Lung Cancer Chemoprevention of Ginsenosides is Mediated by the Glycosidases Activities of the A/J Mouse Intestinal Microbiome Defined by Pyrosequencing

A Dissertation Presented to the

Faculty of the Department of Pharmacological and Pharmaceutical Sciences

College of Pharmacy, University of Houston

In Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

By

Tao Niu

June, 2015

Acknowledgements

I would like to take this opportunity to acknowledge many people who have made my stay at the College of Pharmacy memorable.

Special thanks are reserved for my advisor, Dr. Ming Hu, who has helped and encouraged me all through my graduate years and continues to support, inspire and guide me. I know I will continue to lean on you for supports wherever my career takes me to. Thank you for taking me as your student.

I would also like to sincerely thank one of my committee members, Dr. Joseph F. Petrosino and his graduate student, Diane L. Smith, for their continuous supports and invaluable inputs.

I would also like to sincerely thank my dissertation committee members, Drs. Diana Chow, Vincent Tam and Yi Wang, for their kindly supports.

I have to also thank Dr. Song Gao from University of Houston, and Dr. Hon-chiu Eastwood Leung from Baylor College of Medicine. They provided excellent comments and suggestions to this thesis work.

I also want to thank fellow graduate students, especially Kaustubh Kukarlni, Zhen Yang, and Baojian Wu, of the Department of Pharmacological and Pharmaceutical Sciences for all the supports you gave me.

Lastly, many thanks for the support from my parents.

Abstract

[***Purpose***] Red ginseng extract (RGE) has been reported to possess non-organ-specific preventative effects against several types of cancer, including lung cancer. The ginsenosides responsible for the activity of RGE are poorly defined as primary ginsenosides are inactive. Since most of the primary ginsenosides need the action of bacterial glycosidases to become active one, an important question is if functional activity of bacterial glycosidases would be altered following RGE administration. A more intriguing question is if and how RGE treatment can affect the growth of certain bacteria population in the intestinal microbiome, and whether alteration of such bacteria population would impact the chemoprevention efficacy of RGE. Therefore, the central hypothesis of this thesis is that the lung cancer chemoprevention of ginsenosides is mediated by the glycosidases activites of the A/J mouse intestinal microbiome defined by pyrosequencing. To test the central hypothesis, three specific aims are i) to characterize microbiota dependent metabolism of RGE in A/J mouse fecal lysate; ii) to determine the impacts of RGE administration on bacterial glycosidase activity and intestinal microbiome in A/J mouse; iii) to purifiy and identify bacterial glycosidase(s) from A/J mouse feces that catalyze the rate limiting step in the production of ginesnoside Compound K (one of the most active ginsenosides in RGE).

[***Methods***] The kinetics of microbiota mediated biotransformation of ginesnosides was characterized and kinetic parameters (metabolite formation rates) were determined. The anti-proliferative activity of ginsenosides was tested using the mouse lung cancer LM1 cells. Permeabilities of ginesnosides were also evaluated in Caco-2 cell monolayers.

The dose dependent enzymatic functions was tested by giving RGE orally at three doses (daily oral gavage at 5, 50, and 500 mg/kg respectively). The impacts of RGE administration on bacterial glycosidase and intestinal microbiome of A/J mice were evaluated by giving RGE daily to A/J mice for 7 days followed by measuring kinetic parameters (metabolite formation rates), diversity metrics (α diversity and β diversity) and relative abundance of bacteria in the gut microbiome. Bacterial glycosidases were enriched from A/J mouse feces by a classic protein chemistry approach. The identity of the enzymes was examined by LC-MS/MS analysis followed by gene synthesis, molecular cloning, and expression of the enzymes. The functional activity of the enzymes against ginsenoside Rd was tested.

