol (END) and enterolactone (ENL), known as enterolignans.[1] I will focus on the

degradation of SDG, but other dietary lignans are metabolized to enterolignans similarly, sharing many of the same reactions and intermediates.[12]

SDG undergoes several transformations mediated by gut microbes: *O*-linked deglycosylation to SECO, which can be *O*-link demethylated to dihydroxyenterodiol (DHEND), which can be dehydroxylated to END, which can be dehydrogenated to ENL. DHEND can also be dehydrogenated to dihydroxyenterolactone (DHENL), which can be dehydroxylated to ENL.[4,13,14] These reactions appear to be enantiomer-specific, with chirality of the precursor lignan maintained throughout the reaction series.[15] Either the (+) or (-)-enantiomer of SDG appears to dominate depending on the flax species.[16,17]

Evidence suggests that most lignan absorption occurs in the colon.[18] Enterolignans can be absorbed, conjugated, and excreted by enterocytes into the hepatic portal vein.[19,20] Unconjugated lignans can also be absorbed from the lumen into the lymph and enter directly into systemic circulation.[21,22] Enterolignans can be conjugated in the liver, excreted in the bile, acted on again by gut microbes, and reabsorbed – undergoing enterohepatic circulation.[23] Enterolignans conjugated in the liver can also be released into systemic circulation. Once in the systemic bloodstream, they can be deposited in tissues throughout the body.[24,25] Enterolignans are excreted from the body via urine – mainly in the mono-glucoronidated form – and feces.[26]

1.3 HEALTH EFFECTS OF LIGNANS

As reviewed by Clavel and Mapesa in 2013,[27] evidence suggests that enterolignans may reduce the risk of cardiovascular disease and prostate, breast, and colon cancer. Several human dietary interventions have been conducted showing reduced prostate cancer cell proliferation,[28] breast cancer tumor growth,[29,30] and low-density lipoprotein (LDL) concentration[31,32] with flaxseed supplementation. In a prospective cohort study of women (221 cases, 886 controls), lignan intake was associated with decreased colon cancer risk.[33] Despite observed results, it is difficult to isolate the effects of lignans in these studies, since dietary sources and supplements contain many bioactive compounds in addition to lignans.[27]

Few randomized controlled trials (RCT) using pure lignans have been carried out on humans. When 22 post-menopausal women were given a flaxseed-extracted lignan complex containing 500 mg/day SDG for 6 weeks they had significantly lower plasma C-reactive protein levels.[34] In another study, 78 subjects with benign prostatic hyperplasia that were given a

flaxseed extract containing 300 mg/day SDG for 4 months had improved symptoms and quality of life scores.[35]

There have also been several interventions conducted in animals. In murine models, SDG administration has lowered mammary tumorigenesis,[36,37] lung metastasis,[38] and colon carcinogenesis.[39] SDG administration was shown to lower rates of atherosclerosis in rabbits[40] and have cardio-protective effects in rats.[41]

Several studies using *in vitro* models help explain the bioactive properties of enterolignans. ENL can bind to human pregnane X receptor, affecting xenobiotic detoxification and steroid metabolism.[42] ENL can weakly stimulate or block estrogen receptors – especially ER-α – potentially modulating estrogen signaling.[43,44] ENL and END can also bind to sex hormone-binding globulin which may affect circulating steroid hormone levels.[45] Enterolignans can affect cell proliferation in colon, prostate, and breast *in vitro*-cultured cells.[27] They have been shown to have stronger antioxidant properties *in vitro* than lignan precursors.[46,47]

1.4 INTER-INDIVIDUAL DIFFERENCES IN ENTEROLIGNAN PRODUCTION

Average enterolignan blood concentration for an unsupplemented individual is estimated to be 10 - 25 nmol/l,[4] but varies considerably, and an ENL concentration over 1000 nmol/l has been documented in one individual. Dietary patterns can account for some of this variation and intake of vegetables and grains has been associated with higher serum enterolignan levels.[48] Furthermore, human dietary intervention studies have shown higher serum enterolignan levels with increased intake of lignan-rich foods.[18,49] Similarly, human dietary flaxseed supplementation has been shown to increase urinary enterolignan excretion.[49,50]

Lignan intake levels, however, do not fully account for inter-individual variation. When 12 individuals were administered the same quantity of SDG, there was great variation in plasma enterolignan profiles – both total concentration of enterolignans and ratio of END:ENL.[18] Inter-individual variation in urinary and plasma END:ENL ratios has also been documented in several dietary intervention studies.[49,50] When stool from 100 different participants was incubated *in vitro* with a flaxseed extract (final concentration in fecal suspension: 432 – 864 mg SDG per liter), END and ENL were produced in 63% and 39% of the samples, respectively.[5] Differences in enterolignan status could be explained by many factors in addition to level of dietary intake –

including gut transit time, antibiotic use, and smoking status – but gut microbial composition appears to play a key role.[18,51-53]

1.5   ORGANISMS INVOLVED IN LIGNAN METABOLISM

The populations of bacteria responsible for SECO to END conversion are subdominant in the human gut, on the order of 10^8 cells/g stool, compared to 10^{12} total bacterial cells/g stool.[5,51] Those responsible for SECO to ENL conversion, however, appear to be on the order of 10^5 cells/g stool.[18,51] Women appear to possess more enterolignan-producing bacteria than men.[51] One explanation for this is that lignan-metabolizing microbes may also be capable of acting on female sex hormones, such as progesterone. Thus, higher levels of colonic female sex hormones might enrich for lignan-metabolizing microbes.[4]

In a cross-sectional study of 115 premenopausal women (ages 40 - 45), the bacterial 16S rRNA gene was amplified and sequenced from stool sample-extracted DNA, and urinary excretion of ENL was measured.[54] ENL production was associated with higher community alpha-diversity. Furthermore, presence of genera *Moryella, Acetanaerobacterium, Fastidiosipila, and Streptobacillus* in stool samples was significantly greater in the highest tertile of ENL excreters. The genera *Fastidiosipila* and *Streptobacillus* are related to bacteria that are capable of cleaving glucose from complex substrates[55-57] and hence may be capable of performing SDG to SECO conversion.

In another cross-sectional study of 84 men, serum concentration of ENL was measured using a time-resolved fluoroimmunoassay, and several different bacterial groups were enumerated in stool samples using fluorescence *in situ* hybridization. Those with the highest serum ENL concentration had a significantly higher fecal bacterial cell count, particularly of the *Lactobacillus-Enterococcus* group, when compared to those with the lowest serum ENL concentration. *Enterococcus faecalis* PDG-1 has been shown to convert pinoresinol to lariciresinol.[15] *Lactobacilli* and *Enterococci* both exhibit β-glucosidase activity, and thus may be capable of deglycosylating SDG.[58] Ruminal *Prevotella* spp. have been shown to deglycosylate SDG as well.[59]

Using *in vitro* incubation of human stool, at least 29 species or strains of bacteria capable of performing lignan metabolism have been isolated and identified, as recently summarized by Yoder et al.*[12] Bacteroides fragilis, Bacteroides ovatus, Clostridium cocleatum, Clostridium*

saccharogumia, Clostridium ramosum, Bacteroides distasonis, Bifidobacterium bifidum WC 418, Bifidobacterium breve WC 421, Bifidobacterium catenulatum ATCC 27539 , Bifidobacterium longum subsp. infantis ATCC 15697, Bifidobacterium longum subsp. longum WC 436, Bifidobacterium longum subsp. longum WC 439 , Bifidobacterium pseudocatenulatum WC401, Bifidobacterium pseudocatenulatum WC 402, Bifidobacterium pseudocatenulatum WC 403, and *Bifidobacterium pseudocatenulatum WC 407* have been shown capable of lignan deglycosylation.[8,60-62] *Butyribacterium methylotrophicum, Eubacterium callanderi, Eubacterium limosum, Blautia producta,* and *Clostridiaceae bacterium END-2* are able to demethylate lignans.[8,61,63]

Several organisms have been shown to dehydroxylate and dehydrogenate lignans, in some cases enantioselectively. *Eggerthella lenta, Clostridium scindens, and Lactonifactor longoviformis* can dehydroxylate lignans.[61] *Eubacterium* sp. ARC-2 is capable of dehydroxylating (-) DHEND and (+) DHENL. *Eggerthella* sp. SDG-2 can dehydroxylate (-) DHENL.[64] *Lactonifactor longoviformis* is capable of dehydrogenating lignans.[60] *Clostridiaceae bacterium* END-2 can dehydrogenate (+) END, while *Ruminococcus* sp. END-1 can dehydrogenate (-) END.[65] *Enterococcus faecalis* and *Eggerthella lenta* can reduce lignans.[8,15] It is probable that there are more undiscovered microbes that play a role in lignan metabolism.[4]

1.6   H$_2$ Consumers in the Human Gut

In the human gut, the main form of microbial energy acquisition is fermentation. Poly- and oligosaccharides are broken down by endogenous and microbial enzymes, and resulting monosaccharides are fermented, generating short chain fatty acids (SCFAs) – including acetate, propionate, and butyrate – CO_2 , and H_2. H_2 is primarily generated because of the need to reoxidize reduced pyridine and flavin nucleotides. Accumulation of H_2 would lead to a partial pressure that would thermodynamically inhibit this reoxidation. Thus, it is crucial that H_2-consuming microbes are part of the intestinal community. These organisms fall into three categories based on their form of H_2 metabolism: acetogens, methanogens, and sulfate-reducing bacteria.

Acetogens utilize the Wood-Ljungdahl pathway (WLP) to reduce CO_2 to acetate. This is a multi-reaction process by which a methyl group is transferred from enzyme subunit to subunit. Overall, with H_2 as an electron donor, the total standard free energy change ($\Delta G°'$) for the

reduction of CO_2 to acetate is -95 kJ per mol. No net ATP is generated, but this process is coupled with chemiosmotic ion gradient-driven phosphorylation and generates 1-2 ATP overall. This yield is estimated to be lower at physiological conditions.[66]

Methanogens can also utilize the WLP to synthesize acetyl-CoA or methane (CH_4). The synthesis of CH_4 involves the first half of the WLP, from which the methyl group is transferred to a different subset of enzymes. Overall, the synthesis of CH_4 from CO_2 and H_2 has a $\Delta G^{\circ\prime}$ of -135.6 kJ per mol[67] and is estimated to generate 1 ATP per mol of methane.[68]

Sulfate-reducing bacteria (SRB) reduce sulfate to sulfide for energy conservation. The SRB that live in the human gut generally oxidize H_2, but can use other organic compounds as well.[69] Overall, with H_2 as an electron donor, the $\Delta G^{\circ\prime}$ for the reduction of sulfate to sulfide is -152.2 kJ per mol.[68] This process is estimated to generate 1 ATP per molecule of sulfate reduced, and even less when energy requirement for uptake of sulfate is taken into account.[70]

The extent to which each form of metabolism exists in the human intestine is unclear.[69] Results of some studies suggest that methanogenesis and sulfate reduction dominate because they are thermodynamically more favorable;[71] however, one metagenomic study of human stool found acetogenesis to be the most prevalent form of microbial H_2 utilization in the gut.[72]

1.7   THE POTENTIAL ROLE OF ACETOGENS IN LIGNAN METABOLISM

Acetogens use acetyl-CoA, generated from the WLP, for biomass or ATP generation.[66,73] Many acetogens belong to the genus *Clostridium*, but the group is phylogenetically diverse. It is likely that many bacteria have yet to be characterized as acetogens.[72]

Interestingly, lignan demethylation, dehydroxylation, and dehydrogenation appear to be carried out by bacteria that have been characterized as, or are phylogenetically related to, acetogens. *B. methylotrophicum, B. producta, E. limnosum* are all documented acetogens.[73-75] *Ruminococcus sp. END-1* shares a 98% 16S rRNA gene sequence similarity to the lignan-metabolizing *B. producta* strain.[64] *L. longoviformis* and *Clostridiaceae* bacterium END-2 are highly related to acetogenic species in *Clostridium* cluster XIVa, including *B. producta*.[60] *Eubacterium sp*. ARC-2 shares a 99% 16S rRNA gene sequence similarity to *E. limnosum*, and 98% similarity to *B. methylotrophicum*.[76] *E. lenta* possesses a formyltetrahydrofolate synthetase (FTHFS) that has a 93% similarity to acetogenic FTHFS. FTHFS PCR amplification is typically used to assess presence of acetogens.[77] *Eggerthella sp*. SDG-2 is closely related to *E. lenta*.[64] *E.*

callanderi is in *Clostridium* Cluster XV, and most other species in this cluster are acetogens.[78] The genus *Moryella* forms a sub-lineage with *Blautia coccoides*, an acetogen, in *Clostridium* cluster XIVa and produces acetate *in vitro*.[79] As its name suggests, the genus *Acetoanaerobium* contains at least three known acetogens including *A. ruminis, A. noterae,* and *A. romashkovii*[73]

1.8   THE POTENTIAL ROLE OF METHANOGENS IN LIGNAN METABOLISM

To date, no archaea have been identified as capable of lignan metabolism; however, their methods of energy acquisition are highly similar to those of acetogens. Archaea utilize the WLP, many of the same substrates, and highly similar enzymes.[67,73]

Approximately 50% of the human population possesses intestinal archaea, including *Methanobrevibacter smithii, Methanosphaera stadtmanae, Candidatus Methanomethylophilus alvus, Candidatus Methanomassiliicoccus intestinalis* and *Methanomassiliicoccus luminyensis.* Another order of archaea (Mx-lineage) that contains species closely related to the Rumen Cluster C has also recently been identified.[80] *M. smithii* is both the most abundant and prevalent archaea in the human gut, followed by *M. stadtmanae*.[81] The remaining three identified species, belonging to the order *Methanomassiliicoccales*, appear to be less prevalent.[82]

All archaea found in the human gut thus-far are methanogens,[82] organisms for which methanogenesis is their sole energy-yielding metabolic process. There are two primary forms of methanogenesis that occur in the human gut, both H_2-dependent.[82] Hydrogenotrophic methanogenesis involves using H_2 to reduce CO_2 to methane through the WLP[83] and is utilized by *M. smithii*.[82] The main form of methylotrophic methanogenesis in the human gut involves demethylation of a variety of compounds – including methanol and methylamines – using H_2 to reduce the methyl group to CO_2.[84] This pathway is utilized by *M. stadtmanae* and species of the order *Methanomassiliicoccales*.[81,82,85,86]

1.9   MICROBIAL ENZYMES IN LIGNAN METABOLISM

Deglycosylation is the only step of lignan metabolism that has been connected to a specific bacterial enzyme.[62] Identities of other enzymes involved in lignan degradation can only be hypothesized based on the reactions. Below I use two different number classification schemes: 1) Enzyme Commission (EC) numbers refer to groups of enzymes that carry out the same chemical

reaction,[87] 2) KEGG Orthology gene families (KOs) refer to groups of genes that carry out similar biochemical functions.[88]

1.9.1   *O-linked Deglycosylation*

Microbial β-glucosidases (BG)[EC 3.2.1.21] can carry out lignan deglycosylation.[62] Bacteria harboring BG appear to be quite prevalent in the human gut, on the order of 10^{10} cells/g stool.[89] A wide range of genera from human stool, including *Bacteroides, Bifidobacteria, Butyrivibrio, Coprococcus, Eubacterium, Roseburia,* and *Ruminococcus,* have been shown to possess BG activity.[90] Many of these genera contain known acetogens.[73] BuBGL8, coded by the *bgl8* gene from *Bacteroides uniformis* ZL1, was shown to be a BG capable of deglycosylating SDG.[91]

1.9.2   *O-linked Demethylation*

Acetogens, methanogens, and sulfate-reducing bacteria harbor highly homologous O-demethylases that utilize corrinoid proteins to cleave methyl groups from methoxy groups. A common assay used to select for acetogens is to culture on methoxylated aromatics such as vanillate.[92] Acetogenic O-demethylases cleave off oxygen-bound methyl groups and funnel them into the WLP. *Moorella thermoacetica*, the model acetogenic bacteria, is capable of metabolizing over 20 different types of methoxylated aromatic compounds.[93] Lignan demethyating *E. limnosum, B. methylotrophicum,* and *B. producta* have been shown to O-demethylate several different methoxylated compounds.[73,94] Because the conversion of SECO to DHEND is a methoxylated aromatic O-demethylation, it is quite plausible that it could be catalyzed by a wide range of acetogens.

Acetogens harbor different types of similarly related O-demethylases. *M. thermoacetica* possesses a demethylase composed of 3 subunits coded by MtvA, -B and -C (K00548).[93] There have also been several Mtv homologs identified in the genome of *M. thermoacetica*.[95] The four-subunit O-demethylase possessed by *A. dehalogenans* is composed of an activating enzyme, a corrinoid protein (K16179), and two methyl transfer proteins coded by *odm*A (K14082) and *odm*B (K00548). [96] *Holophaga foetida* contains several genes belonging to these KOs as well.[97] Its demethylating activity has been shown to have a broad range of substrate-specificity.[72]

Methanogens possess demethylases that are highly similar to the acetogenic enzymes coded by mtvABC.[72,96] All identified human gut methanogens have the gene that encodes methyl-

transferase *mtaABC* (K14080, K04480, K14081) which performs demethylation of methanol.[82] This methyl-transferase may demethylate SECO as well.

Sulfate-reducing bacteria can demethylate methoxylated aromatics. *Desulfitobacterium hafniense,* which has been shown to perform these reactions, appears to possess many putative O-demethylase encoding operons. The proteins of one of these operons, Dhaf_4610, Dhaf_461, and Dhaf_4612, have been characterized and express high similarity to O-demethylase proteins of *A. dehalogens.*[98]

1.9.3   *Dehydroxylation*

Bile acids are metabolized by human gut microbes in the intestinal lumen. Interestingly, microbial bile acid metabolism is highly similar to lignan metabolism, in that bile acids are aromatic compounds that are first deconjugated, and subsequently dehydroxylated and dehydrogenated.

Both lignans and bile acids need to be deconjugated before they can be dehydroxylated.[13,99] Bile acid dehydroxylation is an energy-yielding process by which bile acids act as electron acceptors.[100] Most bacteria capable of 7α-dehydroxylating bile acids are highly related and fall within *Clostridium* cluster XIVa, a phylogenic group that contains many acetogens.[101] 7α-dehydroxylating species include *C. scindens, C. hiranonis, C. hylemonae, C. sordellii, B. thetaiotaomicron,* and certain species in the genus *Eubacterium.*[100,102-104] The bile acid inducible (*bai*) genes (K15868, K15869, K15870, K15871, K15872, K15873, K15874) code for enzymes involved in this process.[102-104] *Eggerthella lenta* possesses corticosteroid 21-dehydroxylase [EC 1.14.99.10], involved in 21-dehydroxylating biliary steroids.[105] An *E. lenta*-like strain was also identified as having 7α-dehydroxylating ability.[106] Since *C. scindens*, *Eubacterium* spp.*,* and *Eggerthella spp*. have been shown to dehydroxylate SECO, it has been suggested that bile acid dehydroxylating enzymes may be involved in the dehydroxylation of lignans.[4]

1.9.4   *Dehydrogenation*

Organisms that catalyze bile acid dehydrogenation may benefit from the reducing equivalents generated from the process.[99] Bile acid 3-,7-, 12-, 20- and 21- hydroxyl groups are dehydrogenated by microbial hydroxysteroid dehydrogenases (HSDH) (EC 1.1.1.357, EC 1.1.1.51, EC 1.1.1.159, EC 1.1.1.201, EC 1.1.1.176, EC 1.1.1.238).[99,107] The taxonomic

distribution of HSDH genes is more widespread than the *bai* genes, but the acetogen *B. productus* does possess three types of HSDH. Other organisms closely related to known acetogens, including *Eubacterium aerofaciens*, *Ruminococcus gnavus* N53, *Clostridium absonum, Clostridium sordelii, Clostridium scindens,* and *Clostridium baratii,* can dehydrogenate bile acids.[108,109,110] Since the conversion of END to ENL requires dehydrogenation of aromatic diols, and all three of the lignan dehydrogenating strains identified – *Clostridium* sp. END-2, *Ruminococcus sp. END-1*, and *L. longiformis*– are highly related to *B. producta*, HSDH may be involved.

Another instance of aromatic hydroxyl dehydrogenation is the conversion of estrone to estradiol. Several intestinal bacteria have been identified that can carry this out including *Akaligenes faecalis, Pseudomonas aeruginosa, Staphylococcus aureus*, and *Streptococcus faecalis*.[111] These bacteria may harbor enzymes capable of lignan dehydrogenation.

1.10   AIM, HYPOTHESIS, AND IMPLICATIONS

For the current study we incubated human stool samples anaerobically with a mineral salts media (TCAP2) and SDG in an *in vitro* batch system. We hypothesized that incubation of stool with media containing SDG would enrich for lignan-metabolizing microbes and corresponding genes. We assessed the associated metabolites of SDG (SECO, ENL, and END) and presence of microbial taxa and genes. Through this, we were able to taxonomically and functionally describe a community that metabolizes lignans. Identification of these microbes and enzymes will help characterize an individual's enterolignan-producing capability. This information has implications for our understanding of cancer and cardiovascular disease susceptibility, and may help in development of dietary interventions and preventative probiotics.

Chapter 2. METHODS

2.1 SUBJECTS

Subjects 6105, 6127, 6128, 6129, 6171, 6178, and 6188 from the FlaxFX study (ClinicalTrials.gov identifier: NCT01619020) provided seven of the eight stool samples used for our experiments. Men and women were eligible to participate if they were 20-45 years, healthy non-smokers with a low vegetable and fiber intake, consented to be in the trial, and consented to provide a fresh stool sample. Individuals were excluded if they had been prescribed antibiotics within the last 3 months, currently taking prescription or over-the-counter medications (including oral contraceptives), consumed more than two alcoholic drinks per day, pregnant or lactating, diagnosed with a chronic medical illness or gastrointestinal disorder, or had experienced a recent weight change of 4.5 kg or greater. The eighth stool sample was derived from the same individual (subject 5100) who provided samples used for methods development of our *in vitro* platform. Not all samples were used for all analyses (Table 1).

2.2 *IN VITRO* INCUBATION

2.2.1 *Media*

TCAP2 media (Table 2) was prepared by adding heat-stable ingredients to de-ionized water, adjusting to a pH of 7.3, and autoclaving for 20 minutes at 121 °C. After cooling, resasurine, vitamin solution, trace element solution, glucose, and cysteine were added by filter-sterilization through a 0.2-micron disc filter. A gas manifold hose was fitted with a 0.2-micron disc filter, followed by a small sterile plastic hose, which was inserted into the bottle of media. The media was sparged with 100% high-purity N_2 (Airgas, Radnor Township, PA) and heated to 70 °C in a water bath until the purple color of resasurine, an anaerobic indicator, disappeared (approximately 20 - 90 minutes). The bottle was capped and immediately placed in a Bactron Anaerobic Chamber that contained a gas mixture of 80% N_2, 15% CO_2, and 5% H_2.

2.2.2   *Incubation Protocol*

Each participant provided a fresh stool sample for *in vitro* analysis. The stool was collected in a plastic tub, and within 30 minutes, brought into a Bactron Anaerobic Chamber. Inside the chamber the stool was weighed and combined with anaerobic TCAP2 media to a final concentration of 1.66% weight:volume. The stool was broken apart with a spatula and further mixed with a homogenizer. The fecal suspension was then filtered through sterile cheesecloth and electronically pipetted into 50 mL glass serum bottles (20 mL aliquots). These bottles were inoculated with lignan substrate dissolved in 100% methanol (Thermo Fisher Scientific, Waltham, MA) to a final concentration of 6.55 µM SDG, 6.54 µM ENL, 6.45 µM END, or left un-spiked (a blank "control").

The bottles were capped with rubber stoppers and brought outside the chamber, where the tops were fitted with aluminum seals and sterilized with 100% ethanol. A gas manifold hose was fitted with a 0.2 micron disc filter followed by a sterile Luer-lok needle, which was inserted into the rubber stopper at the top of the serum bottle. Another Luer-lok needle was inserted into the rubber stopper as a vent and the serum bottle headspace was replaced with 100% high-purity N_2 for 10 minutes at the rate of 0.5-1.0 L/min. Bottles were placed in a rotating incubator (C24 Incubator, New Brunswick Scientific, Enfield, CT) and incubated at 37°C/300 rpm. Samples from 6105, 6127, 6128, 6129, and 5100 were incubated for 7 days. Samples from 6171, 6178, and 6188 were incubated for 6 days.

2.2.3   *Sample Collection*

Daily aliquots of fecal suspensions from individual incubating serum bottles were taken at approximately the same time every day using anaerobic technique, and stored at -20 °C in aliquots of 500 µL/day for lignan analysis and 100 µL/day for bacterial enumeration. Samples for DNA analysis were taken on Day 1 and the final day of the incubation (Day 6 or 7), and stored at -80 °C. The Day 1 sample was two pea-sized scoops of fresh stool dispersed in 5 mL sterile RNA*later* (Thermo Fisher Scientific, Waltham, MA) collected in the anaerobic chamber during the previously described procedure. The final-day sample was obtained from pelleted SDG-incubated fecal suspensions. To collect the pellet, fecal suspensions from serum bottles were combined in a sterile 50 mL centrifuge tube and spun down in a Beckman Coulter

Centrifuge. Suspensions from 6105, 6127, 6128, 6129, and 5100 were centrifuged at 6000 rpm/20°C for 10 minutes, supernatant was discarded, and the pellet was re-suspended with 5 mL of sterile RNA*later* for storage. Suspensions from 6171, 6178, and 6188 were centrifuged at 14,000 rpm/4°C for 10 minutes. Supernatant was discarded, pellets were re-suspended in TCAP2, and tubes were re-centrifuged using the same conditions. Supernatant was discarded and the pellet was re-suspended with 5 mL of sterile RNA*later* for storage.

2.3 BACTERIAL ENUMERATION VIA QPCR

Eubacterial 16S rRNA gene copy number concentration in SDG-inoculated fecal suspensions was estimated via quantitative polymerase chain reaction (qPCR) (7900HT Fast Real-Time PCR System by Applied Biosystems, Carlsbad, CA). Universal primers 330F (5'-ACT CCT ACG GGA GGC AGC AGT-3') and 530R (5'-GTA TTA CCG CGG CTG CTG GCAC-3') were used to amplify the 16S rRNA gene[112,113] with SYBR Green (Invitrogen, Carlsbad, CA), a total reaction volume of 20 μL, and cycling conditions: 50 °C for 2 min; 95 °C for 10 min; [95 °C for 30 sec; 58 °C for 30 sec; 72 °C for 1 min] x40; 95 °C for 15 sec; 60 °C for 15 sec; and 95 °C for 15 sec. We extended the first 95 °C step to 10 minutes in order to ensure cell lysis.

2.4 LIGNAN METABOLITE MEASUREMENT VIA GCMS

For lignan metabolite analysis, SECO, END, and ENL concentrations from fecal suspensions were measured via gas chromatography mass spectrometry (GCMS) using modified methods from Frankenfeld *et al*[114] as described elsewhere.[14] Time points from Days 1 through 6 were analyzed for some participants, while time points from only Days 1 and 6 were analyzed for other participants (Table 1). Day 7 was additionally analyzed for participant 6105. It was decided to only use measured metabolite concentrations through Day 6 for analyses because: a) metabolite levels did not seem to change when measured through Day 7 (Figure 1) and b) ENL appeared to be the final metabolite in the pathway.

For each time point, measured metabolite concentration was divided by inoculated substrate concentration to calculate percent conversion. In a separate analysis for samples that had been analyzed by qPCR (Table 1), moles of metabolite was divided by qPCR-estimated copy number at each time point to normalize for bacterial biomass.

2.5 DNA EXTRACTION AND SEQUENCING

DNA was extracted from Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension with the Qiagen Stool DNA Mini Kit (Qiagen, Hilden, Germany)[115] using a modified protocol of the manufacturer's instructions. DNA was sent to Research and Testing (Shallowater, TX) for 16S rRNA gene and whole metagenome shotgun sequencing. The 12 samples (from 6 participants) that were V4 16S rRNA gene-sequenced and used for our analyses [6x Day 1 initial stool and 6x final-day pelleted SDG-incubated fecal suspension (Table 1)] were sequenced with 19 samples that were not used for analyses in this paper. There were 6 samples (from 3 participants) [3x Day 1 initial stool and 3x final-day pelleted SDG-incubated fecal suspension (Table1)] that were sent for metagenomic sequencing and used in our analyses.

The V4 region of the 16S rRNA gene was paired-end (2x300) sequenced with the target of 30,000 sequences-per-sample. Primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3' + barcode) and 806R (5'-GGACTACVSGGGTATCTAAT-3'),[116] were used in a 30 cycle PCR with HotStarTaq Plus Master Mix Kit (Qiagen, USA) to amplify the V4 region of the 16S rRNA gene with the cycling conditions: 94°C for 3 min; [94°C for 30 sec] x28; 53°C for 40 sec; 72°C for 1 min; 72°C for 5 min. PCR products were run in 2% agarose gel to ensure high quality amplification. Samples were pooled in equal ratios based on DNA concentration and molecular weight. Pooled samples were purified with Ampure XP beads (Beckman Coulter, Brea, CA). Libraries were prepared according to the Illumina TruSeq DNA library preparation protocol.[117] Illumina MiSeq instruments and reagents were used following manufacturer's instructions.

Samples were sequenced using whole metagenome shotgun sequencing[118] with the target of 6 million sequences-per-sample. Sample concentrations were measured using the Qubit® dsDNA HS Assay Kit (Life Technologies, Carlsbad, CA). Libraries were prepared using the Nextera DNA Sample preparation kit (Illumina, San Diego, CA) following the manufacturer's instructions. Specifically, samples were diluted to a concentration of 2.5 ng/uL. An aliquot containing 50 ng of DNA was fragmented and tagged using the Nextera Enzyme Mix, which contains a transposase and transposon ends bound to sequencing primer sites. A 5-cycle PCR was carried out, in which unique index was added to the samples. The final concentration and average size of the library was measured using the Qubit® dsDNA HS Assay Kit and the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), respectively. The library (12.5 pM)

was paired-end (2x300) sequenced in MiSeq (Illumina) with the 600 Cycles v3 Reagent Kit (Illumina), and following the manufacturer's instructions.

2.6   TAXONOMIC AND FUNCTIONAL ANALYSES

2.6.1   *Processing and Taxonomic Annotation of 16S rRNA Gene Amplicon Sequences with QIIME*

16S rRNA gene sequences were processed through an in-house QIIME[116] pipeline. Paired-end reads were joined, removed if <200 basepairs (bp) or >700 bp, and truncated with QIIME scripts and a quality score sliding window of 50 bp. They were then clustered into operational taxonomic units (OTUs) at 97% similarity and classified using MOTHUR[119] and the SILVA database(www.arb-silva.de). For community structure analyses, annotated sequences were grouped into phylum and genus-level phylotypes.[120] The final output was a data matrix of metagenomic sample by organism abundance.

2.6.2   *Processing of Metagenomic Sequences with MG-RAST*

Unassembled metagenomic sequences were annotated by the Metagenomics Rapid Annotation using Subsystem Technology (MG-RAST) pipeline.[121] Briefly, paired-end reads were joined, and low quality regions were trimmed from FASTQ data using SolexaQA[122] so that sequences contained at most 5 bases below a Phred quality score of 15. The data were screened, using Bowtie,[123] to remove reads that mapped to the human genome. Genes were called with FragGeneScan.[124]

2.6.3   *Taxonomic Annotation of Metagenomic Protein Sequences with MG-RAST*

Proteins were clustered at 90% similarity with uclust[125] in QIIME.[116] sBLAT[126] was used to map the longest sequence in each cluster to the M5nr database.[127] Phylum-level taxonomy was assigned to M5nr-annotated reads with an e-value less than 1e-10, an identity greater than 80%, and an alignment length greater than 75 base-pairs[128] using the Representative Hit tool. Organism abundance tables were collated with in-house R scripts to yield a data matrix of metagenomic sample by organism abundance.

2.6.4   *Taxonomic Annotation of Metagenomic Protein Sequences with MetaPhlAn*

Reads that had been processed through the trimming step of the MG-RAST pipeline were uploaded to the Huttenhower Galaxy server and classified with MetaPhlAn[129] using the "very sensitive (global)" setting of BowTie2[130] as recommended by MetaPhlAn developers for human-associated microbiomes. The global option utilizes end-to-end alignment, which uses all nucleotides in a query sequence for alignment. The "very sensitive" option maximizes sensitivity and accuracy of the alignment with preset parameters: -D 20 -R 3 -N 0 -L 20 -i S,1,0.50 as defined elsewhere.[130]

Reads were mapped to a collection of clade-specific marker genes derived from the IMG database,[131] providing species-level taxonomic assignment. The number of marker genes per clade varies, but the number of hits per clade was normalized by the sum of nucleotides in all marker genes corresponding to that clade. Due to a problem with the server, we were unable to upload the sample from 6128 Day 1, which has been left out of our MetaPhlAn taxonomic analysis. The final output was a data matrix of metagenomic sample by organism abundance.

2.6.5   *Functional Annotation of Metagenomic Protein Sequences with HUMAnN*

Reads that had been processed through the screening step of the MG-RAST pipeline were annotated by DIAMOND[132] using KEGG database v56. Annotations with an e-value < 1e-5 were analyzed using HUMAnN.[133] This program normalizes abundance for each KO and then assigns KOs to KEGG modules and pathways. Modules and pathways are groups of KOs that work together in a biochemical pathway.[88] Modules generally contain fewer KOs and involve fewer reactions than pathways.

During HUMAnN processing, abundance for each KO was weighted by the e-value of its hits, normalized by average sequence length of its gene family, and assigned to KEGG modules and pathways using MinPath.[134] The mapping of KOs to modules and pathways resulted in the output of (KO, module/pathway) pairs. Modules/pathways that were in clear disagreement with a rough taxonomy of reads in a sample were eliminated. Specifically, a very approximate organism abundance was calculated. Each organism abundance was derived by summing all normalized, weighted hits for that organism. "Expected abundance" for each module/pathway was then calculated by summing the abundances for each organism that reads in that module/pathway had

mapped to. (KO, module/pathway) pairs for KOs that mapped to two or more modules/pathways *and* corresponded to a module/pathway that possessed an abundance below the "expected abundance" were removed.

The abundance of each organism-specific KO was further normalized by organism-specific gene family copy number. Specifically, KOs with an abundance 1.5 inter-quartile ranges below the median number of hits across all the genes in the module/pathway were replaced with this median. Finally, module/pathway abundance was calculated as the mean of the upper half of its component KO abundances. For individual gene analysis, the "01b-hit-keg-cat" file was used. For module and pathway analysis, the "04b-hit-keg-mpm-cop-nul-nve-nve" and "04b-hit-keg-mpt-cop-nul-nve-nve" files were used, respectively. The final output was a data matrix of metagenomic sample by KO, module, or pathway abundance.

2.6.6 *Analyses of Taxonomic and Functional Output*

All abundance tables that were not already in the form of relative abundance expressed as a percent, were converted. Relative abundance was calculated by dividing specific organism, KO, module, or pathway abundance by the sum of all organism, KO, module, or pathway (respectively) abundances in the sample and multiplying by 100. Shannon's index[135] was used to for assessment of taxonomic alpha diversity. For 16S rRNA gene amplicon sequences, this was carried out as part of our in-house QIIME pipeline with OTUs rarified to 30,766 sequences per sample. For metagenomic sequences, this was carried out using species-level assignments from MetaPhlAn with the VEGAN R package.[136] Organisms and KOs identified or hypothesized to be involved in lignan metabolism (Table 3 and 4) were searched for using in-house R scripts.

2.7 DECISION TO USE METAPHLAN INSTEAD OF MG-RAST

Classification of protein sequences into phyla was carried out using both MG-RAST and MetaPhlAn (Figure 2). Notably, MG-RAST classified sequences into many phyla that had never been identified in human stool. Relative abundances of dominant phyla were substantially different. For example abundance of Bacteriodetes was much higher, and Firmicutes and Euryarcheaota much lower, in MG-RAST output.

One sample, 6128 Day 7, was sequenced using both V4 16S rRNA gene amplicon and metagenomic shotgun sequencing. When classification of sequences into dominant phyla was

compared between our in-house V4 16S rRNA gene pipeline, MG-RAST, and MetaPhlAn, MGRAST was the outlier (Figure 3). Because of this, and because MG-RAST output contained spurious phyla, MetaPhlAn was chosen for further taxonomic analysis of metagenomic data.

In an additional analysis (data not shown), for which we used the MG-RAST Representative Hit tool to classify 16S rRNA gene sequences from the 6128 Day 7 shotgun-sequenced metagenomic sample (Database: M5rna,[127] e-value <1e-10, identity >97%, alignment length >150 bp), relative abundances of major phyla were similar to the MG-RAST protein output.

2.8   DECISION TO USE SUMMED KOs INSTEAD OF HUMANN MODULE OUTPUT

Several KEGG pathways hypothesized to be involved in lignan metabolism were searched for in the HUMAnN pathway output. Of those that were found (Figure 21a), methane metabolism (ko00680) was present at high levels in all samples. Subsequently, KEGG modules corresponding to ko00680 were searched for in HUMAnN module output (Figure 4). Methane oxidation (M00174) and Coenzyme M biosynthesis (M00358) were the only modules found. This did not make sense according to the V4 16S rRNA gene and metagenomic taxonomic classification. For example, methanogens were highly abundant in Day 7 samples from 6128 and 5100, but the modules for methanogenesis were nonexistent as annotated by HUMAnN.

Furthermore, M00174 was found at very low levels (<0.000025%) in the 6105 samples, and did not seem to explain the relatively high abundance (~0.6%) of ko00680 that appeared in the pathway output. Component KOs of M00174 were then searched for in the HUMAnN KO output and none were found. It was concluded that this must have been a HUMAnN error that occurred during assignment of KOs to modules. It was thus decided to sum component KO relative abundances (from the "01b-hit-keg-cat" file) to yield module and pathway relative abundances. We used this method along with HUMAnN pathway output to assess pathway abundance, but did not use HUMAnN module output for analyses. Modules unique to pathway ko00680 were investigated to further understand the components contributing to its abundance.

Chapter 3. RESULTS

3.1 BACTERIAL ENUMERATION VIA QPCR

SDG incubations from the four participants assessed exhibited a similar pattern of eubacterial cell growth over time, although 16S rRNA gene copy number per µl for 6128 and 5100 changed less over time than for 6105 and 6129. 16S rRNA gene copy number per µl peaked on Day 2 or 3 and ranged between 2,583,751 and 20,342,383 (Figure 5). For most serum bottles, copy number declined after Day 3. For some, especially one bottle from participant 5100, concentrations began to rise from Day 5 to Day 7.

3.2 LIGNAN METABOLITE MEASUREMENTS VIA GCMS

Lignan concentration data from serum bottles inoculated with END, ENL, and left un-inoculated were used as controls to isolate END to ENL conversion, any metabolism of ENL, and presence of any lignans in the original stool sample, respectively. Stool from four participants converted nearly all inoculated END to ENL (86 – 92%) *in vitro* (Figure 6). 6128 converted roughly 30% END to ENL almost immediately, whereas the rest of the END-incubated samples converted no substrate by the first time point. Most of the inoculated ENL was recovered at Day 6 (minimum 83 %), and no other metabolites were measured in this treatment. Transient low levels of lignans were measured in the "blank" controls. The maximum of these measurements was 0.53 µM ENL (8% of the SDG concentration added to SDG-inoculated bottles).

Stool of all 8 participants converted the majority of SDG to ENL by Day 6 (Figure 7). The average and standard deviation of SDG to ENL percent conversion from all SDG-inoculated serum bottles was 88.4% ± 6.9. For 6105 and 6129, the time course of measured metabolites reflected several steps of the SDG conversion pathway. SDG was metabolized almost immediately to SECO, which was metabolized to END and subsequently ENL. 6105 produced by far the most END. Negligible END was measured for 5100 and 6128, and negligible SECO was measured for 5100.

3.3   qPCR-normalized GCMS measurements

Normalizing measured metabolites by eubacterial 16S rRNA gene copy number changed the rate of conversion moderately (Figure 8). Most notably, ENL:copy number ratio for 6105 rose steadily and did not plateau, when compared to un-adjusted measurements, and to adjusted measurements of incubations from other participants. The maximum value for this measurement was also substantially higher than adjusted measurements of incubations from other participants. Since all incubations had similarly high levels of percent conversion to ENL, this was largely due to declining copy number during the latter half of the 6105 incubation. ENL:copy number ratio plateaued for incubations from other participants, however for both 6129 and 5100 there was a slight decrease in this value from Day 5 to Day 6. This was mainly due to increase in copy number during this time period for these incubations. Another notable effect of normalization was a high initial SECO:copy number ratio for 6129.

3.4   V4 16S rRNA Gene Taxonomic Analysis (n = 6)

3.4.1   *Sequence Processing*

For the 31 samples sequenced using Illumina MiSeq paired-end sequencing of the V4 region, 4.2 million raw sequences with an average length of 299 bp were processed. Of these sequences, 1.9 million with the average length of 272 bp passed quality control. The twelve samples (out of 31) that were used for analyses, all possessed over 38,000 quality-processed sequences (Table 5).

3.4.2   *Alpha diversity, Phyla and Genera*

OTU-level alpha diversity, for the most part, did not change consistently with a ~1 week incubation; however, there were moderate decreases in diversity for incubations 6128 and 6129 when comparing Day 1 initial stool to Day 7 pelleted fecal suspension (Figure 9). Taxonomy was first assessed at the level of phylum (Figure 10). Bacteriodetes grew in all incubations, with the exception of 6128. Firmicutes decreased in all incubations. Euryarchaeota grew substantially in 6127, 6128, 6171. Incubations from the rest of the participants demonstrated growth of Euryarchaeota, but at very low levels (Figure 11). None of the samples contained

Verrucomicrobia with a relative abundance over 1.5%, although in incubations from 6127 and 6188, this phylum increased notably.

Taxonomy was then assessed at the level of genus. Growth of Bacteroidetes was primarily accounted for by an expansion of *Bacteroides*, and to a lesser degree, *Parabacteroides* and *Alistipes* (Figure 12). Decrease in Firmicutes was not attributable to one particular genus (Figure 13). Growth of Euryarchaeota was almost entirely due to *Methanobrevibacter* (Figure 14). Verrucomicrobia was primarily composed of *Akkermansia*, with the exception of the 6127 Day 7 sample, which possessed an uncultured bacterium of the order vadinHA64 (Figure 15).

3.4.3   *Organisms Identified or Hypothesized to be Involved in Lignan Metabolism*

Subsequently, genera that were identified or hypothesized to be involved in lignan metabolism were investigated for *in vitro* enrichment (Figure 16, Table 6). The only organism out of these genera that grew substantially in all incubations was *Bacteroides*. A sulfate-reducing organism of the family Desufovibrionaceae grew markedly in the incubation from 6129, and to a lesser degree in the incubation from 6188. *Bifidobacterium* was also present at high levels – especially in samples from 6127 – but did not grow consistently in all incubations. An organism of the order Eubacteriaceae grew substantially in the 6128 incubation. Other genera of interest were present at lower abundance. Many genera searched for were not present or detected in our samples.

Most samples contained at least one genus identified or hypothesized to be involved in each step of SDG to ENL metabolism. Lignan deglycosylating, demethylating, and dehydroxylating organisms were found at higher abundance than dehydrogenating organisms. As a group, deglycosylating organisms grew *in vitro*. As a group, demethylating organisms grew in all incubations except 6178. The 6128 Day 7 sample had particularly high levels of demethylating and dehydroxylating organisms, which was heavily accounted for by the order Eubacteriaceae. Dehyroxylating organisms grew in the 6127, 6128, and 6171 incubations.

3.5   METAGENOMIC TAXONOMIC ANALYSIS (N=3)

3.5.1   *Sequence Processing*

For the 6 samples sequenced using Illumina MiSeq paired-end metagenomic shotgun sequencing, 35.3 million raw sequences were processed, with an average length of 301 bp. After quality control, 28.7 million sequences with the average length of 262 bp were used for analyses.

3.5.2   *Alpha diversity, Top 10 Phyla and Genera*

The Day 7 pelleted SDG-incubated fecal suspension had a marginally lower alpha diversity than Day 1 initial stool at the species-level for incubations from 6105 and 5100 (Figure 17). Phyla Bacteroidetes and Firmicutes dominated initial stool samples (Figure 18). Firmicutes relative abundance decreased with *in vitro* incubation. Euryarchaeota relative abundance increased substantially in the 5100 incubation and was dominant in 6128 Day 7. At the genus-level, this was largely due to the presence of *Methanobrevibacter* in 6128 and *Methanobrevibacter* and *Methanosphaera* in 5100. Euryarchaeota was measured at very low levels in 6105 Day 7, but abundance was negligible compared to Day 7 from the other two participants. Verrucomicrobia and Bacteroidetes grew substantially with stool incubated from 6105. The genus *Akkermansia* accounted for most Verrucomicrobia abundance. At the species-level, these reads were assigned to the species *Akkermansia muciniphila* (data not shown).

3.5.3   *Organisms Identified or Hypothesized to be Involved in Lignan Metabolism*

Genera identified or hypothesized to be involved in lignan metabolism were further investigated (Figure 19). Several of these genera were present in at levels above 1.5%. As aforementioned, *Methanosphaera* and *Methanobrevibacter* grew substantially in the 5100 incubation, and the latter was present at high levels in 6128 Day 7. *Bifidobacterium* grew in the 6105 incubation and was present in 5100 Day 1. *Ruminococcus* was abundant in all samples except 5100 Day 7. *Eubacterium* generally decreased with incubation, but was highly abundant in 6128 Day 7. *Bacteroides* was highly abundant in all samples, although less so in 6128 Day 7.

Several genera investigated were present at levels below 1.5%. *Clostridium* increased substantially in the 6105 incubation, and was also abundant in 6128 Day 7. *Lactobacillus* was

abundant in both samples from 5100. *Eggerthella* grew substantially in the 6105 incubation. *Streptococcus* abundance decreased in the 5100 incubation and was present in 5100 and 6128 Day 7. *Desulfovibrio* was present in 5100 Day 7.

Species identified or hypothesized to be capable of lignan metabolism were further investigated (Figure 20). Most notably, *Eubacterium limnosum* grew substantially with incubation of stool from 5100. It was also present at high levels in the 6128 Day 7 sample. According to the V4 16S rRNA gene data, relative abundance of the family Eubacteriaceae increased by several orders of magnitude (.02 to 23%) from Day 1 to Day 7 for the 6128 incubation. Our V4 16S rRNA gene pipeline was unable to classify this phylotype at a higher resolution than family, but it is likely that it contains the same reads that MetaPhlAn classified as *E. limosum*. *Methanosphaera stadmanae* and *Methanobrevibacter smithii* accounted for the full abundance of their corresponding genera described above. *Ruminococcus bromii* was most abundant in both 6105 samples, but also notably present in 6128 Day 7 and 5100 Day 1. *Ruminococcus lactaris* was present in both 6105 samples and 6128 Day 7.

Several other species identified or hypothesized to be capable of lignan metabolism were present at levels below 1.5%. *Bifidobacterium longum*, *Bacteroides ovatus*, *Eggerthella lenta*, *Bifidobacterium pseudocatenulatum*, *Clostridium leptum, Bacteroides thetaiotaomicron*, and *Escherichia coli* grew substantially in the 6105 incubation. *Eggerthella lenta* and *Bifidobacterium pseudocatenulatum* were undetectable or at negligible levels in Day 1 samples, but present in all measured Day 7 *in vitro* samples. *Bifidobacterium catenulatum* and *Bacteroides fragilis* grew in 5100. *Clostridium scindens, Clostridium leptum, Bacteroides thetaiotaomicron* were present at notable levels in 6128 Day 7. *Desulfovibrio desulfuricans* was found at low levels in 5100 Day 7. Incubated samples from 6128 and 5100 showed increases in highly abundant (>1.5%) species of interest that were not present in the 6105 samples. 6105 showed the largest increase in low abundance (<1.5%) species of interest.

3.6   METAGENOMIC FUNCTIONAL ANALYSIS (N=3)

3.6.1   *Pathways*

Sequences in our samples mapped to a total of 7276 KOs and 231 pathways using HUMAnN. Of the KEGG pathways hypothesized to be involved in lignan metabolism searched for in

HUMAnN pathway output (Figure 21), methane metabolism (ko00680), starch and sucrose metabolism (ko00500), steroid hormone biosynthesis (ko00140), and primary bile acid biosynthesis (ko00120) were present in the samples analyzed. Methane metabolism and starch and sucrose metabolism were the most abundant of the pathways investigated. Methane metabolism was present in all samples and increased with incubation of the samples from participants 5100 and 6128. The 3 most abundant KOs in the Day 1 sample from these two participants and both samples from 6105 were 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (K01834), 6-phosphofructokinase 1 (K00850), and fructose-bisphosphate aldolase, class II (K01624). Starch and sucrose metabolism was present in all samples and decreased marginally with the 7-day incubation. Steroid hormone biosynthesis was present in all samples and increased slightly with the 7-day incubation. Primary bile acid synthesis was present in all samples except 5100 Day 7. This pathway increased with incubation of samples from 6105 and decreased with incubation of samples from 5100 and 6128.

The results of manually summed KOs for these pathways was almost identical, with two exceptions: 1) abundances for methane metabolism and starch and sucrose metabolism were at lower levels, but maintained roughly the same inter-sample proportions; 2) primary bile acid synthesis was present in all samples at roughly the same abundance. Although several KOs were found in the primary bile acid biosynthesis (ko00120) pathway, choloylglycine hydrolase (K01442) and alpha-methylacyl-CoA racemase (K01796) were the most abundant. The former was the only KO detected in the secondary bile acid biosynthesis pathway (ko00121).

3.6.2 *Modules*

Total abundance of all detected KOs in the WLP did not change substantially (Figure 22). However, acetyl-CoA synthase (K14138) and carbon-monoxide dehydrogenase subunits (K00198, K00197, K00194) increased considerably with the 7-day incubation of stool from 5100 and 6128, but decreased less markedly for participant 6105.

Only two KOs – 11beta-hydroxysteroid dehydrogenase (K00071) and cytochrome P450, family 17, subfamily A (K00512) – belonging to the C21-Steroid hormone biosynthesis module (M00109) were detected in our samples. K00071, which was substantially more abundant than K00512, was present in all samples, and increased with incubation of stool from 6105 (Figure 23).

Summed KOs of each module unique to the methane metabolism pathway all increased after the 7-day incubation of stool from 6128 and 5100 (Figure 24). 3 KOs – acetyl-CoA synthetase (K01895), acetate kinase (K00925), and phosphate acetyltransferase (K00625) – from M00357 were found in high abundance in all Day 1 samples and 6105 Day 7.

3.6.3 *KOs Corresponding to Enzymes Hypothesized to Act Directly on Lignans*

KOs corresponding to enzymes hypothesized to act directly on lignans were further investigated (Figure 25). KOs corresponding to beta-glucosidases were present at high levels in all samples, the most abundant being K05349. All beta-glucosidase KO abundances decreased in samples after the 7-day incubation. K01195 corresponding to beta-glucuronidase, was present in all samples. KOs corresponding to acetogen demethylases were present in all samples, and increased substantially with a 7-day incubation of stool from 6128 and 5100. 5-methyltetrahydrofolate-homocysteine methyltransferase (K00548) was the most abundant of these KOs. KOs corresponding to methanogen demethylases were present at high levels in 5100 and 6128 Day 7 samples. KOs corresponding to bile acid dehydrogenases, as a group, increased in 5100 and 6128 incubations, and decreased slightly in the 6105 incubation. 3alpha(or 20beta)-hydroxysteroid dehydrogenase (K00038) and 7-alpha-hydoroxysteroid dehydrogenase (K00076) were the most abundant of this group.

Chapter 4. DISCUSSION

4.1 METABOLITE MEASUREMENTS AND EUBACTERIAL ENUMERATION

The fact that most of the inoculated ENL was recovered at Day 6, and no other metabolites were measured in this treatment, indicates that this is the last metabolite in the process of SDG metabolism. Lack of full recovery could be due to pipetting error or less-than-complete extraction in GCMS preparation, since measured ENL for this treatment remains relatively constant over the 6-day period.

Our incubations performed SDG to ENL conversion to very efficiently compared other studies with similar incubations,[5,62,137] however the amount of SDG we inoculated with was substantially lower. The fecal inoculum concentration we used was chosen – based on our methods development – to maximize SDG to ENL conversion as described elsewhere.[14] The negligible levels of SECO and END measured in incubations 6128 and 5100 could indicate more rapid or alternate routes of metabolism. For example DHEND can be metabolized to DHENL and then to ENL, bypassing END.[14] The occurrence of this pathway is unknown in our experiments, since we did not measure DHENL. The close-to-full conversion to ENL in END-inoculated incubations for these participants shows that their stool is at least capable of converting END to ENL, yet their microbes could prefer an alternate route of metabolism when provided different precursors. The high conversion rate to ENL in the 6128 END-inoculated incubation, could indicate a faster metabolism as an explanation for lack of END observed in the 6128 SDG-inoculated incubation. Conversely, the 5100 incubation did not show this difference in rate.

4.2 TAXONOMIC ANALYSIS

MetaPhlAn appeared to outcompete MG-RAST in accurately producing sample phylotype abundance. There are probably several explanations for this. MetaPhlAn uses clade-specific markers instead of all protein reads. This helps account for gene transfer and misclassification of ubiquitous proteins.[129] By using marker genes and normalizing by nucleotide length, MetaPhlAn eliminates the problem of variable genome size. MG-RAST does nothing to account for this,

which may explain the difference in abundance of sequences classified to Bacteroidetes, Firmicutes, and Euryarchaeota.

Abundance of Bacteriodetes, in fact, has been associated with a larger average genome size (AGS), while abundance of Firmicutes has been associated with a smaller AGS.[138] Thus, failing to correct by genome size would probably yield an inflated Bacteroidetes:Firmicutes ratio as seen with MG-RAST output. More metabolically specialized organisms appear to have smaller genomes than generalist organisms.[138] For example, *Methanobrevibacter smithii* has a much smaller genome than *Bacteroides thetaotamicron*. This could explain the reason Euryarchaeota had a lower abundance according to MG-RAST. Finally, MetaPhlAn uses nucleotides rather than amino acids for classification, which provides more taxonomic information, since several different nucleotide triplets can code for one amino acid.[139]

OTU and species-level alpha diversity did not decrease markedly for most samples after a ~1 week incubation of stool with a mineral salts media. We had expected to observe a decrease in alpha diversity since the growth substrates available *in vitro* were substantially limited and less diverse compared to the colonic biome. The change in species-level alpha diversity is small when compared to a similar study in which human stool was incubated with crude oil and a dispersant.[140] The fact that MetaPhlAn species-level alpha diversity was lower per sample than QIIME OTU-level alpha diversity may indicate that the 16S rRNA gene methods used provide a more diverse representation of a microbial community that the metagenomic methods used.

An increase in Bacteroidetes:Firmicutes ratio after incubation was observed in taxonomic classification of both the V4 16S rRNA gene and metagenomic data. This is similar to results of *in vitro* incubations with polyphenols from tea[141,142] and a red wine/grape juice mix.[142] In both studies, Bacteroidetes:Firmicutes ratio increased *in vitro* when media was supplemented with polyphenols. It is possible that the small amount of SDG in our *in vitro* incubations could drive this effect. Regardless, since a large number of SDG-deglycosylating bacteria fall in the phylum Bacteroidetes, and degylcosylation is necessary for the subsequent steps of SDG to ENL conversion, it follows that enriching for Bacteroidetes might promote SDG to ENL conversion.

Euryarchaeota increased in several of the incubations, but it has yet to be determined whether these microbes are involved in lignan metabolism. It is possible that by consuming hydrogen, they are driving fermentation reactions, and in turn promoting the growth of organisms capable of cleaving and utilizing glucose moieties from lignans.

Verrucomicrobia increased notably in several samples after incubation. According to metagenomic analysis, this phylum was almost primarily accounted for by *A. muciniphila. A. muciniphila* is a gram-negative, strictly anaerobic, mucin degrading bacteria isolated from human feces.[143] Interestingly, dietary administration of polyphenol extracts from green tea,[144] Concord grape,[145] and cranberry[146] increased cecal levels of *A. muciniphila* in mice. Similarly, *in vitro* incubation of human feces with polyphenol extracts from a red wine/grape juice combination or black tea increased relative abundance of *A. muciniphila*.[142]

The mechanism by which polyphenols promote the growth of *A. muciniphila* is still undetermined. It may be due to an antimicrobial effect, which could reduce the viability of competitors and thus allow *A. muciniphila* to thrive.[145,146] This is supported by the fact that the microbe has been shown to dominate in human feces following antibiotic administration.[147] Another possibility is that polyphenols promote production of mucin, which is feeding *A. muciniphila*.[145,146] This conflicts, however, with findings that mucin gene expression was not altered in mice fed Concord grape polyphenols,[145] and would not explain the results observed *in vitro*. A third possibility is that since *A. muciniphila* is strictly anaerobic, and hence sensitive to free oxygen radicals, polyphenols could provide protection by scavenging these radicals.[145] A final option is that polyphenols are a substrate utilized by *A. muciniphila. A. muciniphila* does possess glycosidases[143] which could potentially act on glycosylated polyphenols, but this has yet to be investigated.[146]

Organisms hypothesized to be involved in lignan metabolism, as a group, were not enriched after all incubations. Enrichment and presence of these organisms was sample-specific, with each sample containing some of the organisms of interest. Most samples contained at least one genus identified or hypothesized to be involved in each step of SDG to ENL metabolism. This is in agreement with the fact that SDG metabolism is carried out by a consortia of microbes and that each step can be performed by several different types of organisms.[12] The observation that dehydrogenating organisms were found at lower levels than organisms capable of performing other reactions involved in lignan metabolism is in line with studies that have measured abundance of lignan metabolizing organisms in human stool.[5,18,51] The fact that *Ruminococcus bromii* and *Ruminococcus lactaris* were found in our samples is interesting in light of a very recent study showing an association between the presence of these species in stool from participants supplemented with flaxseed, and ENL production.[148]

Bacteroides prominently increased after most incubations. Though our metagenomic taxonomic classification of incubated samples from three participants did not reveal that this genus was primarily accounted for by species identified in lignan metabolism, it is possible that other species in this genus are involved as well. If this was the case, an increase in their abundance could be promoting SDG deglycosylation and ultimately helping to drive ENL production.

4.3   FUNCTIONAL ANALYSIS

There were some differences in HUMAnN pathway vs. KO-summed pathway output. The generally lower abundance of HUMAnN pathway output is likely due to the fact that it averages the most abundant genes in the pathway, instead of summing total abundance. The differences inter-sample ratios observed with primary bile acid synthesis might have been due to taxonomic limitation and/or copy number normalization.

KOs unique to methane metabolism were greatly enriched in 6128 and 5100 Day 7 samples, mirroring the presence of methanogens in these samples detected by our taxonomic analysis. It is possible that methanogens are demethylating SECO in these incubations. The fact that the methane metabolism pathway (ko00680) was present in samples 5100 and 6128 Day 1, and both samples from 6105, in which methanogens were not detected by taxonomic analysis, was surprising. Upon further investigation, the 3 most abundant KOs in these samples are KOs found in many other pathways, in addition to methane metabolism. All KOs corresponding to modules unique to methane metabolism were found at very low levels in the non-methanogen samples, with the exception of three KOs belonging to M00357. These KOs are also part of several other pathways. It can thus be concluded that our detection of ko00680 in non-methanogen containing samples is mainly due to ubiquitous KOs, and that methane metabolism is not occurring at notable levels in these samples.

The purpose of MinPath is essentially to eliminate pathways that are falsely detected due to ubiquitous KOs. Since the HUMAnN pathway output was highly similar to the manually KO-summed pathway output, it seems that MinPath failed to eliminate the correct pathways. Furthermore, HUMAnN eliminated modules that contained non-ubiquitous KOs; yet when these modules were assessed by summing KOs, they were found to be present. Finally, KOs corresponding to a module that was present in HUMAnN module output, were not present in

HUMAnN KO output. Future studies should be conducted with an awareness of these limitations.

Since several acetogens are capable of lignan metabolism, especially the demethylation step, enriching for these organisms may promote the conversion of SDG to ENL. Four KOs that correspond to the last step in the acetyl CoA pathway (Figure 26) were more abundant in the 6128 and 5100 Day 7 samples. K00194, K00197, and K00198 are also utilized in methanogenesis, while K14138 is not. Co-occurrence of these four specific KOs has been shown to be a good indicator of the presence of the WLP in microbes.[72,95] This tracks well with the fact that *E. limnosum* − an acetogenic bacterium − and methanogens were highly enriched in these samples, since methanogens and the acetogenic bacteria both utilize the WLP. Furthermore acetogens are a phylogenetically diverse group and it is likely that there are many uncharacterized acetogens in the human gut.[72] Thus there may be undetected acetogens in these samples that are additionally accounting for the presence of the WLP.

Furthering this hypothesis, KOs corresponding to acetogenic demethylases increased substantially with incubation of samples from 6128 and 5100. The enzymes at the end of the WLP are those that receive the methyl group from these demethylases, thus the pathway for methoxylated aromatic demethylation appears to be enriched for (Figure 26).

KOs corresponding to beta-glucosidases and beta-glucuronidases were detected at high abundance in our samples and may be responsible for the SDG deglycosylation in our samples. No *bai* genes were found in our investigation. Two abundant KOs involved in bile acid biosynthesis were detected. Choloylglycine hydrolase (K01442) is a hydrolase that cleaves carbon-nitrogen bonds other than peptide bonds. Alpha-methylacyl-CoA racemase (K01796) is a racemase that is involved in converting cholesterol to cholate and chenodeoxycholate.[88] Though these enzymes probably would not act directly on lignans, their presence indicates that bile acid metabolizing microbes are present in our samples.

KOs corresponding to bile acid dehydrogenases were enriched in the 6128 and 5100 incubations, and a KO corresponding to a steroid hormone dehydrogenase was enriched in the 6105 incubation. It is possible that these dehydrogenases are acting to dehydrogenate END in our incubations.

4.4   MEDIA FOR *IN VITRO* INCUBATIONS

Several different growth media have been used to culture enterolignan-producing microbes. In a study described previously, Brain Heart Infusion broth was used to culture feces from 100 different individuals, yielding variable degrees of enterolignan production.[5] Mt Broth has also been used, yielding an average SECO to enterolignan (END+ENL) percent conversion of 59.9%.[137] Human feces incubated with General Anaerobic Medium converted 90% of SDG to ENL.[13] When MRS broth was used *in vitro*, microbes were only capable of metabolizing SDG to SECO and did not produce enterolignans.[62] The efficient conversion of SDG to ENL observed in our incubations may be due in part to the ability of our media to promote the survival of lignan-metabolizing organisms.

The reagents in our media provided substrates necessary for acetogens and methanogens to thrive. Our media contained formate. Acetogens do not necessarily possess all enzymes in the WLP, and not all acetogens are capable of growing solely on H_2 and CO_2. Since formate is the intermediate after CO_2 in the WLP, it follows that several acetogens have been shown to grow well on formate,[73] and addition of formate can drive acetogenesis.[149] The acetogen *Marvinbryantia formatexigens* (isolated from human stool), does not harbor the selenium-containing subunit of formate dehydrogenase, and requires formate for acetogenesis.[72] Addition of sodium formate to the culture media for this species has been shown to greatly increase its growth.[150] Methanogens can also utilize formate, and addition of formate to methanogen culture media has been shown to promote methanogenesis.[67,149]

Our media also contained acetate. Acetate is standardly used in enrichment media for methanogens, and can be entered into the incomplete reductive tricarboxylic acid (TCA) cycle for energy yield. Methanogens can also utilize ammonium and sulfur, both provided by ammonium sulfate in our media.[67]

Our incubations contained a substantial amount of methanol, which can be used as a substrate by acetogens, such as *E. limnosum,*[73] and methanogens.[67] The standard concentration of methanol used for culturing methanogens is 0.4%,[67] while our media contained 1.5%. Trace metals provided in our media are important for the metalloenzymes that acetogens and methanogens possess.[67,73] Glucose, contained in our media, can sustain both primary fermenters

and acetogens. The fermentation of glucose would generate CO2, H2, and short chain fatty acids, which could subsequently be utilized by acetogens and methanogens.[69,151]

Acetogens and methanogens are considered obligate anaerobes.[67,73] We ensured our media was anaerobic by adding an anaerobic indicator, sparging with nitrogen, and adding the reducing agent Cys-HCl. Both methanogens and acetogens thrive in a neutral pH[67,73] and our media was brought to a pH of 7.3.

It is unclear why *A. muciniphila* grew well in our incubations. Potentially the small amount of SDG, being a polyphenol, could promote its growth. Additionally, methanol could be exerting antimicrobial effects, and reducing competition.

4.5   LIMITATIONS

There are several limitations to our methods. For one, the sample size is small, and limits our ability to detect differences between treatments. However, this is part of an ongoing study and future analyses will be conducted with more samples.

Eubacterial 16S rRNA gene copy number estimated by qPCR changed less over time for incubations that contained high levels of methanogens. This is indicative of the fact that we are only measuring the eubacterial, and not the archaeal, 16S rRNA gene. Future studies should use primers that amplify both the archaeal and eubacterial 16S rRNA gene.

We have observed that the community from four different participants converted SDG to ENL *in vitro*. We have also observed presence or enrichment of acetogens, methanogens, bile acid metabolizing microbes, and KOs corresponding to the WLP, methanogenesis, beta-glucosidases, beta-glucuronidases, demethylases, and bile acid and steroid dehydrogenases in these incubations. We may hypothesize that these organisms and pathways are responsible for SDG metabolism, but there is no firm evidence to conclude such. The community in our cultures was very large and it could be argued that the relative abundance of microbes involved in lignan metabolism actually decreases *in vitro*, but remains at just a high enough level to enable conversion of SDG to ENL. Further steps should be taken to solidify understanding.

Isolation of microbes and microbial enzymes, and testing for lignan-metabolizing capability would provide the most concrete answers. Organismal isolation could be carried out as described in Dr. Clavel's thesis.[137] If investigation of acetogens or methanogens was desired, substrates used for their enrichment could be added to plates and media. For acetogens one might add

methoxylated aromatic compounds or formate.[73] Probably the most relevant and useful methoxylated aromatic to add would be SECO. For methanogens one might add methanol, formate, acetate, CO_2, and H_2.[67] Methane, acetate, H_2, and CO_2 in serum bottles could be measured over time to further characterize hydrogenotrophic metabolism.

Putative lignan-metabolizing enzyme genes could be amplified from identified lignan-converting microbes and subsequently cloned, expressed, and characterized as in Tao et al.[91] Furthermore, isolates belonging to groups of microbes hypothesized to be involved in lignan metabolism could be ordered from a supplier and subsequently tested for lignan-degrading capabilities.

It may also be advantageous to inoculate media with enough SDG (or other lignans) to potentially promote growth of microbes that are using it as a substrate. The concentration of SDG (6.55 μM) accounts for less than 0.01% of the total carbon in our media. In addition, hydrogenotrophic metabolism is low-energy yielding.[67,73] For these reasons it is unlikely that 6.55 μM SDG would be enough to serve as a substantial food source. It may be difficult, however, to add enough lignan substrate to promote growth. For example, in one study, compounds demethylated by acetogenic bacteria had no effect on growth.[152] This may have been due to several factors, including less ATP generated from metabolism than is required for uptake of the compound or disposal of metabolites. Also, higher levels of SDG might prove toxic to microbes. Flaxseed-cake extract with a concentration of 364 μM SDG was shown to slow growth of *Staphylococcus aureus* in Luria Broth.[153]

Incubation in media that does not support SDG to ENL conversion might generate useful insights. Microbial populations grown in this media could be compared to populations grown in TCAP2, with the idea that TCAP2-grown communities would be richer in lignan-metabolizers. Differences in phenotype, however, could be gene expression-mediated, in which case taxonomy and functional gene analysis would provide little insight.

This precipitates another limitation and future direction. Changes in functional gene abundance may belie actual changes in function, in that many of the genes measured metagenomically are not expressed. Adding a metatranscriptomic component would give a much clearer picture of actual functions at work. It may be expected that metatranscriptomic changes would be more transient, thus it would be critical to identify time points of interest for functional characterization.

Subsequently, we were only comparing initial microbial community to Day 7, but there were likely changes in community structure throughout the incubation. It also may be the case that DNA sequenced from the pelleted fecal suspension originated from microbes that died early in the incubation. Future studies should pellet fecal suspensions on several days to capture the transformation of the microbial community.

Several taxa in addition to those identified or hypothesized may be involved in lignan metabolism. By focusing our analyses on specific taxa, we may have missed important phylogenetically related organisms. Furthermore there are many microbial genomes that have not been fully sequenced. These microbes would not have been identified by our methods. Similarly we may have missed functional genes that are closely related to the KOs searched for, but have a different KO numbers or are unclassified. Analyzing our data using phylogenetic trees would help to remedy these problems.

Using KOs as a gene identifier also poses problems. A single KO can be found in many different enzymes and pathways. For example K00548, which is part of an acetogenic demethylase, is also used in several pathways including, cysteine and methionine metabolism, selenocompound metabolism, one carbon pool by folate, and biosynthesis of amino acids[88] and is found in 67% of gut microbes.[154]

4.6 CONCLUSION

In summary, SDG was almost fully converted to ENL in our incubations. In most incubations, Bacteroidetes:Firmicutes ratio increased; methanogens, deglycosylators – especially *Bacteroides* – and demethylators were enriched; and KOs corresponding to beta-glucosidases were present at high levels. Several phenomena were true of some incubations. *Eubacterium limnosum*, *Akkermansia muciniphila,* methane metabolism, and the WLP were highly enriched. KOs corresponding to demethylases, dehydroxylases and dehydrogenases were enriched. The acetogenic demethylation pathway may have been enriched. To conclude, organisms and functional genes identified or hypothesized to be involved in lignan metabolism were not unanimously enriched, but sample-specific. These findings may shed light on organisms, enzymes, and biochemical pathways involved in lignan metabolism. Future studies with larger sample sizes may help to confirm these results.

REFERENCES

1.	Webb AL, McCullough ML. Dietary lignans: potential role in cancer prevention. *Nutr Cancer.* 2005;51(2):117-131.
2.	Peterson J, Dwyer J, Adlercreutz H, Scalbert A, Jacques P, McCullough ML. Dietary lignans: physiology and potential for cardiovascular disease risk reduction. *Nutr Rev.* 2010;68(10):571-603.
3.	Milder IE, Arts IC, van de Putte B, Venema DP, Hollman PC. Lignan contents of Dutch plant foods: a database including lariciresinol, pinoresinol, secoisolariciresinol and matairesinol. *Br J Nutr.* 2005;93(3):393-402.
4.	Clavel T, Dore J, Blaut M. Bioavailability of lignans in human subjects. *Nutr Res Rev.* 2006;19(2):187-196.
5.	Possemiers S, Bolca S, Eeckhaut E, Depypere H, Verstraete W. Metabolism of isoflavones, lignans and prenylflavonoids by intestinal bacteria: producer phenotyping and relation with intestinal community. *FEMS Microbiol Ecol.* Vol 61. England2007:372-383.
6.	Penalvo JL, Heinonen SM, Aura AM, Adlercreutz H. Dietary sesamin is converted to enterolactone in humans. *J Nutr.* Vol 135. United States2005:1056-1062.
7.	Lampe JW, Gustafson DR, Hutchins AM, et al. Urinary isoflavonoid and lignan excretion on a Western diet: relation to soy, vegetable, and fruit intake. *Cancer Epidemiol Biomarkers Prev.* 1999;8(8):699-707.
8.	Clavel T, Borrmann D, Braune A, Dore J, Blaut M. Occurrence and activity of human intestinal bacteria involved in the conversion of dietary lignans. *Anaerobe.* 2006;12(3):140-147.
9.	Nose M, Fujimoto T, Takeda T, Nishibe S, Ogihara Y. Structural transformation of lignan compounds in rat gastrointestinal tract. *Planta Med.* 1992;58(6):520-523.
10.	Eeckhaut E, Struijs K, Possemiers S, Vincken JP, Keukeleire DD, Verstraete W. Metabolism of the lignan macromolecule into enterolignans in the gastrointestinal lumen as determined in the simulator of the human intestinal microbial ecosystem. *J Agric Food Chem.* 2008;56(12):4806-4812.
11.	Setchell KD, Lawson AM, Borriello SP, et al. Lignan formation in man--microbial involvement and possible roles in relation to cancer. *Lancet.* Vol 2. England1981:4-7.
12.	Yoder SC, Lancaster SM, Hullar MAJ, Lampe JW, Rio KTD. Chapter 7 - Gut Microbial Metabolism of Plant Lignans: Influence on Human Health. *Diet-Microbe Interactions in the Gut.* San Diego: Academic Press; 2015:103-117.
13.	Wang LQ, Meselhy MR, Li Y, Qin GW, Hattori M. Human intestinal bacteria capable of transforming secoisolariciresinol diglucoside to mammalian lignans, enterodiol and enterolactone. *Chem Pharm Bull (Tokyo).* 2000;48(11):1606-1610.
14.	Yoder S. *Metabolism of Plant Lignans by Human Intestinal Bacteria.* Seattle, WA: Nutritional Sciences Program, University of Washington; 2013.
15.	Xie LH, Akao T, Hamasaki K, Deyama T, Hattori M. Biotransformation of pinoresinol diglucoside to mammalian lignans by human intestinal microflora, and isolation of Enterococcus faecalis strain PDG-1 responsible for the transformation of (+)-pinoresinol to (+)-lariciresinol. *Chem Pharm Bull (Tokyo).* 2003;51(5):508-515.
16.	Knust U, Hull WE, Spiegelhalder B, Bartsch H, Strowitzki T, Owen RW. Analysis of enterolignan glucuronides in serum and urine by HPLC-ESI-MS. *Food Chem Toxicol.* Vol 44. England2006:1038-1049.
17.	Sicilia T, Niemeyer HB, Honig DM, Metzler M. Identification and stereochemical characterization of lignans in flaxseed and pumpkin seeds. *J Agric Food Chem.* 2003;51(5):1181-1188.
18.	Kuijsten A, Arts IC, Vree TB, Hollman PC. Pharmacokinetics of enterolignans in healthy men and women consuming a single dose of secoisolariciresinol diglucoside. *J Nutr.* 2005;135(4):795-801.
19.	Jansen GH, Arts IC, Nielen MW, Muller M, Hollman PC, Keijer J. Uptake and metabolism of enterolactone and enterodiol by human colon epithelial cells. *Arch Biochem Biophys.* Vol 435. United States2005:74-82.
20.	Axelson M, Setchell KD. The excretion of lignans in rats -- evidence for an intestinal bacterial source for this new group of compounds. *FEBS Lett.* Vol 123. Netherlands1981:337-342.
21.	Umeda-Sawada R, Ogawa M, Igarashi O. The metabolism and distribution of sesame lignans (sesamin and episesamin) in rats. *Lipids.* 1999;34(6):633-637.
22.	Hirose N, Inoue T, Nishihara K, et al. Inhibition of cholesterol absorption and synthesis in rats by sesamin. *J Lipid Res.* 1991;32(4):629-638.
23.	Murkies AL, Wilcox G, Davis SR. Phytoestrogens. *The Journal of Clinical Endocrinology & Metabolism.* 1998;83(2):297-303.
24.	Hong SJ, Kim SI, Kwon SM, Lee JR, Chung BC. Comparative study of concentration of isoflavones and lignans in plasma and prostatic tissues of normal control and benign prostatic hyperplasia. *Yonsei Med J.* Vol 43. Korea South2002:236-241.
25.	Boccardo F, Lunardi GL, Petti AR, Rubagotti A. Enterolactone in breast cyst fluid: correlation with EGF and breast cancer risk. *Breast Cancer Res Treat.* 2003;79(1):17-23.
26.	Adlercreutz H, van der Wildt J, Kinzel J, et al. Lignan and isoflavonoid conjugates in human urine. *J Steroid Biochem Mol Biol.* Vol 52. England1995:97-103.
27.	Clavel T, Mapesa JO. *Phenolics in Human Nutrition: Importance of the Intestinal Microbiome for Isoflavone and Lignan Bioavailability.* Springer Berlin Heidelberg; 2013.
28.	Demark-Wahnefried W, Polascik TJ, George SL, et al. Flaxseed supplementation (not dietary fat restriction) reduces prostate cancer proliferation rates in men presurgery. *Cancer Epidemiol Biomarkers Prev.* Vol 17. United States2008:3577-3587.
29.	Velentzis LS, Cantwell MM, Cardwell C, Keshtgar MR, Leathem AJ, Woodside JV. Lignans and breast cancer risk in pre- and post-menopausal women: meta-analyses of observational studies. *Br J Cancer.* Vol 100. England2009:1492-1498.
30.	Thompson LU, Chen JM, Li T, Strasser-Weippl K, Goss PE. Dietary flaxseed alters tumor biological markers in postmenopausal breast cancer. *Clin Cancer Res.* Vol 11. United States2005:3828-3835.
31.	Bassett CM, Rodriguez-Leyva D, Pierce GN. Experimental and clinical research findings on the cardiovascular benefits of consuming flaxseed. *Appl Physiol Nutr Metab.* Vol 34. Canada2009:965-974.
32.	Pan A, Yu D, Demark-Wahnefried W, Franco OH, Lin X. Meta-analysis of the effects of flaxseed interventions on blood lipids. *Am J Clin Nutr.* Vol 90. United States2009:288-297.

33. Ward HA, Kuhnle GG, Mulligan AA, Lentjes MA, Luben RN, Khaw KT. Breast, colorectal, and prostate cancer risk in the European Prospective Investigation into Cancer and Nutrition-Norfolk in relation to phytoestrogen intake derived from an improved database. *Am J Clin Nutr.* Vol 91. United States2010:440-448.

34. Hallund J, Tetens I, Bugel S, Tholstrup T, Bruun JM. The effect of a lignan complex isolated from flaxseed on inflammation markers in healthy postmenopausal women. *Nutr Metab Cardiovasc Dis.* Vol 18. Germany2008:497-502.

35. Zhang W, Wang X, Liu Y, et al. Effects of dietary flaxseed lignan extract on symptoms of benign prostatic hyperplasia. *J Med Food.* 2008;11(2):207-214.

36. Chen H, O'Reilly E, McCullough ML, et al. Consumption of dairy products and risk of Parkinson's disease. *Am J Epidemiol.* 2007;165(9):998-1006.

37. Rickard SE, Yuan YV, Chen J, Thompson LU. Dose effects of flaxseed and its lignan on N-methyl-N-nitrosourea-induced mammary tumorigenesis in rats. *Nutr Cancer.* 1999;35(1):50-57.

38. Li D, Yee JA, Thompson LU, Yan L. Dietary supplementation with secoisolariciresinol diglycoside (SDG) reduces experimental metastasis of melanoma cells in mice. *Cancer Lett.* Vol 142. Ireland1999:91-96.

39. Jenab M, Thompson LU. The influence of flaxseed and lignans on colon carcinogenesis and beta-glucuronidase activity. *Carcinogenesis.* 1996;17(6):1343-1348.

40. Prasad K. Regression of hypercholesterolemic atherosclerosis in rabbits by secoisolariciresinol diglucoside isolated from flaxseed. *Atherosclerosis.* Vol 197. Ireland2008:34-42.

41. Penumathsa SV, Koneru S, Thirunavukkarasu M, Zhan L, Prasad K, Maulik N. Secoisolariciresinol diglucoside: relevance to angiogenesis and cardioprotection against ischemia-reperfusion injury. *J Pharmacol Exp Ther.* Vol 320. United States2007:951-959.

42. Jacobs MN, Nolan GT, Hood SR. Lignans, bacteriocides and organochlorine compounds activate the human pregnane X receptor (PXR). *Toxicol Appl Pharmacol.* Vol 209. United States2005:123-133.

43. Penttinen P, Jaehrling J, Damdimopoulos AE, et al. Diet-derived polyphenol metabolite enterolactone is a tissue-specific estrogen receptor activator. *Endocrinology.* Vol 148. United States2007:4875-4886.

44. Mueller SO, Simon S, Chae K, Metzler M, Korach KS. Phytoestrogens and their human metabolites show distinct agonistic and antagonistic properties on estrogen receptor alpha (ERalpha) and ERbeta in human cells. *Toxicol Sci.* Vol 80. United States2004:14-25.

45. Schottner M, Spiteller G, Gansser D. Lignans interfering with 5 alpha-dihydrotestosterone binding to human sex hormone-binding globulin. *J Nat Prod.* Vol 61. United States1998:119-121.

46. Kitts DD, Yuan YV, Wijewickreme AN, Thompson LU. Antioxidant activity of the flaxseed lignan secoisolariciresinol diglycoside and its mammalian lignan metabolites enterodiol and enterolactone. *Mol Cell Biochem.* 1999;202(1-2):91-100.

47. Prasad K. Antioxidant Activity of Secoisolariciresinol Diglucoside-derived Metabolites, Secoisolariciresinol, Enterodiol, and Enterolactone. *Int J Angiol.* Vol 92000:220-225.

48. Adlercreutz H, Bannwart C, Wahala K, et al. Inhibition of human aromatase by mammalian lignans and isoflavonoid phytoestrogens. *J Steroid Biochem Mol Biol.* 1993;44(2):147-153.

49. Nesbitt PD, Lam Y, Thompson LU. Human metabolism of mammalian lignan precursors in raw and processed flaxseed. *Am J Clin Nutr.* 1999;69(3):549-555.

50. Lampe JW, Martini MC, Kurzer MS, Adlercreutz H, Slavin JL. Urinary lignan and isoflavonoid excretion in premenopausal women consuming flaxseed powder. *Am J Clin Nutr.* 1994;60(1):122-128.

51. Clavel T, Henderson G, Alpert CA, et al. Intestinal bacterial communities that produce active estrogen-like compounds enterodiol and enterolactone in humans. *Appl Environ Microbiol.* Vol 71. United States2005:6077-6085.

52. Lampe JW. Isoflavonoid and lignan phytoestrogens as dietary biomarkers. *J Nutr.* 2003;133 Suppl 3:956S-964S.

53. Sonestedt E, Wirfalt E. Enterolactone and breast cancer: methodological issues may contribute to conflicting results in observational studies. *Nutr Res.* Vol 30. United States: 2010 Elsevier Inc; 2010:667-677.

54. Hullar MA, Lancaster SM, Li F, et al. Enterolignan Producing Phenotypes are Associated with Increased Gut Microbial Diversity and Altered Composition in Premenopausal Women in the United States. *Cancer Epidemiol Biomarkers Prev.* 2014.

55. Elliott SP. Rat bite fever and Streptobacillus moniliformis. *Clin Microbiol Rev.* 2007;20(1):13-22.

56. Falsen E, Collins MD, Welinder-Olsson C, Song Y, Finegold SM, Lawson PA. Fastidiosipila sanguinis gen. nov., sp. nov., a new Gram-positive, coccus-shaped organism from human blood. *Int J Syst Evol Microbiol.* 2005;55(Pt 2):853-858.

57. Larcia LL, 2nd, Estacio RC, Dalmacio LM. Bacterial diversity in Philippine fermented mustard (burong mustasa) as revealed by 16S rRNA gene analysis. *Benef Microbes.* 2011;2(4):263-271.

58. Mroczynska M, Libudzisz Z. Beta-glucuronidase and beta-glucosidase activity of Lactobacillus and Enterococcus isolated from human feces. *Pol J Microbiol.* 2010;59(4):265-269.

59. Schogor AL, Huws SA, Santos GT, et al. Ruminal Prevotella spp. may play an important role in the conversion of plant lignans into human health beneficial antioxidants. *PLoS One.* Vol 9. United States2014:e87949.

60. Clavel T, Lippman R, Gavini F, Dore J, Blaut M. Clostridium saccharogumia sp. nov. and Lactonifactor longoviformis gen. nov., sp. nov., two novel human faecal bacteria involved in the conversion of the dietary phytoestrogen secoisolariciresinol diglucoside. *Syst Appl Microbiol.* 2007;30(1):16-26.

61. Clavel T, Henderson G, Engst W, Dore J, Blaut M. Phylogeny of human intestinal bacteria that activate the dietary lignan secoisolariciresinol diglucoside. *FEMS Microbiol Ecol.* 2006;55(3):471-478.

62. Roncaglia L, Amaretti A, Raimondi S, Leonardi A, Rossi M. Role of bifidobacteria in the activation of the lignan secoisolariciresinol diglucoside. *Appl Microbiol Biotechnol.* 2011;92(1):159-168.

63. Jin JS, Hattori M. Human intestinal bacterium, strain END-2 is responsible for demethylation as well as lactonization during plant lignan metabolism. *Biol Pharm Bull.* 2010;33(8):1443-1447.

64. Jin JS, Zhao YF, Nakamura N, et al. Enantioselective dehydroxylation of enterodiol and enterolactone precursors by human intestinal bacteria. *Biol Pharm Bull.* Vol 30. Japan2007:2113-2119.

65. Jin JS, Kakiuchi N, Hattori M. Enantioselective oxidation of enterodiol to enterolactone by human intestinal bacteria. *Biol Pharm Bull.* 2007;30(11):2204-2206.

66. Schuchmann K, Muller V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat Rev Microbiol.* Vol 12. England2014:809-821.

67.   Hedderich R, Whitman WB. Physiology and Biochemistry of the Methane-Producing Archaea. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, eds. *Prokaryotes: A Handbook on the Biology of Bacteria, Vol 2, Third Edition:ECOPHYSIOLOGY AND BIOCHEMISTRY*: Springer, 233 Spring Street, New York, Ny 10013, United States; 2006:1050-1079.

68.   Thauer RK, Jungermann K, Decker K. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev.* 1977;41(1):100-180.

69.   Carbonero F, Benefiel AC, Gaskins HR. Contributions of the microbial hydrogen economy to colonic homeostasis. *Nat Rev Gastroenterol Hepatol.* Vol 9. England2012:504-518.

70.   Muyzer G, Stams AJ. The ecology and biotechnology of sulphate-reducing bacteria. *Nat Rev Microbiol.* 2008;6(6):441-454.

71.   Christl SU, Murgatroyd PR, Gibson GR, Cummings JH. Production, metabolism, and excretion of hydrogen in the large intestine. *Gastroenterology.* Vol 102. United States1992:1269-1277.

72.   Rey FE, Faith JJ, Bain J, et al. Dissecting the in vivo metabolic potential of two human gut acetogens. *J Biol Chem.* Vol 285. United States2010:22082-22090.

73.   Drake HL, Küsel K, Matthies C. Acetogenic Prokaryotes. 2006.

74.   Kerby R, Zeikus JG. Catabolic enzymes of the acetogen Butyribacterium methylotrophicum grown on single-carbon substrates. *J Bacteriol.* 1987;169(12):5605-5609.

75.   Schiel-Bengelsdorf B, Durre P. Pathway engineering and synthetic biology using acetogens. *FEBS Lett.* 2012;586(15):2191-2198.

76.   Jin JS, Zhao YF, Nakamura N, Akao T, Kakiuchi N, Hattori M. Isolation and characterization of a human intestinal bacterium, Eubacterium sp. ARC-2, capable of demethylating arctigenin, in the essential metabolic process to enterolactone. *Biol Pharm Bull.* Vol 30. Japan2007:904-911.

77.   Ottesen EA, Leadbetter JR. Diversity of formyltetrahydrofolate synthetases in the guts of the wood-feeding cockroach Cryptocercus punctulatus and the omnivorous cockroach Periplaneta americana. *Appl Environ Microbiol.* 2010;76(14):4909-4913.

78.   Willems A, Collins MD. Phylogenetic relationships of the genera Acetobacterium and Eubacterium sensu stricto and reclassification of Eubacterium alactolyticum as Pseudoramibacter alactolyticus gen. nov., comb. nov. *Int J Syst Bacteriol.* 1996;46(4):1083-1087.

79.   Carlier JP, K'Ouas G, Han XY. Moryella indoligenes gen. nov., sp. nov., an anaerobic bacterium isolated from clinical specimens. *Int J Syst Evol Microbiol.* Vol 57. England2007:725-729.

80.   Brugere JF, Borrel G, Gaci N, Tottey W, O'Toole PW, Malpuech-Brugere C. Archaebiotics: Proposed therapeutic use of archaea to prevent trimethylaminuria and cardiovascular disease. *Gut Microbes.* 2013;5(1).

81.   Borrel G, Harris HM, Tottey W, et al. Genome sequence of "Candidatus Methanomethylophilus alvus" Mx1201, a methanogenic archaeon from the human gut belonging to a seventh order of methanogens. *J Bacteriol.* Vol 194. United States2012:6944-6945.

82.   Gaci N, Borrel G, Tottey W, O'Toole PW, Brugere JF. Archaea and the human gut: new beginning of an old story. *World J Gastroenterol.* 2014;20(43):16062-16078.

83.   Liu Y, Whitman WB. Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann N Y Acad Sci.* Vol 1125. United States2008:171-189.

84.   Ferry JG. Enzymology of one-carbon metabolism in methanogenic pathways. *FEMS Microbiol Rev.* Vol 23. Netherlands1999:13-38.

85.   Borrel G, Harris HM, Parisot N, et al. Genome Sequence of "Candidatus Methanomassiliicoccus intestinalis" Issoire-Mx1, a Third Thermoplasmatales-Related Methanogenic Archaeon from Human Feces. *Genome Announc.* 2013;1(4).

86.   Dridi B, Fardeau ML, Ollivier B, Raoult D, Drancourt M. Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int J Syst Evol Microbiol.* Vol 62. England2012:1902-1907.

87.   International Union of Biochemistry and Molecular Biology, Nomenclature Committee, Webb E. *Enzyme nomenclature 1992 : recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes.* San Diego: Academic Press; 1992.

88.   Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* Vol 28. England2000:27-30.

89.   McBain AJ, Macfarlane GT. Ecological and physiological studies on large intestinal bacteria in relation to production of hydrolytic and reductive enzymes involved in formation of genotoxic metabolites. *J Med Microbiol.* 1998;47(5):407-416.

90.   Dabek M, McCrae SI, Stevens VJ, Duncan SH, Louis P. Distribution of beta-glucosidase and beta-glucuronidase activity and of beta-glucuronidase gene gus in human colonic bacteria. *FEMS Microbiol Ecol.* Vol 66. England2008:487-495.

91.   Tao YL, Yang DH, Zhang YT, et al. Cloning, expression, and characterization of the beta-glucosidase hydrolyzing secoisolariciresinol diglucoside to secoisolariciresinol from Bacteroides uniformis ZL1. *Appl Microbiol Biotechnol.* 2014;98(6):2519-2531.

92.   Bache R, Pfennig N. SELECTIVE ISOLATION OF ACETOBACTERIUM-WOODII ON METHOXYLATED AROMATIC-ACIDS AND DETERMINATION OF GROWTH YIELDS. *Archives of Microbiology.* 1981;130(3):255-261.

93.   Naidu D, Ragsdale SW. Characterization of a three-component vanillate O-demethylase from Moorella thermoacetica. *J Bacteriol.* 2001;183(11):3276-3281.

94.   Sharak Genthner BR, Bryant MP. Additional characteristics of one-carbon-compound utilization by Eubacterium limosum and Acetobacterium woodii. *Appl Environ Microbiol.* 1987;53(3):471-476.

95.   Pierce E, Xie G, Barabote RD, et al. The complete genome sequence of Moorella thermoacetica (f. Clostridium thermoaceticum). *Environ Microbiol.* Vol 10. England2008:2550-2573.

96.   Kaufmann F, Wohlfarth G, Diekert G. O-demethylase from Acetobacterium dehalogenans--cloning, sequencing, and active expression of the gene encoding the corrinoid protein. *Eur J Biochem.* 1998;257(2):515-521.

97.   Anderson I, Held B, Lapidus A, et al. Genome sequence of the homoacetogenic bacterium Holophaga foetida type strain (TMBS4(T)). *Stand Genomic Sci.* Vol 6. United States2012:174-184.

98.   Studenik S, Vogel M, Diekert G. Characterization of an O-demethylase of Desulfitobacterium hafniense DCB-2. *J Bacteriol.* Vol 194. United States2012:3317-3326.

99.   Ridlon JM, Kang DJ, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. *J Lipid Res.* Vol 47. United States2006:241-259.

100.   Ridlon JM, Alves JM, Hylemon PB, Bajaj JS. Cirrhosis, bile acids and gut microbiota: unraveling a complex relationship. *Gut Microbes.* Vol 4. United States2013:382-387.

101.   Buffie CG, Bucci V, Stein RR, et al. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature.* Vol 517. England2015:205-208.

102.   Doerner KC, Takamine F, LaVoie CP, Mallonee DH, Hylemon PB. Assessment of fecal bacteria with bile acid 7 alpha-dehydroxylating activity for the presence of bai-like genes. *Appl Environ Microbiol.* 1997;63(3):1185-1188.

103.   Ye HQ, Mallonee DH, Wells JE, Bjorkhem I, Hylemon PB. The bile acid-inducible baiF gene from Eubacterium sp. strain VPI 12708 encodes a bile acid-coenzyme A hydrolase. *J Lipid Res.* 1999;40(1):17-23.

104.   Sherrod JA, Hylemon PB. Partial purification and characterization of NAD-dependent 7alpha-hydroxysteroid dehydrogenase from Bacteroides thetaiotaomicron. *Biochim Biophys Acta.* 1977;486(2):351-358.

105.   Feighner SD, Hylemon PB. Characterization of a corticosteroid 21-dehydroxylase from the intestinal anaerobic bacterium, Eubacterium lentum. *J Lipid Res.* 1980;21(5):585-593.

106.   Masuda N, Oda H, Hirano S, Masuda M, Tanaka H. 7 alpha-Dehydroxylation of bile acids by resting cells of a Eubacterium lentum-like intestinal anaerobe, strain c-25. *Appl Environ Microbiol.* 1984;47(4):735-739.

107.   Kisiela M, Skarka A, Ebert B, Maser E. Hydroxysteroid dehydrogenases (HSDs) in bacteria: a bioinformatic perspective. *J Steroid Biochem Mol Biol.* Vol 129. England: A 2011 Elsevier Ltd; 2012:31-46.

108.   Lee JY, Arai H, Nakamura Y, Fukiya S, Wada M, Yokota A. Contribution of the 7beta-hydroxysteroid dehydrogenase from Ruminococcus gnavus N53 to ursodeoxycholic acid formation in the human colon. *J Lipid Res.* Vol 54. United States2013:3062-3069.

109.   Gagen EJ, Wang J, Padmanabha J, et al. Investigation of a new acetogen isolated from an enrichment of the tammar wallaby forestomach. *BMC Microbiol.* Vol 142014:314.

110.   Lepercq P, Gerard P, Beguet F, et al. Epimerization of chenodeoxycholic acid to ursodeoxycholic acid by Clostridium baratii isolated from human feces. *FEMS Microbiol Lett.* Vol 235. Netherlands2004:65-72.

111.   Jarvenpaa P, Kosunen T, Fotsis T, Adlercreutz H. In vitro metabolism of estrogens by isolated intestinal micro-organisms and by human faecal microflora. *J Steroid Biochem.* 1980;13(3):345-349.

112.   Pascual A, Perez MH, Jaton K, et al. Mycoplasma hominis necrotizing pleuropneumonia in a previously healthy adolescent. *BMC Infect Dis.* Vol 10. England2010:335.

113.   Ravva SV, Sarreal CZ, Mandrell RE. Bacterial communities in aerosols and manure samples from two different dairies in central and Sonoma valleys of California. *PLoS One.* 2011;6(2):e17281.

114.   Frankenfeld CL, McTiernan A, Tworoger SS, et al. Serum steroid hormones, sex hormone-binding globulin concentrations, and urinary hydroxylated estrogen metabolites in post-menopausal women in relation to daidzein-metabolizing phenotypes. *J Steroid Biochem Mol Biol.* Vol 88. England2004:399-408.

115.   Persson S, de Boer RF, Kooistra-Smid AM, Olsen KE. Five commercial DNA extraction systems tested and compared on a stool sample collection. *Diagn Microbiol Infect Dis.* Vol 69. United States: 2011 Elsevier Inc; 2011:240-244.

116.   Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* Vol 7. United States2010:335-336.

117.   Illumina. TruSeq DNA Sample Preparation Guide. San Diego, CA: Author; 2012.

118.   Staden R. A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res.* 1979;6(7):2601-2610.

119.   Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol.* Vol 73. United States2007:5261-5267.

120.   Kim M, Morrison M, Yu Z. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods.* Vol 84. Netherlands: 2010 Elsevier B.V; 2011:81-87.

121.   Meyer F, Paarmann D, D'Souza M, et al. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics.* Vol 9. England2008:386.

122.   Cox MP, Peterson DA, Biggs PJ. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics.* Vol 11. England2010:485.

123.   Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* Vol 10. England2009:R25.

124.   Rho M, Tang H, Ye Y. FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* Vol 38. England2010:e191.

125.   Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* Vol 26. England2010:2460-2461.

126.   Kent WJ. BLAT--the BLAST-like alignment tool. *Genome Res.* 2002;12(4):656-664.

127.   Wilke A, Harrison T, Wilkening J, et al. The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics.* Vol 13. England2012:141.

128.   Martin J, Sykes S, Young S, et al. Optimizing read mapping to reference genomes to determine composition and species prevalence in microbial communities. *PLoS One.* Vol 7. United States2012:e36427.

129.   Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods.* Vol 9. United States2012:811-814.

130.   Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-359.

131.   Markowitz VM, Chen IM, Palaniappan K, et al. The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res.* 2010;38(Database issue):D382-390.

132.   Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* Vol 12. United States2015:59-60.

133.   Abubucker S, Segata N, Goll J, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol.* Vol 8. United States2012:e1002358.

134.   Ye Y, Doak TG. A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. *PLoS Comput Biol.* 2009;5(8):e1000465.

135.   Magurran A. Measuring biological diversity. Oxford, UK: Blackwell; 2004.

136.   *vegan: Community Ecology Package.* [computer program]. Version R version 2.0–3.2012.

137.   Clavel T. *Metabolism of the dietary lignan secoisolariciresinol diglucoside by human intestinal bacteria*. Potsdam Abteilung Gastrointestinale Mikrobiologie   Deutsches Institut für Ernährungsforschung Potsdam-Rehbrücke   2005.

138.   Nayfach S, Pollard KS. Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol.* Vol 16. England2015:51.

139.   Brown TA. *Genomes*. Oxford: Wiley-Liss & Bios Scientific Publishers; 2002.

140.   Kim JN, Kim BS, Kim SJ, Cerniglia CE. Effects of crude oil, dispersant, and oil-dispersant mixtures on human fecal microbiota in an in vitro culture system. *MBio.* 2012;3(5).

141.   Lee HC, Jenner AM, Low CS, Lee YK. Effect of tea phenolics and their aromatic fecal bacterial metabolites on intestinal microbiota. *Res Microbiol.* Vol 157. France2006:876-884.

142.	Kemperman RA, Gross G, Mondot S, et al. Impact of polyphenols from black tea and red wine/grape juice on a gut model microbiome. *Food Research International.* 2013;53(2):659-669.
143.	Derrien M, Vaughan EE, Plugge CM, de Vos WM. Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol.* Vol 54. England2004:1469-1476.
144.	Axling U, Olsson C, Xu J, et al. Green tea powder and Lactobacillus plantarum affect gut microbiota, lipid metabolism and inflammation in high-fat fed C57BL/6J mice. *Nutr Metab (Lond).* Vol 9. England2012:105.
145.	Roopchand DE, Carmody RN, Kuhn P, et al. Dietary polyphenols promote growth of the gut bacterium Akkermansia muciniphila and attenuate high fat diet-induced metabolic syndrome. *Diabetes.* 2015.
146.	Anhe FF, Roy D, Pilon G, et al. A polyphenol-rich cranberry extract protects from diet-induced obesity, insulin resistance and intestinal inflammation in association with increased Akkermansia spp. population in the gut microbiota of mice. *Gut.* Vol 64. England: Published by the BMJ Publishing Group Limited. 2015:872-883.
147.	Dubourg G, Lagier JC, Armougom F, et al. High-level colonisation of the human gut by Verrucomicrobia following broad-spectrum antibiotic treatment. *Int J Antimicrob Agents.* 2013;41(2):149-155.
148.	Lagkouvardos I, Klaring K, Heinzmann SS, et al. Gut metabolites and bacterial community networks during a pilot intervention study with flaxseeds in healthy adult men. *Mol Nutr Food Res.* 2015.
149.	Hunger S, Schmidt O, Hilgarth M, et al. Competing formate- and carbon dioxide-utilizing prokaryotes in an anoxic methane-emitting fen soil. *Appl Environ Microbiol.* Vol 77. United States2011:3773-3785.
150.	Wolin MJ, Miller TL, Collins MD, Lawson PA. Formate-dependent growth and homoacetogenic fermentation by a bacterium from human feces: description of Bryantella formatexigens gen. nov., sp. nov. *Appl Environ Microbiol.* 2003;69(10):6321-6326.
151.	Kane MD, Breznak JA. Acetonema longum gen. nov. sp. nov., an H2/CO2 acetogenic bacterium from the termite, Pterotermes occidentis. *Arch Microbiol.* 1991;156(2):91-98.
152.	Jansen M, Hansen TA. Non-Growth-Associated Demethylation of Dimethylsulfoniopropionate by (Homo)acetogenic Bacteria. *Appl Environ Microbiol.* 2001;67(1):300-306.
153.	Czemplik M, Kulma A, Bazela K, Szopa J. The biomedical potential of genetically modified flax seeds overexpressing the glucosyltransferase gene. *BMC Complement Altern Med.* Vol 12. England2012:251.
154.	Degnan PH, Taga ME, Goodman AL. Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab.* 2014;20(5):769-778.
155.	Gibson GR, Macfarlane GT, Cummings JH. Sulphate reducing bacteria and hydrogen metabolism in the human large intestine. *Gut.* 1993;34(4):437-439.

FIGURES AND TABLES

Table 1. Participant study number, gender, age, whether qPCR analysis was conducted, time points of aliquots from fecal suspensions that were used for GCMS analysis, time points (Day 1 initial stool and final-day pelleted SDG-inoculated fecal suspension) from which DNA was extracted for sequencing, whether or not DNA was sequenced using 16S rRNA gene sequencing and/or shotgun metagenomic sequencing.

Participant #	Gender	Age	qPCR	GCMS timepoints	DNA extraction timepoints	16S rRNA gene sequencing	Shotgun metagenomic sequencing
5100	Female	50.3	yes	Days 1 - 6	Days 1 & 7	no	yes
6105	Female	41.77	yes	Days 1 - 7	Days 1 & 7	no	yes
6127	Male	40.2	no	Days 1 - 6	Days 1 & 7	yes	no
6128	Female	25.55	yes	Days 1 - 6	Days 1 & 7	yes	yes
6129	Male	30.89	yes	Days 1 - 6	Days 1 & 7	yes	no
6171	Female	43.63	no	Days 1 & 6	Days 1 & 6	yes	no
6178	Male	25.01	no	Days 1 & 6	Days 1 & 6	yes	no
6188	Female	28.27	no	Days 1 & 6	Days 1 & 6	yes	no

Table 2. TCAP2 Media Components

TCAP2	
	g/L
L-Cysteine-HCl (Aldrich)	0.25
Glucose (Sigma)	0.9
Resazurin (1mg/L) (Aldrich)	1
NaCl (Sigma)	2.78
K_2HPO_4 (Sigma)	0.42
$Na_2HPO_4 \cdot 7H_2O$ (Fisher)	3.3
$NaH_2PO_4 \cdot 2H_2O$ (Fluka)	1.2
$(NH_4)_2SO_4$ (Fisher)	1.2
$MgCl_2 \cdot 6H_2O$ (Sigma)	0.08
KH_2PO_4 (Fisher)	0.16
CH_3COONa (Fisher)	0.5
HCOONa (Fluka)	0.5
Trace element solution	10 mL
Vitamin solution	1 mL

Trace Element Solution	mg/L
$MnSO_4 \cdot H_2O$ (Fisher)	1159
$FeSO_4 \cdot 7H_2O$ (Fisher)	3680
$CoCl_2 \cdot 6H_2O$ (MP Biomedicals)	120
$ZnSO_4 \cdot 7H_2O$ (Fisher)	440
$CuSO_4 \cdot 5H_2O$ (Fisher)	98
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$ (Fluka)	17.4

Vitamin Solution	mg/L
Menadione (MP Biomedicals)	1
Biotin (Jackson Immuno Research)	2
Pantothenate (MP Biomedicals)	10
Nicotinamide (Acros)	5
Vitamin B_{12} (unknown)	0.5
Thiamine (MP Biomedicals)	4
ρ-Aminobenzoic acid (MP Biomedicals)	5
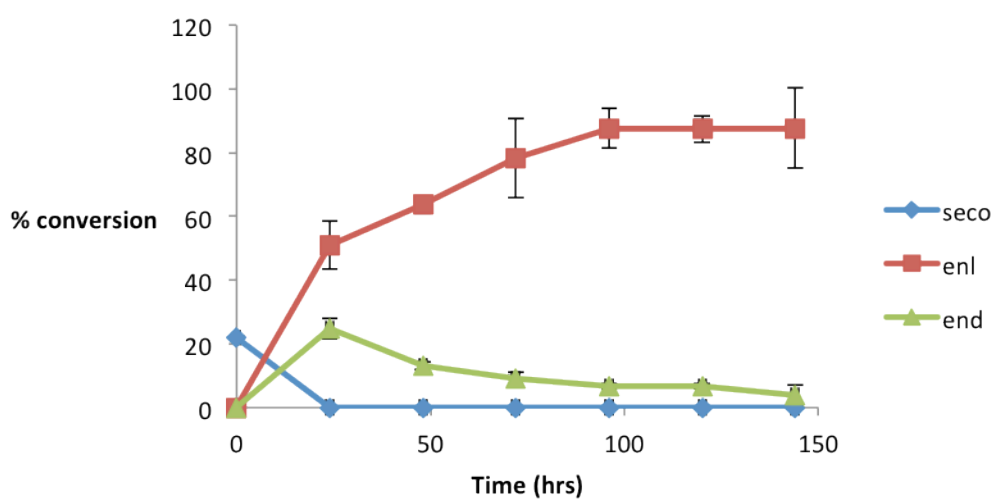
Figure 1. SDG to SECO, END, and ENL percent conversion measured by GCMS from daily aliquots of SDG-inoculated fecal suspension over a 7-day incubation of stool from participant 6105.

Table 3. SDG-metabolizing bacteria,[8,12,54,60-63,148] acetogenic bacteria,[73] methanogenic archaea,[82] bile acid dehydroxylators, [100,102-106] bile acid dehydrogenators,[108,109,107,110] sulfate-reducing bacteria,[155] and estrone to estradiol converters[111] searched for. Search terms, using custom R-scripts in QIIME and MetaPhlAn output, were genera and full species names when available. Exceptions to these search terms were "desul" for sulfate-reducers and "formatexigens", for *Marvinbryantia formatexigens*, since its genus had been renamed.

SDG-metabolizing bacteria	Acetogenic bacteria	Methanogens	Bile acid dehydroxylators	Bile acid dehydrogenators	Sulfate-reducing bacteria	Estrone to estradiol converters
Moryella	Clostridium coccoides	Methanobrevibacter smithii	Bacteroides	Eggerthella lenta	Desulfobacter	Akaligenes faecalis
Acetanaerobacterium,	Blautia hydrogenotrophica	Methanosphaera stadtmanae	Eubacterium	Clostridium perfringens	Desulomonas	Pseudomonas aeruginosa
Fastidiosipila	Marvinbryantia formatexigens	Candidatus Methanomethylophilus alvus	Eggerthella	Ruminococcus productus	Desulfotomaculum	Staphylococcus aureus
Streptobacillus	Ruminococcus	Candidatus Methanomassiliicoccus intestinalis	Clostridium scindens	Bacteroides fragilis	Desulfovibrio	Streptococcus faecalis
Lactobacillus	CS1 Van	Methanomassiliicoccus luminyensis	Clostridium hiranonis	Bacteroides thetaiotaomicron	Desulobulbus	
Enterococcus	CS3Glu		Clostridium hylemonae	Clostridium absonum		
Bacteroides fragilis	CS7H		Clostridium sordellii	Clostridium sordellii		
Bacteroides ovatus,	HA		Bacteroidetes thetaiotaomicron	Clostridium innocuum		
Clostridium cocleatum	S5a2h			Clostridium scindens		
Clostridium saccharogumia	VK64i			VPI 12708		
Clostridium ramosum	I52			Clostridium bifermentans		
Bacteroides distasonis				Clostridium leptum		
Bifidobacterium bifidum				Clostridium paraputrificum		
Bifidobacterium breve				Clostridium limosum		
Bifidobacterium catenulatum				Eubacterium aerofaciens		
Bifidobacterium longum				Clostridium hylemonae		
Bifidobacterium pseudocatenulatum				Clostridium scindens		
Butyribacterium methylotrophicum				Comamonas testosteroni		
Eubacterium callanderi				Pseudomonas		
Eubacterium limosum				Mycobacterium bovis		
Blautia producta				Mycobacterium tuberculosis		
Eggerthella lenta				Streptomyces exfoliatus		
Clostridium scindens				Escherichia coli		
Lactonifactor longoviformis				Brucella		
Eubacterium sp. ARC-2				Bacteroides fragilis		
Eggerthella sp. SDG-2				Burkholderia multivorans		
Clostridiaceae				Thermus thermophilus		
Ruminococcus				Bacillus thuringiensis		
Enterococcus faecalis				Escherichia coli		
Ruminococcus bromii						
Ruminococcus lactaris						

Table 4. KEGG IDs corresponding to KOs, modules, and pathways searched for in HUMAnN output. Modules and pathways that are starred were investigated as summed component KOs (as defined on the KEGG website: http://www.genome.jp/kegg/)

Enzyme, Module or Pathway	KEGG ID
Individual KOs	
β-glucosidases	K01188, K05350, K05349
β-glucoronidases	K01195
demethylases (acetogenic bacteria)	K00548, K15023, K14084, K16179, K14082
demethylases (methanogen - methanol)	K14080, K04480, K14081
corticosteroid 21-dehydroxylase	K00513
bile acid hydroxysteroid dehydrogenases	K05296, K00089, K00037, K04119, K00076, K05295, K00038, K1568, K05883
Modules	
Reductive acetyl-CoA pathway (Wood-Ljungdahl pathway)	M00377*
C21-Steroid hormone biosynthesis, progesterone => corticosterone/aldosterone	M00108*
C21-Steroid hormone biosynthesis, progesterone => cortisol/cortisone	M00109*
Methane oxidation, methanotroph, methane => formaldehyde	M00174*
Formaldehyde assimilation, xylulose monophosphate pathway	M00344
Formaldehyde assimilation, ribulose monophosphate pathway	M00345
Formaldehyde assimilation, serine pathway	M00346
Methanogenesis, methanol => methane	M00356*
Methanogenesis, acetate => methane	M00357*
Coenzyme M biosynthesis	M00358*
F420 biosynthesis	M00378*
Acetyl-CoA pathway, CO2 => acetyl-CoA	M00422*
Methanogenesis, methylamine/dimethylamine/trimethylamine => methane	M00563*
Methanogenesis, CO2 => methane	M00567*
2-Oxocarboxylic acid chain extension, 2-oxoglutarate => 2-oxoadipate => 2-oxopimelate => 2-oxosuberate	M00608
Pathways	
Methane metabolism	ko00680*
Starch and sucrose metabolism	ko00500*
Secondary bile acid biosynthesis	ko00121*
Steroid hormone biosynthesis	ko00140*
Primary bile acid biosynthesis	ko00120*
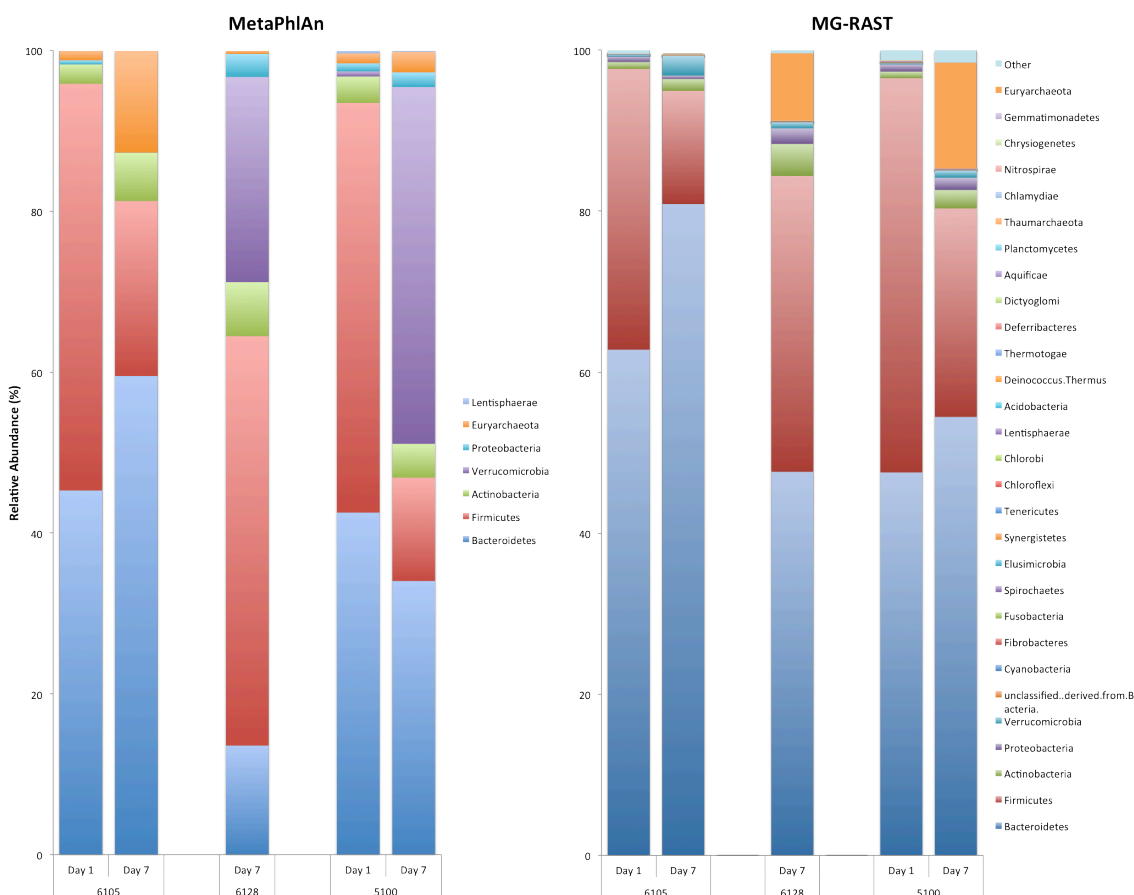
Figure 2. Relative abundance of phyla in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspensions from 3 different participants using MetaPhlAn or MG-RAST for classification.
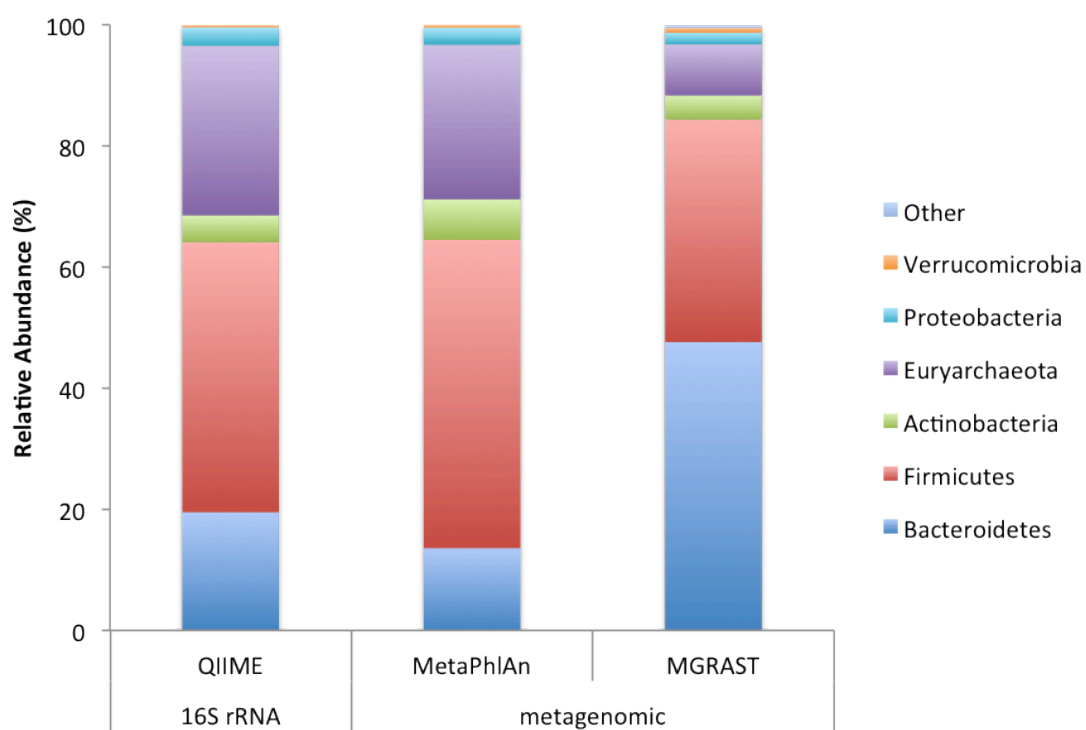
Figure 3. Comparison of dominant phyla relative abundance for the same sample (6128 Day 7) using 3 different sequencing method/classification tool combinations: V4 16S rRNA gene amplicon/in-house QIIME pipeline, metagenomic shotgun/MetaPhlAn, metagenomic shotgun/MG-RAST.
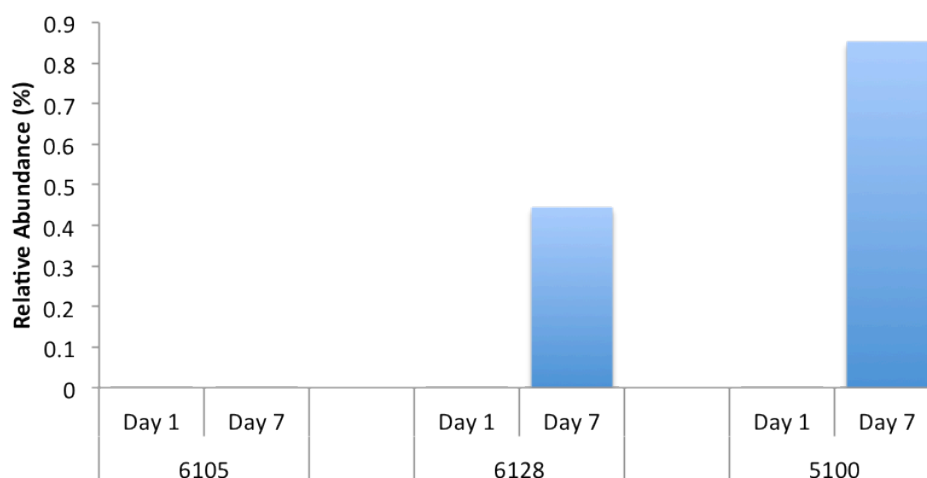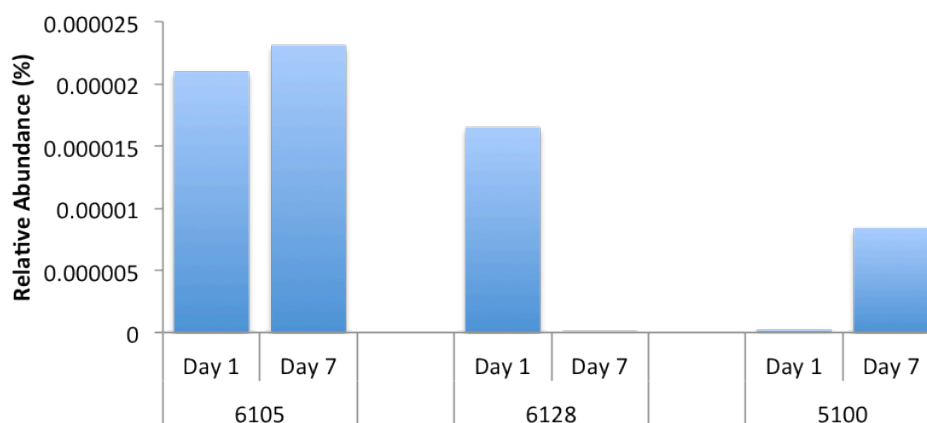
Figure 4. Relative abundance of modules hypothesized to be involved in lignan metabolism in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants calculated using DIAMOND and HUMAnN module output.
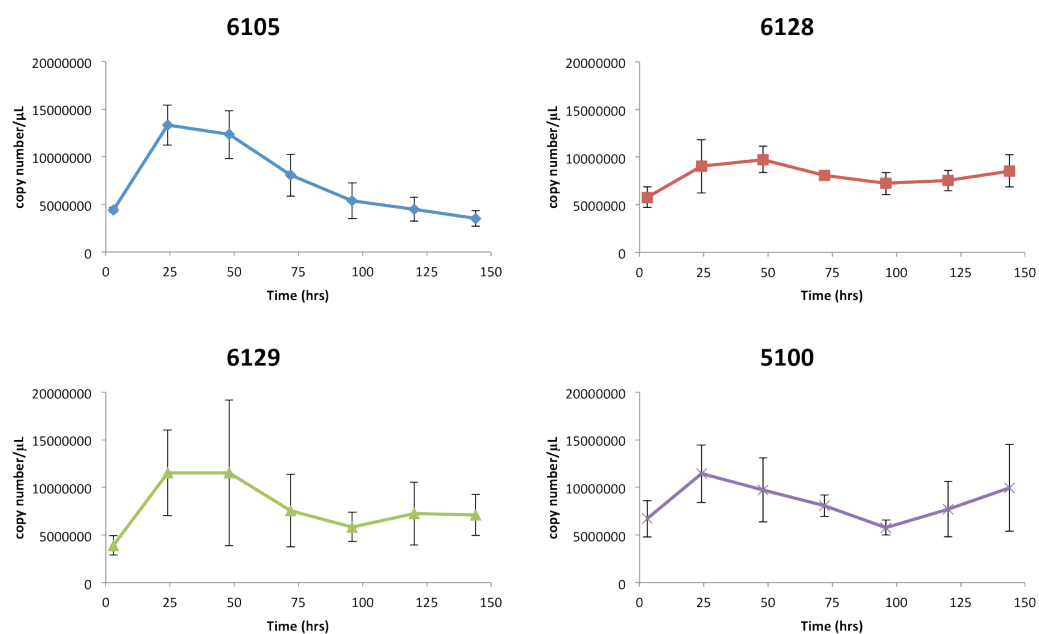
Figure 5. 16S rRNA gene copy number concentration estimated by qPCR from daily aliquots of fecal suspension over a 7-day SDG-inoculated incubation for four different participants.
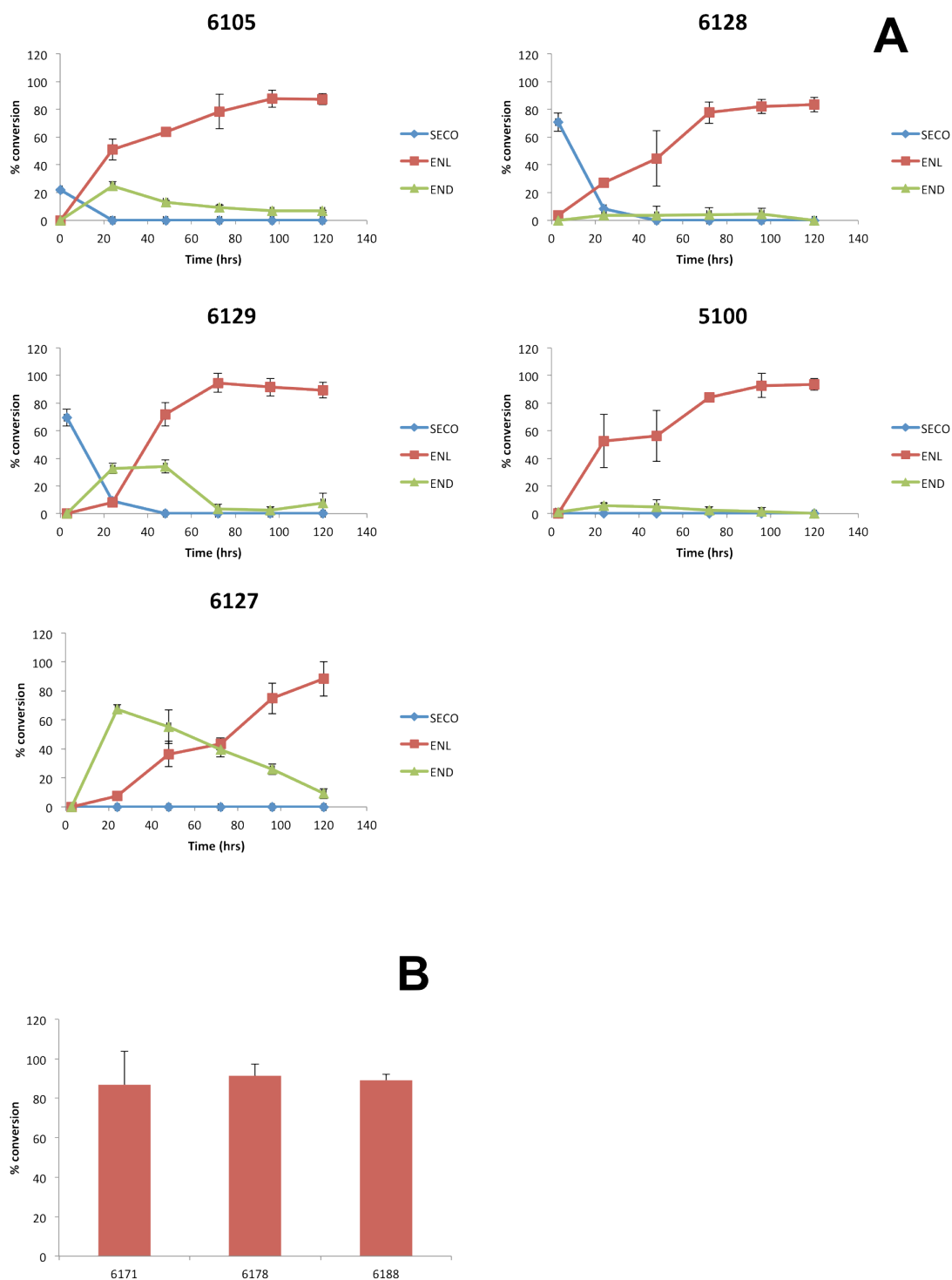
Figure 6. A) SDG to SECO, END, and ENL percent conversion measured by GCMS from daily aliquots of fecal suspension over 6 days of SDG-inoculated incubations from five participants. **B)** SDG to ENL percent conversion measured by GCMS from Day 6 fecal suspension of SDG-inoculated incubation from three remaining participants.
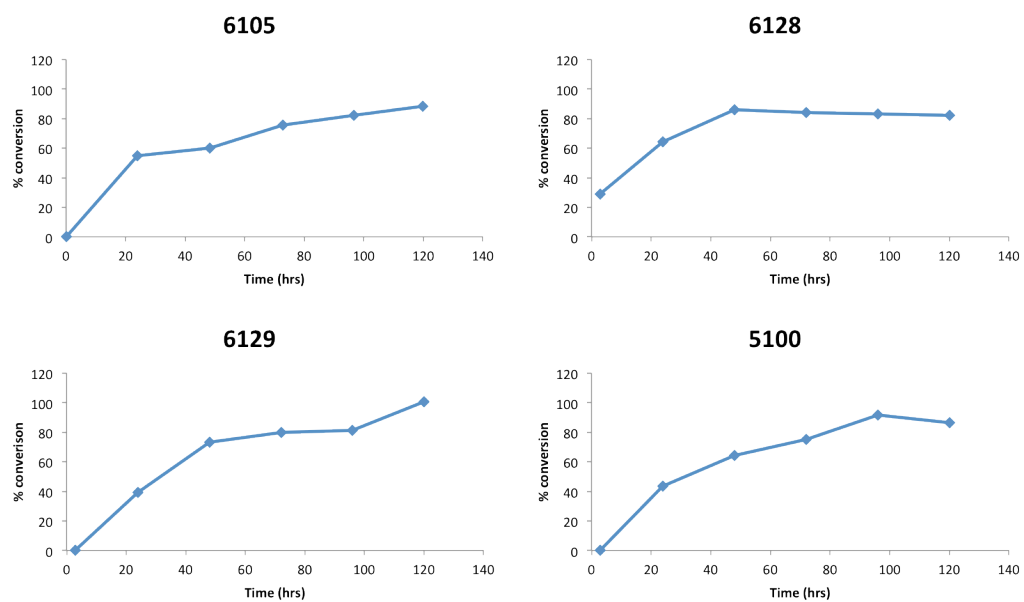
Figure 7. END to ENL percent conversion measured by GCMS from daily aliquots of fecal suspension over 6 days of END-inoculated incubations from four participants.
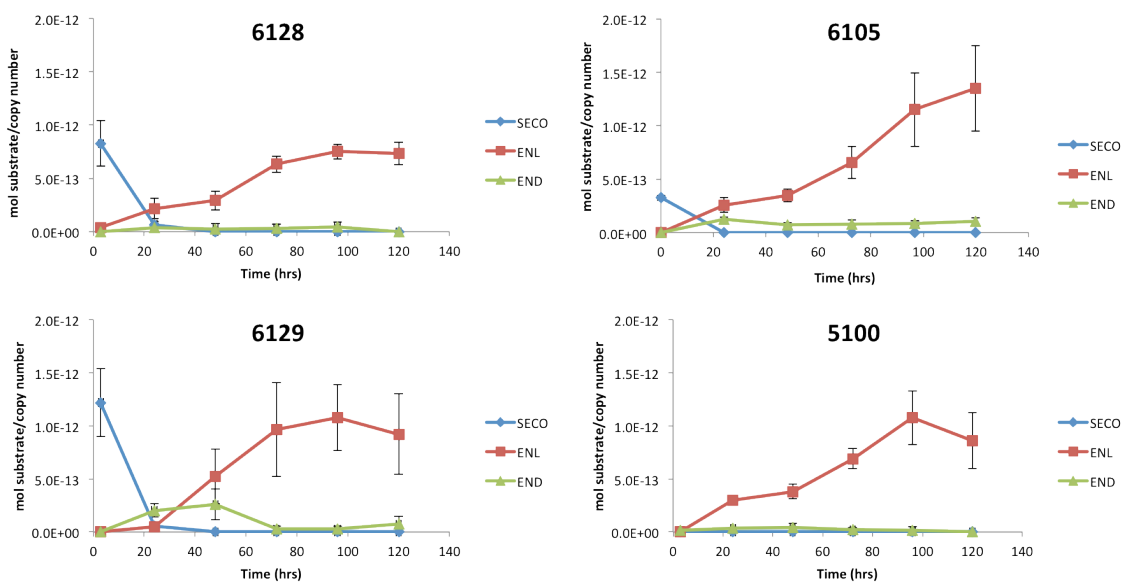
Figure 8. Moles of SECO, END, and ENL (measured by GCMS) divided by 16 rRNA gene copy number (measured by qPCR) from daily aliquots of fecal suspensions over 6 days of SDG-inoculated incubations from four participants.

Table 5. A) QC statistics of 16S rRNA gene amplicon and whole metagenome shotgun sequenced DNA extracted from Day 1 initial stool and final-day pelleted SDG-inoculated fecal suspension from 6 different participants. **B)** Post-QC statistics of 16S rRNA gene amplicon DNA from remaining 19 samples that were NOT USED in our analyses.

A

Participant #	DNA extraction timepoint (Day #)	16S rRNA After QC		Metagenomic QC				
					# of sequences		Length in basepairs	
		# of sequences	Length in basepairs (avg+/- stdev)	MG-RAST ID	Before QC	After QC	Before QC	After QC
5100	1	N/A	N/A	4629691.3	4,876,147	4,631,045	286 ± 82	235 ± 96
	7	N/A	N/A	4629690.3	6,177,329	5,819,240	296 ± 84	232 ± 105
6105	1	N/A	N/A	4551896.3	6,295,055	4,825,377	296 ± 122	267 ± 130
	7	N/A	N/A	4549297.3	5,145,601	3,627,976	306 ± 94	268 ± 114
6127	1	64,146	272 ± 8	N/A	N/A	N/A	N/A	N/A
	7	50,350	273 ± 9	N/A	N/A	N/A	N/A	N/A
6128	1	101,936	272 ± 9	4551897.3	6,440,154	5,112,128	318 ± 115	292 ± 126
	7	65,188	273 ± 10	4549299.3	6,380,077	4,694,486	305 ± 96	277 ± 112
6129	1	68,921	272 ± 11	N/A	N/A	N/A	N/A	N/A
	7	38,966	273 ± 11	N/A	N/A	N/A	N/A	N/A
6171	1	55,137	272 ± 9	N/A	N/A	N/A	N/A	N/A
	6	63,333	273 ± 9	N/A	N/A	N/A	N/A	N/A
6178	1	43,658	272 ± 8	N/A	N/A	N/A	N/A	N/A
	6	60,864	273 ± 9	N/A	N/A	N/A	N/A	N/A
6188	1	68,144	272 ± 10	N/A	N/A	N/A	N/A	N/A
	6	68,742	272 ± 8	N/A	N/A	N/A	N/A	N/A

B

| sample name	number of sequences
FFX.6129.END	63182
FFX.6135.END	58114
FFX.6135.stool	40963
FFX.6146.END	53835
FFX.6146.stool	65898
FFX.6155.END	87469
FFX.6155.stool	85840
FFX.6156.END	88464
FFX.6156.stool	53580
FFX.6171.END	55152
FFX.6171.END.TR	53100
FFX.6171.SDG.M	67507
FFX.6178.END	49154
FFX.6178.SDG.M	49913
FFX.6188.END	50671
FFX.6188.SDG.M	52383
FFX.FHC.A1	51383
FFX.FHC.A2	57573
FFX.FHC.A3	79335
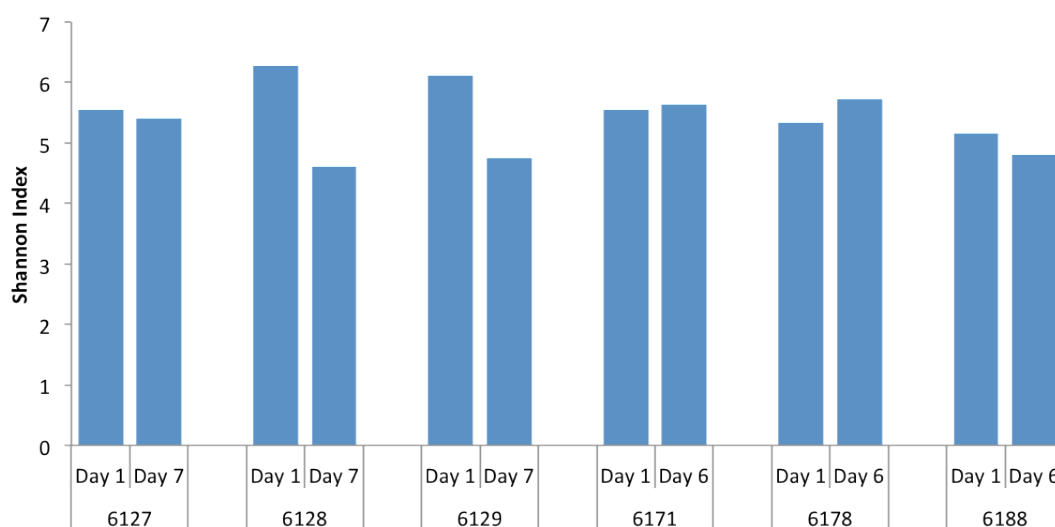
Figure 9. Alpha diversity (Shannon Index) of OTUs calculated using an in-house QIIME pipeline and derived from 16S rRNA amplicon sequences of DNA extracted from Day 1 initial stool and Day 6 or 7 pelleted SDG-inoculated fecal suspension from 6 different participants.
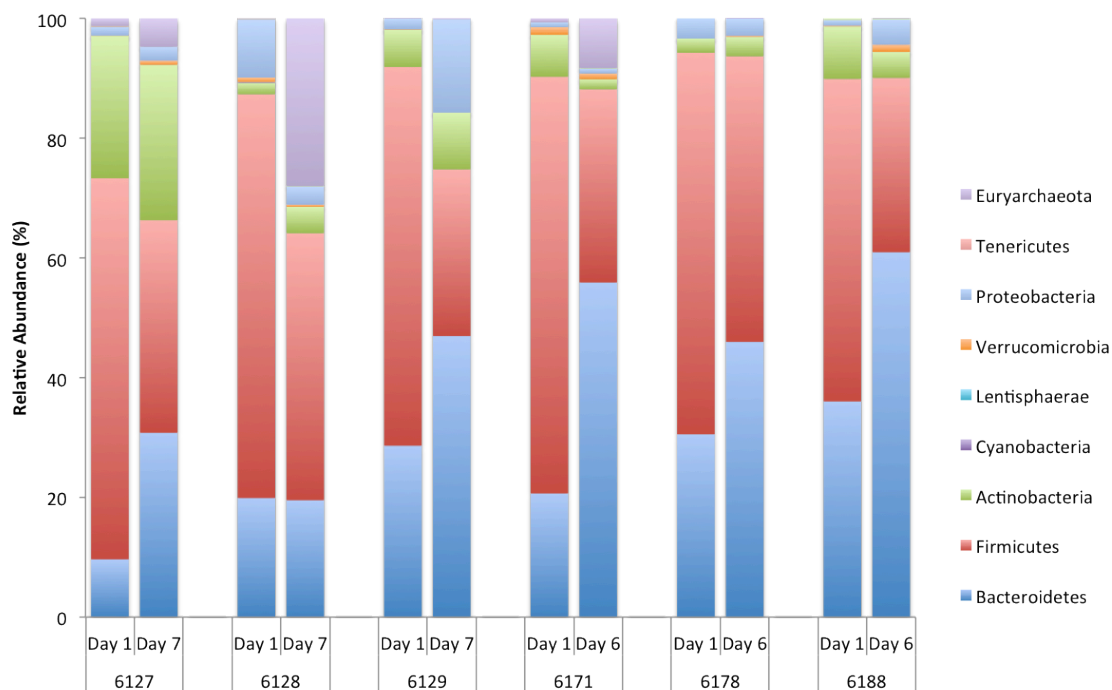
Figure 10. Microbial composition (phyla) of Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house QIIME pipeline classified V4 16S rRNA sequences.
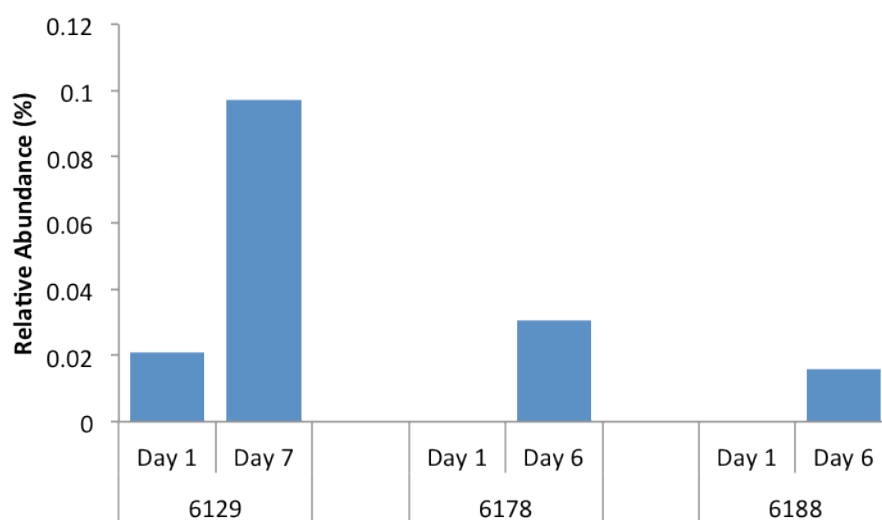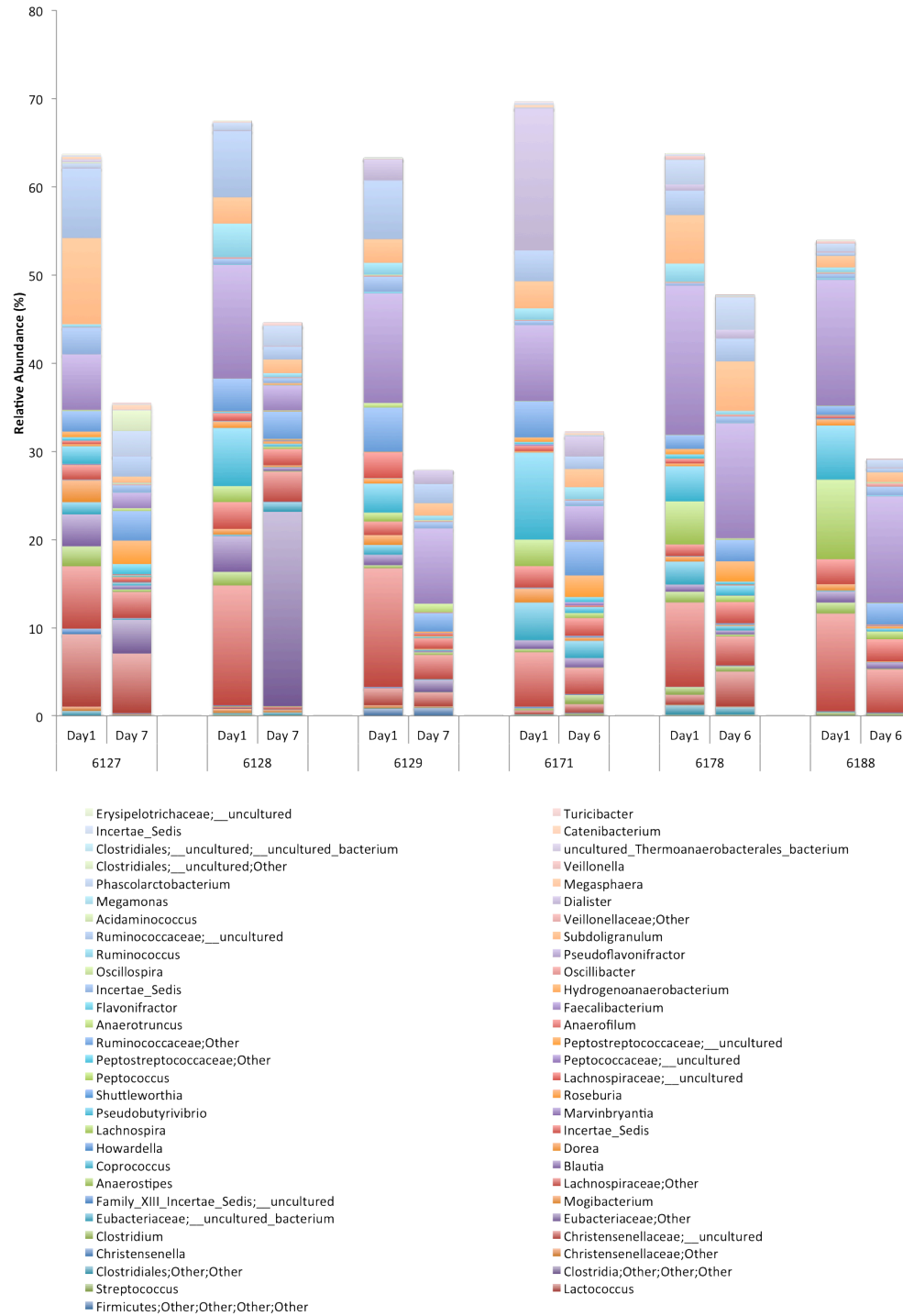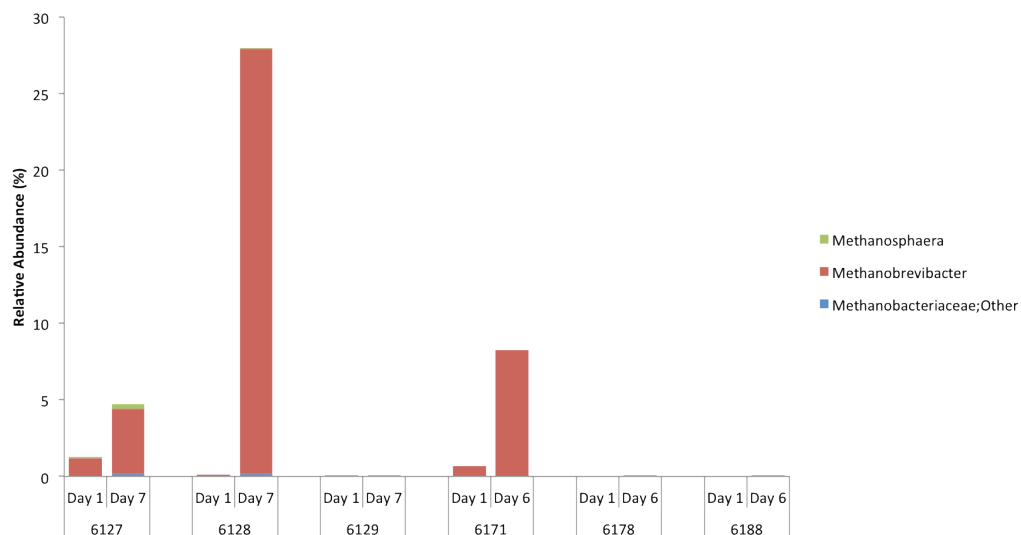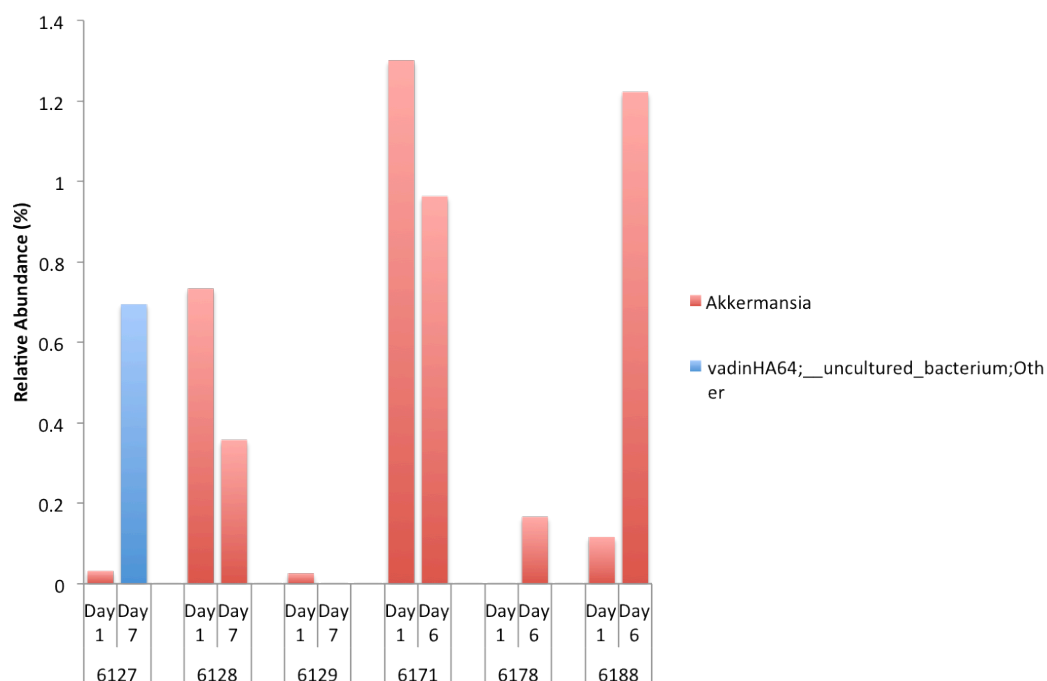

Figure 11. Euryarchaeota relative abundance in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from participants 6129, 6178, 6188 calculated using in-house QIIME pipeline classified V4 16S rRNA sequences.
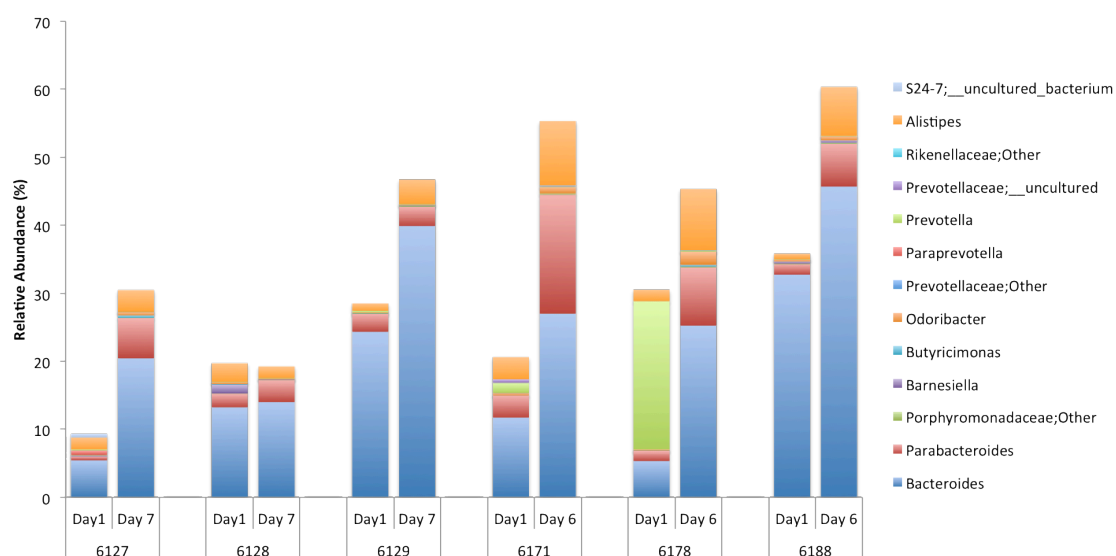
Figure 12. Relative abundance of genera from the phylum Bacteroidetes in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house QIIME pipeline classified V4 16S rRNA sequences.

Figure 13. Relative abundance of genera from the phylum Firmicutes in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using to in-house pipeline classified V4 16S rRNA sequences.

Figure 14. Relative abundance of genera from the phylum Euryarchaeota in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house QIIME pipeline classified V4 16S rRNA sequences.

Figure 15. Relative abundance of genera from the phylum Verrucomicrobia in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house QIIME pipeline classified V4 16S rRNA sequences.
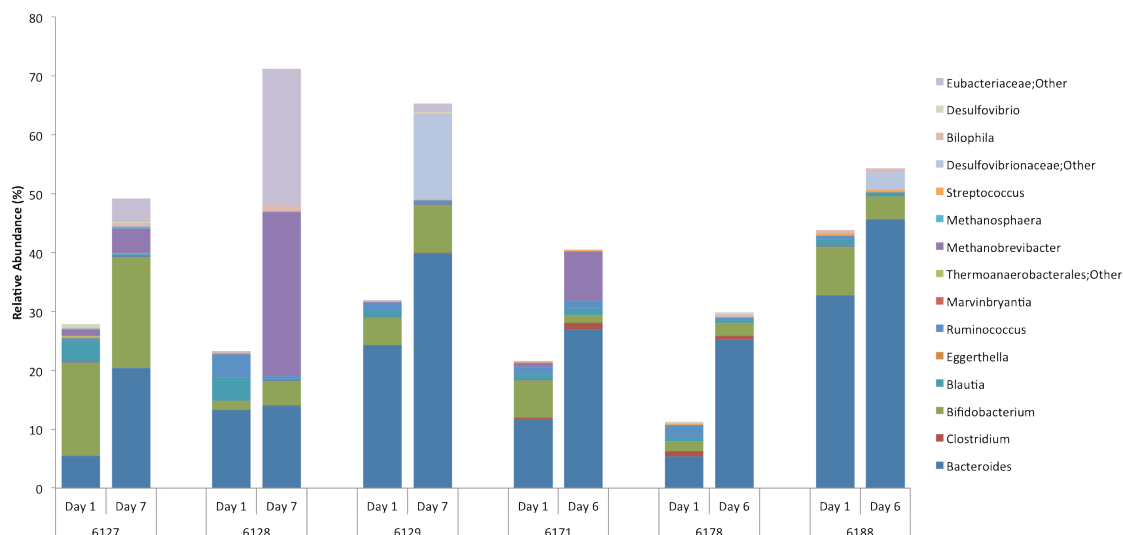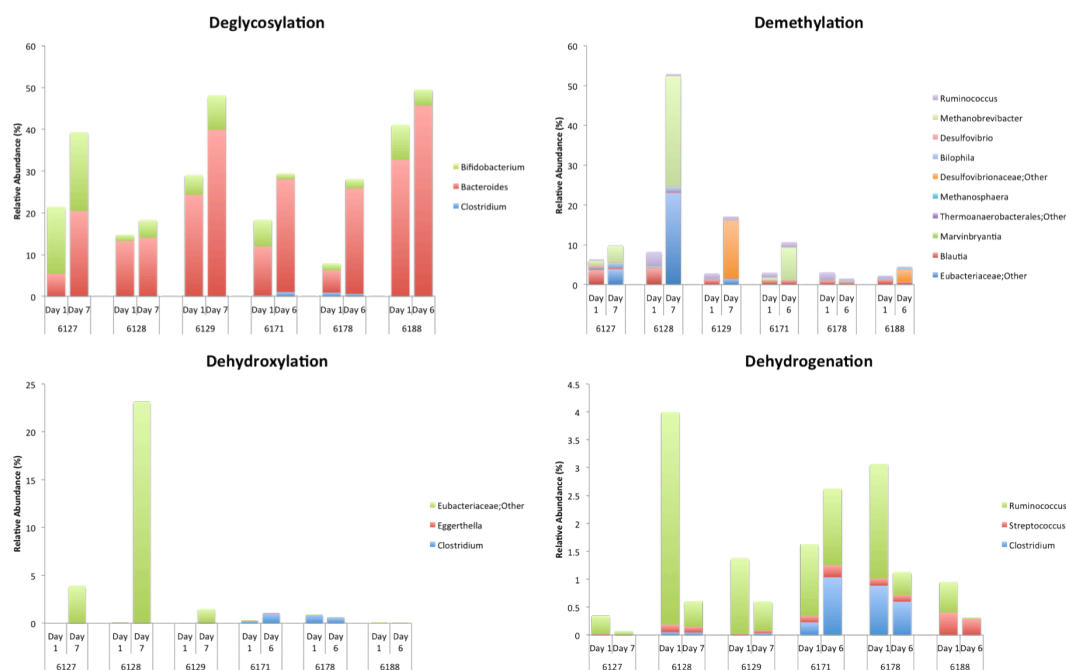
Table 6. Relative abundance of genera that were identified or hypothesized to be involved in lignan metabolism in Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house pipeline classified V4 16S rRNA sequences.

		6127		6128		6129		6171		6178		6188	
		Day 1	Day 7	Day 1	Day 7	Day 1	Day 7	Day 1	Day 6	Day 1	Day 6	Day 1	Day 6
	Bacteroides	5.45	20.44	13.25	14.01	24.35	39.90	11.73	27.01	5.34	25.26	32.75	45.70
	Clostridium	0.00	0.00	0.05	0.04	0.00	0.03	0.23	1.03	0.88	0.60	0.00	0.00
	Bifidobacterium	15.99	18.86	1.51	4.22	4.69	8.22	6.39	1.39	1.67	2.25	8.34	3.84
	Blautia	3.62	0.42	4.07	0.39	1.18	0.15	0.99	1.05	0.82	0.43	1.25	0.72
	Eggerthella	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.02	0.02
	Ruminococcus	0.32	0.07	3.82	0.46	1.36	0.52	1.29	1.37	2.05	0.41	0.55	0.03
Genus	**Marvinbryantia**	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
	Thermoanaerobacterales; Other	0.34	0.08	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	Methanobrevibacter	1.17	4.20	0.10	27.74	0.02	0.10	0.67	8.28	0.00	0.03	0.00	0.02
	Methanosphaera	0.08	0.33	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Streptococcus	0.02	0.00	0.14	0.11	0.02	0.05	0.11	0.22	0.13	0.12	0.40	0.29
	Desulfovibrionaceae; Other	0.03	0.03	0.05	0.04	0.04	14.52	0.03	0.02	0.05	0.06	0.04	3.05
	Bilophila	0.06	0.63	0.18	1.00	0.22	0.22	0.00	0.00	0.04	0.38	0.39	0.51
	Desulfovibrio	0.66	0.16	0.04	0.04	0.00	0.18	0.00	0.00	0.18	0.16	0.00	0.03

Figure 16. Relative abundance of genera that were identified or hypothesized to be involved in lignan metabolism in Day 1 initial stool and final-day pelleted SDG-incubated fecal suspension from 6 different participants calculated using in-house pipeline classified V4 16S rRNA sequences: **A)** grouped by sample **B)** grouped by sample and identified or hypothesized lignan metabolizing capability
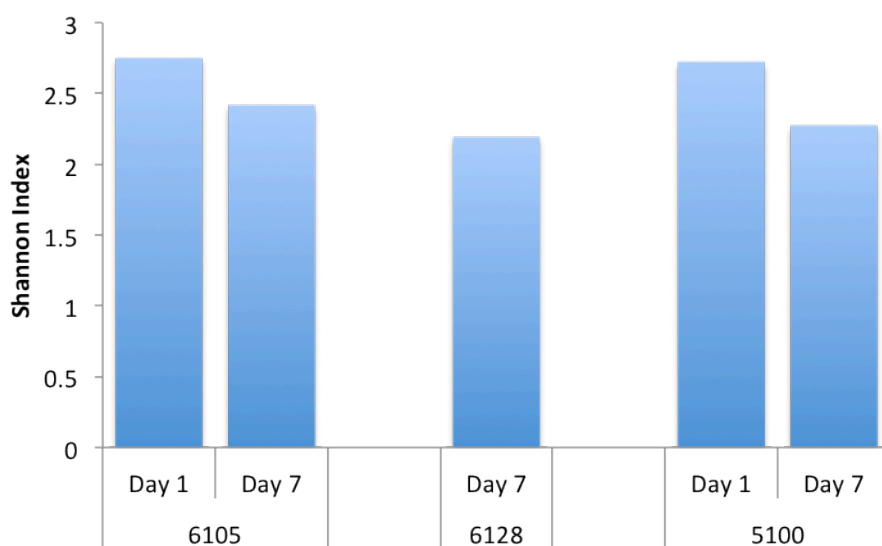
Figure 17. Species-level alpha diversity (Shannon index) calculated with the VEGAN R package from metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants based on MetaPhlAn output.
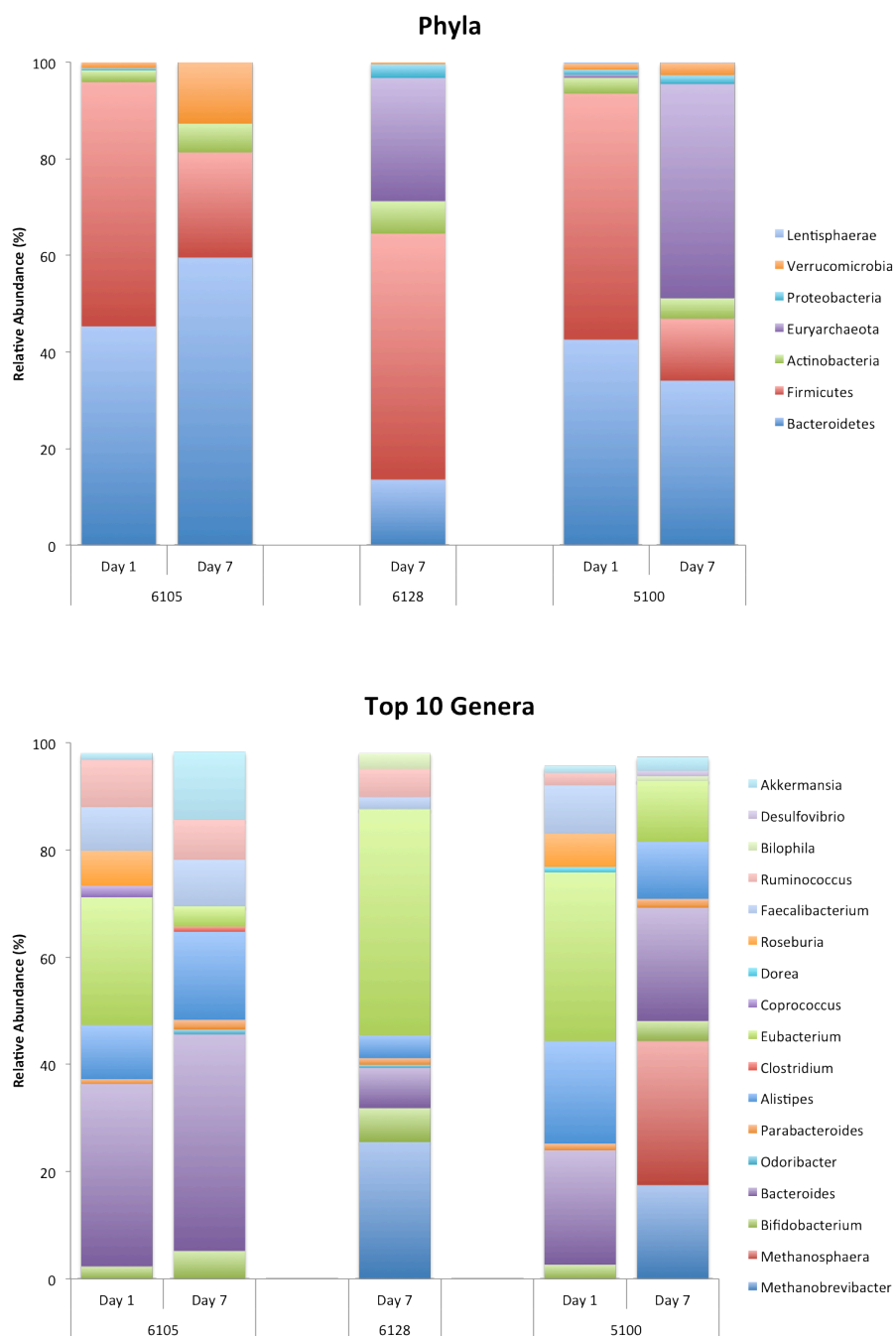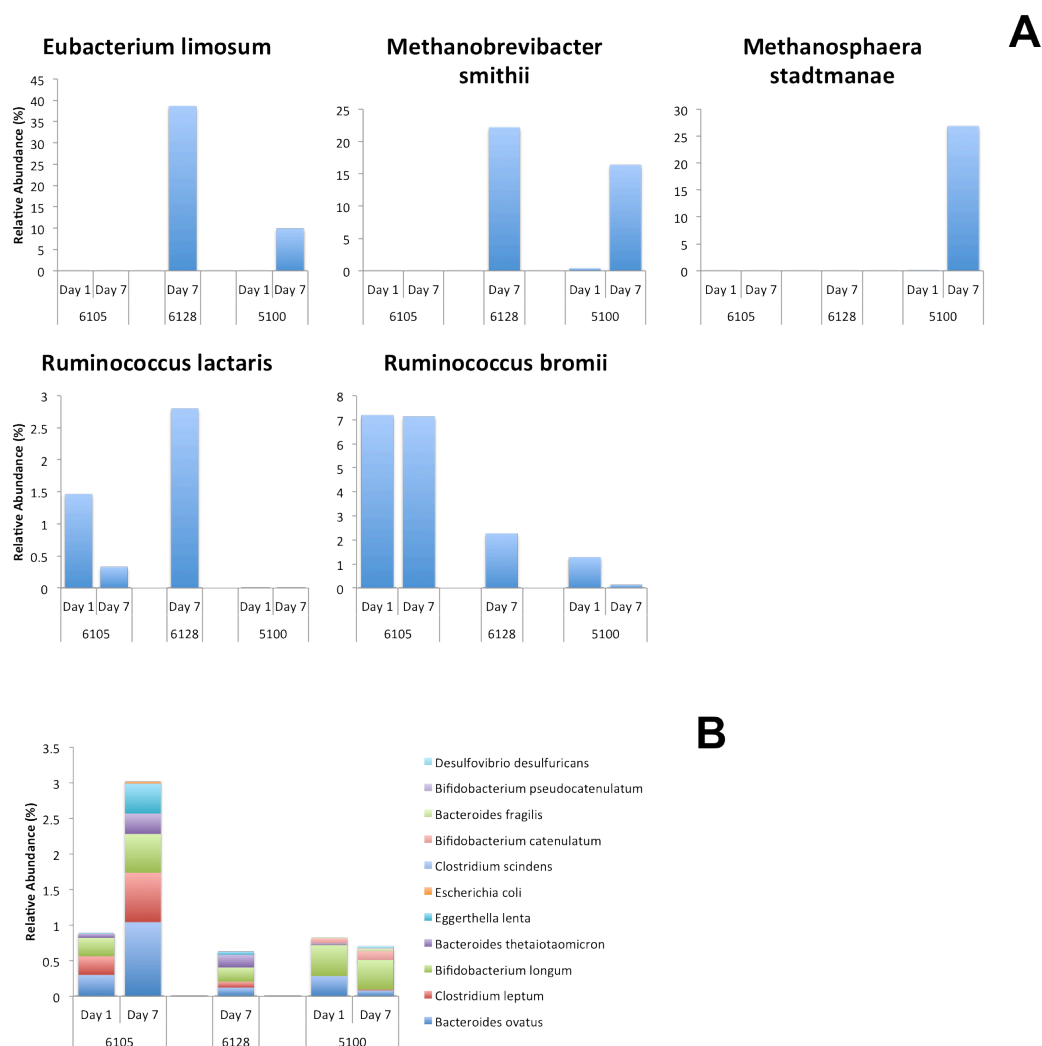
Figure 18. Phyla and top 10 genera relative abundance in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using MetaPhlAn for classification.
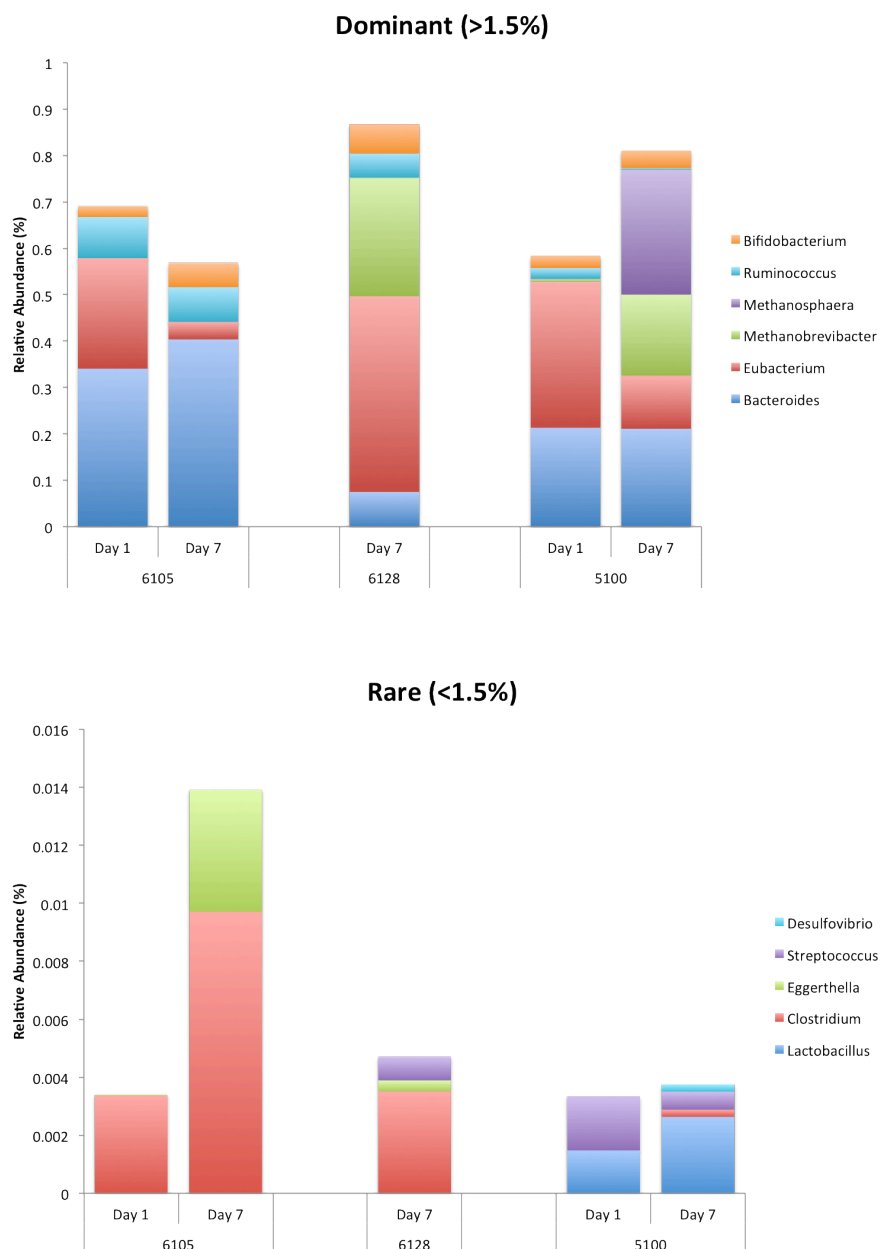
Dominant (>1.5%)

Rare (<1.5%)

Figure 19. Relative abundance of dominant (>1.5%) and rare (<1.5%) genera identified or hypothesized to be involved in lignan metabolism in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using MetaPhlAn for classification.

Figure 20. Relative abundance of **A)** dominant (>1.5%) and **B)** rare (<1.5%) species identified or hypothesized to be involved in lignan metabolism in metagenomic shotgun-sequenced DNA samples extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using MetaPhlAn for classification.
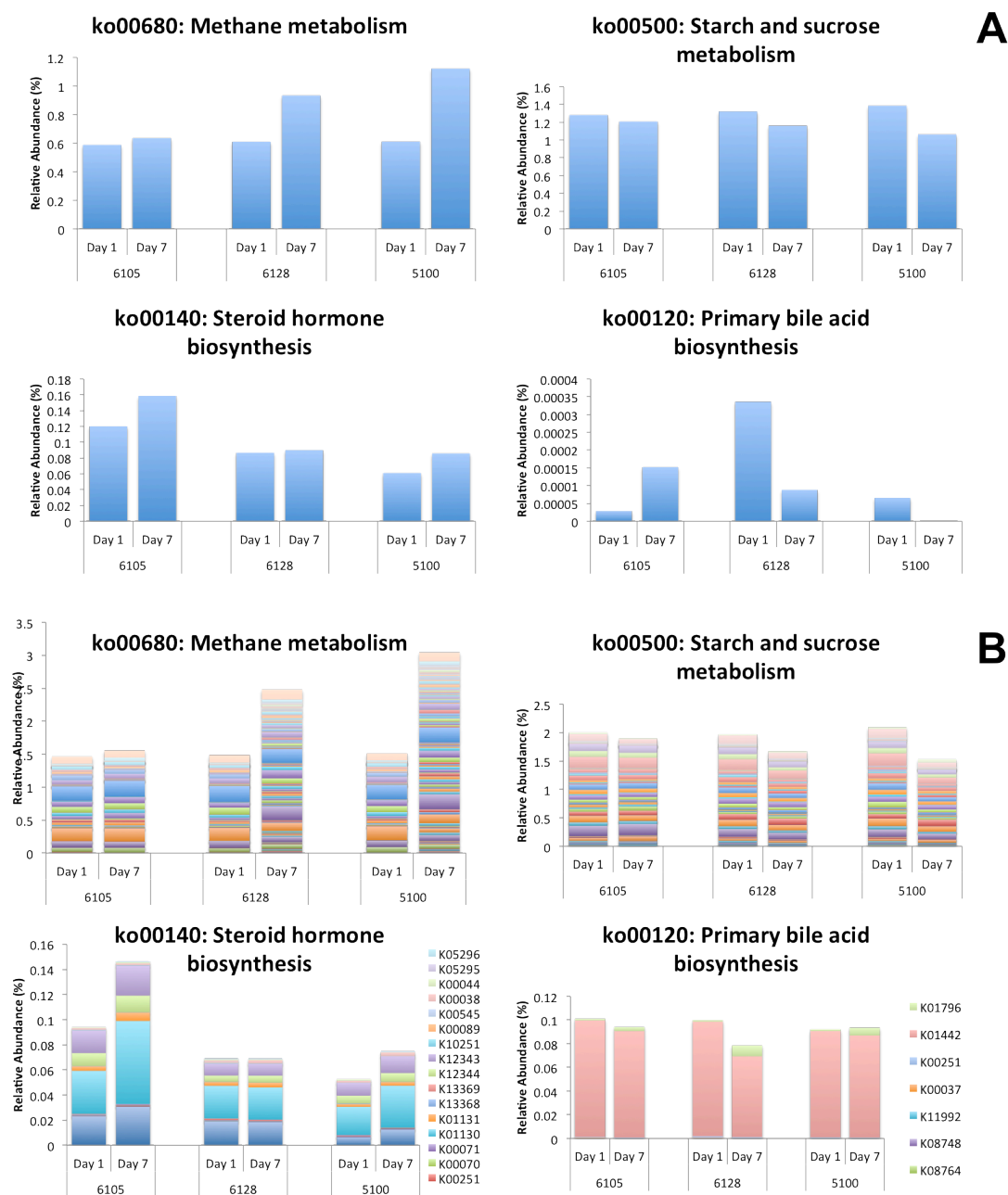
Figure 21. Relative abundance of pathways hypothesized to be involved in lignan metabolism derived from metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspensions from 3 different participants using DIAMOND and HUMAnN for functional assignment: **A)** derived from HUMAnN pathway output **B)** derived from HUMAnN KO output (Note: KOs corresponding to ko00680 and k000500 not listed because too numerous – over 80 per pathway detected in our samples).
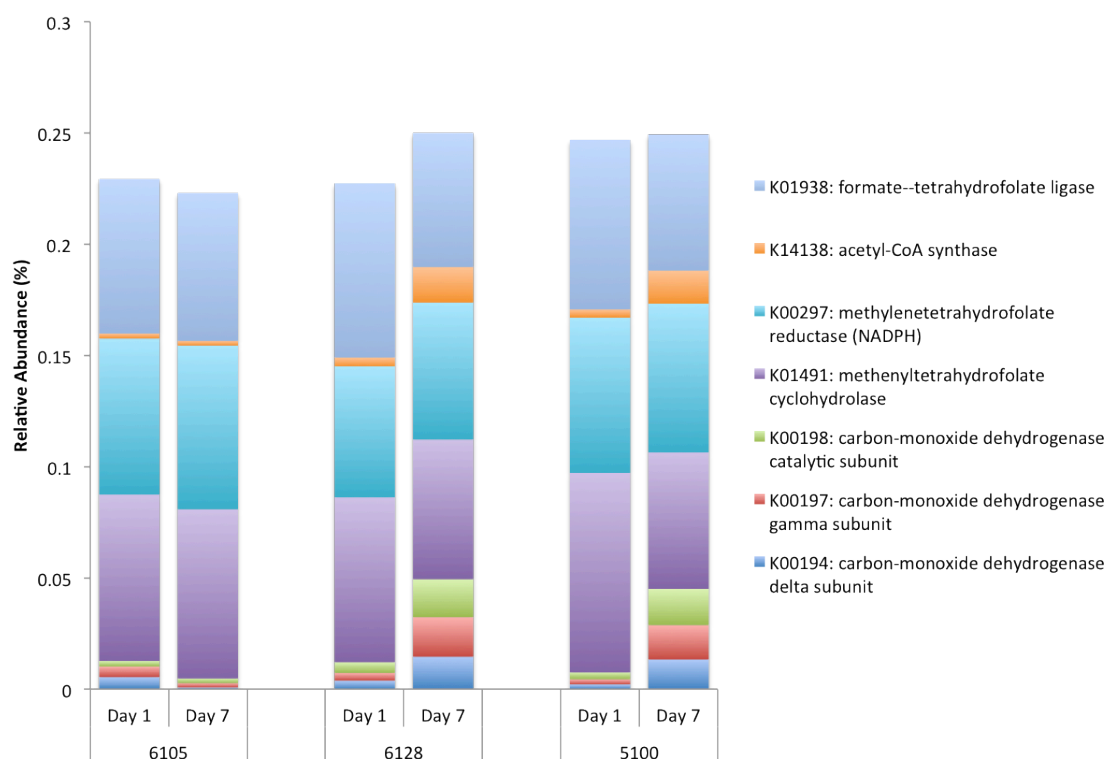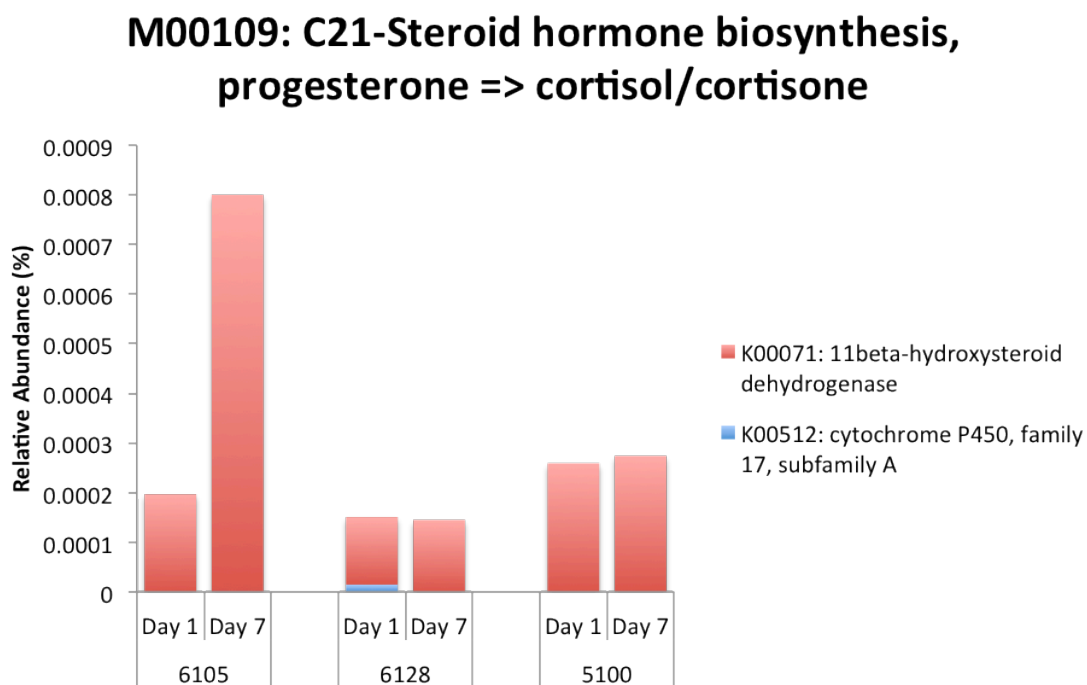
Figure 22. Relative abundance of component KOs in the reductive acetyl-CoA pathway (Wood-Ljungdahl pathway) - M00377 - in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using DIAMOND and HUMAnN for functional assignment.
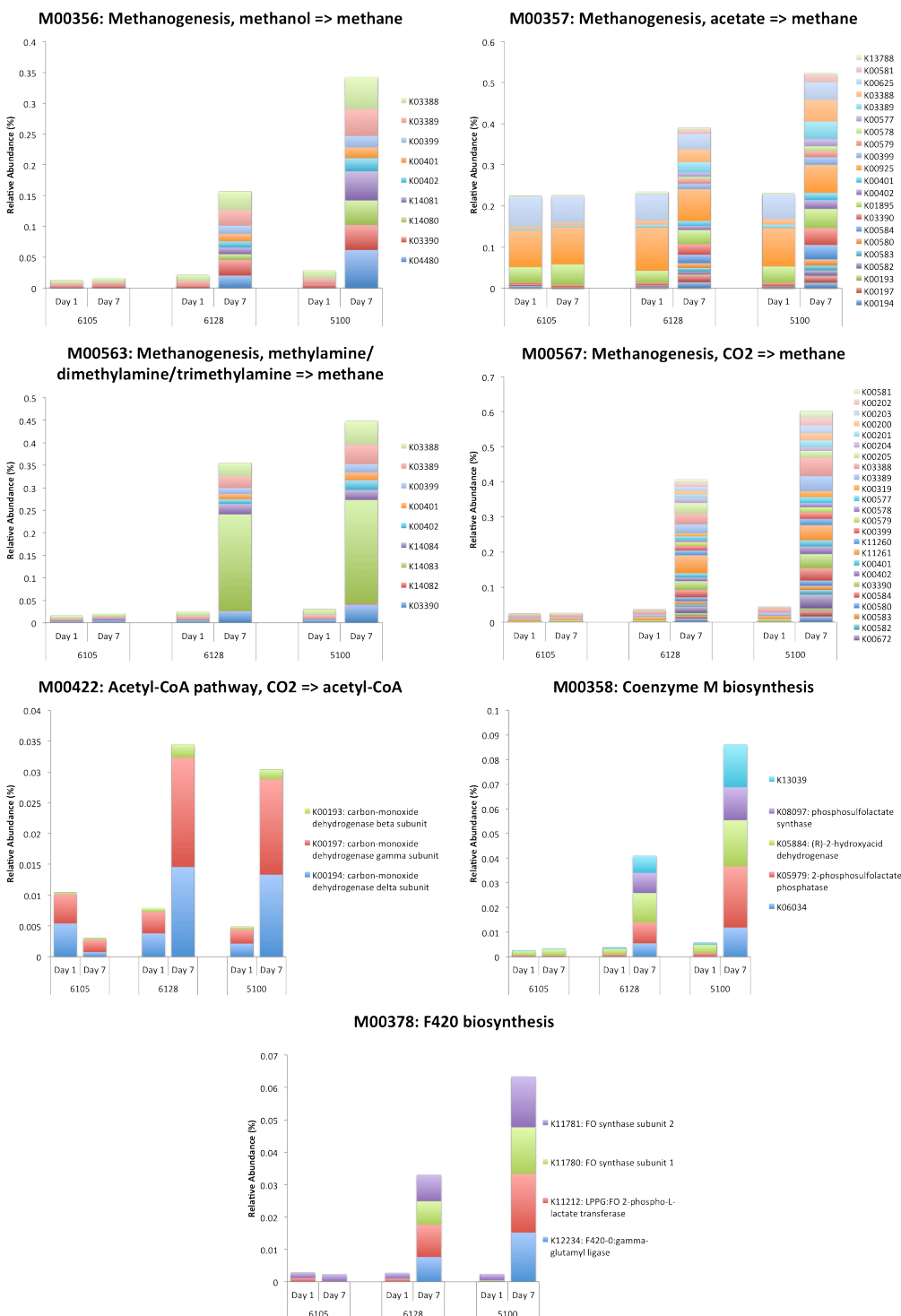
Figure 23. Relative abundance of component KOs of C21-Steroid hormone biosynthesis (progesterone => cortisol/cortisone) - M00109 - in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using DIAMOND and HUMAnN for functional assignment.

Figure 24. Relative abundance of KOs corresponding to modules unique to the methane metabolism pathway (ko00680) in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using DIAMOND and HUMAnN for functional assignment.
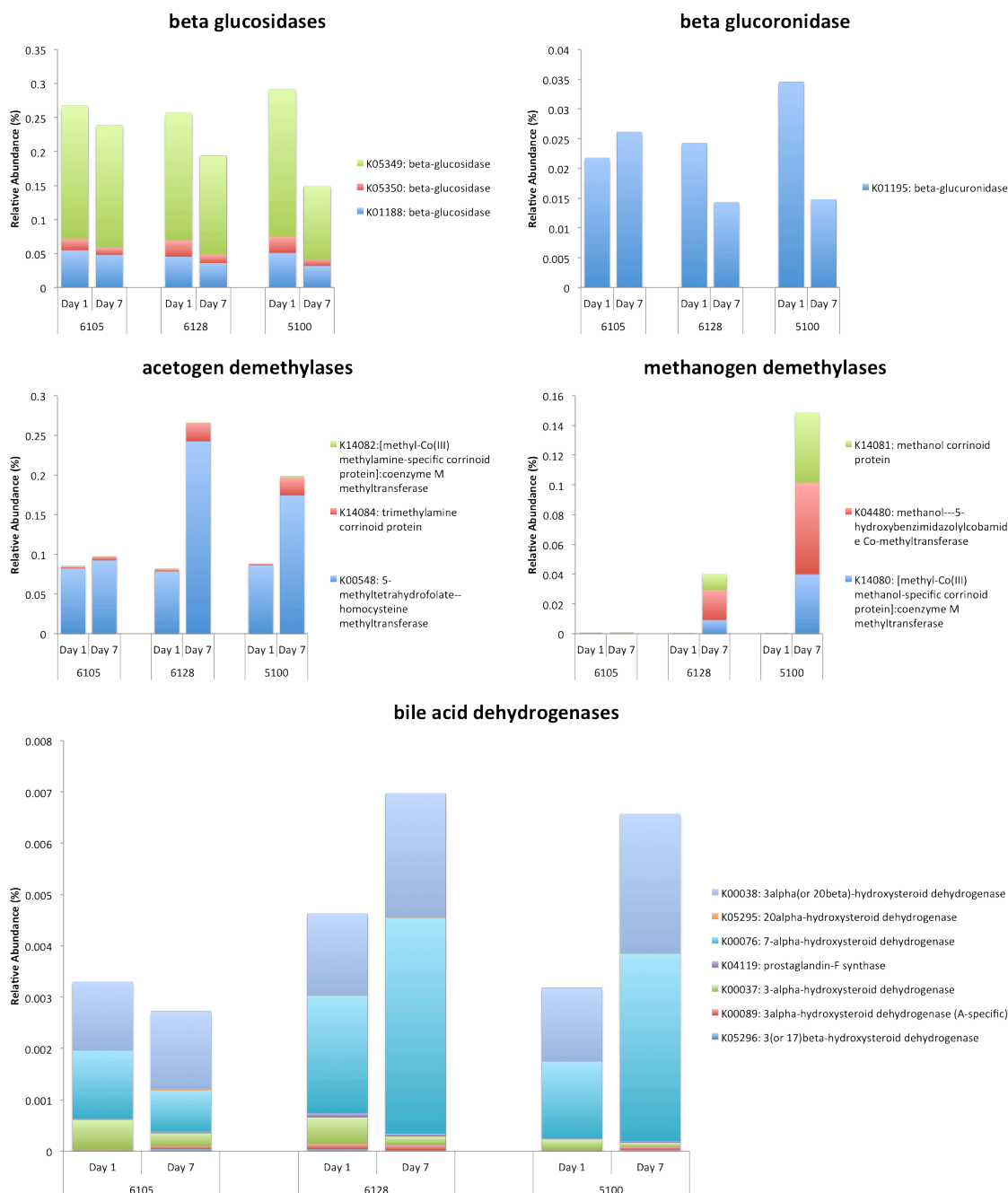
Figure 25. Relative abundance of KOs hypothesized to be involved in lignan metabolism in metagenomic shotgun-sequenced DNA extracted from Day 1 initial stool and Day 7 pelleted SDG-incubated fecal suspension from 3 different participants using DIAMOND and HUMAnN for functional assignment.
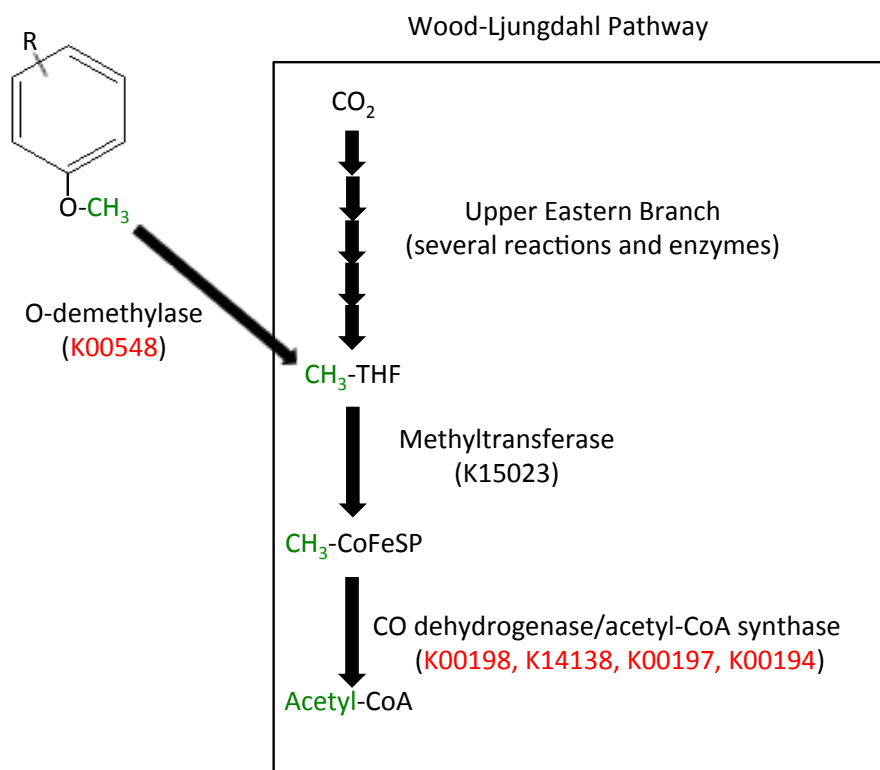
Figure 26. The demethylation of a methoxylated aromatic compound by an O-demethylase (K00548) belonging to *M. thermoacetica.* The methyl group is shunted into the Wood-Ljungdahl pathway. Specifically, it is transferred to the 5-methyltetrahydrofolate corrinoid/iron sulfur protein methyltransferase (K15023) and then to CO dehydrogenase/acetyl-CoA synthase, which forms acetyl-CoA. For incubations from 6128 and 5100, KOs highlighted in red had increased relative abundance in Day 7 pelleted SDG-incubated fecal suspension, when compared to Day 1 initial stool.