[***Results***] **I**) Compound K exhibited higher anti-proliferative activity (IC50 ~13 μg/mL) and better permeability (1 $\times 10^{-6}$ cm/s) than primary ginsenosides. Primary ginsenoside Rb1 was converted to Rd, F2, and then Compound K by A/J mouse fecal lysate in a stepwise fashion. Formation of F2 from Rd (metabolite formation rate 0.09 ± 0.003, 0.09 ± 0.01 nmol/min/mg at 20, 5 μM substrate concentration respectively) was the slowest step in the biotransformation of Rb1 to Compound K. **II)** Bacterial glycosidase activity in response to RGE treatment exhibited a dose dependent manner and the optimal dose of RGE was found to be 50 mg/kg. Oral administration of RGE at 50 mg/kg for 7 days significantly enhanced glycosidase activity of A/J mice by a markedly change ($p<0.0001$) of metabolite formation rate in RGE treatment group. Noted inter-subject variability of glycosidase activity was observed among the A/J mice. While none of the mice in the Dose Response study exhibited changes in the microbiome following RGE treatment,

distinct changes in microbiome composition and richness were observed after 50 mg/kg RGE treatment in the RGE Interaction study. We also identified significant changes in relative abundance of the genus *Lactobacillus*, which contains species that can hydrolyze RGE.**III)** Specific activity of enriched enzymes increased from 0.757 to 27.5 μmol/mg/min after enrichment. The overall enrichment fold and yield was 36 and 5.81%, respectively. The SDS-PAGE and LC-MS/MS analysis revealed that one unique peptide NGVLFPR (mass=801.4497, z=2) correlating to bacterial glycosdases was found. Two bacterial glycosidases (gi: 501268188 and 147736211), when overexpressed were found to hydrolyze ginsenoside Rd to F2 and C-K.

[***Conclusion***] We have demonstrated that in vivo conversion of primary ginsenosides in RGE to the secondary and bioactive ginsenoside C-K was only mediated by microbial glycosidases. The formation of ginsenoside F2 from Rd was found, for the first time, to be the rate-limiting step in the biotransformation of Rb1 to C-K. Two bacterial glycosidases were enriched from A/J mouse feces and confirmed for the first time to hydrolyze ginsenoside Rd to F2 and C-K. Lung cancer chemoprevention of RGE is mediated by the bacterial glycosidases activites of the A/J mouse intestinal microbiota. Measurement of activity (formation of ginsenoside F2 from Rd) of such bacterial glycosidases may help differentiate potential responders of chemoprevention of RGE and non-responders, suggested by the large inter-subject variability of bacterial glycosidase activity.

List of Tables

List of Figures

List of Abbreviations

16s rRNA	16S ribosomal RNA
ANOVA	Analysis of Variance
APS	Ammonium persulfate
BCA	Bicinchoninic acid
BCRPs	Breast Cancer Resistance Proteins
BEH	Ethylene bridged hybrid
bp	base pair
CAI	Codon Adaptation Index
CE	Collision energy
CEP	Collision cell entrance potential
C-K	Compound K
COX-2	Cyclooxygenase-2
CXP	Collision cell exit potential;
CYP450	Cytochrome P450
DD	Double distilled
DMA	Desmethylangolensin
DME	Drug Metabolizing Enzyme
DMSO	Dimethyl sulfoxide
DP	Decluster potential;
DTT	Dithiothreitol
EC	Enzyme Commission
EGCG	Epigallocatechin-3-gallate
ELSD	Evaporative light scattering detector
ESI	Electrospray Ionization
FDA	Food and Drug Administration
FMO	Flavin-containing Monooxygenase
GFC	Gel Filtration Chromatography
GH	Glycoside Hydrolase
GI	Gastrointestinal
GRAS	Generally-recognized-as-safe
GSTs	Glutathione S-transferases

HBSS	Hank's Balanced Salt Solution
HDAC	Histone deacetylase
HMP	Human Microbiome Project
IARC	International Agency for Research on Cancer
IBD	Inflammatory bowel disease
IEX	Ion Exchange Chromatography
IS	Internal Standard
LLOQ	Lower limit of quantification
MRPs	Multidrug resistance proteins
MS	Mass Spectrometer
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
MW	Molecular Weight
N	RGE Non-responder
NATs	N-acetyltransferases
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NSAIDs	Nonsteroidal anti-inflammatory drugs
OD	Optical density
O-DMA	O-desmethylangolensin
OUT	Operational Taxonomic Unit
PAGE	Polyacrylamide gel electrophoresis
PBS	Potassium phosphate buffer
PEITC	Phenethyl isothiocyanate
PDA	Photodiode-arrayed
PPD	Protopanaxadiol
ppm	One part per million
PTM	Posttranslational modification
QC	Quality Control
R	RGE Responder
RGE	Red Ginseng Extract
ROS	Reactive oxygen species
SCFAs	Short chain fatty acids
SD	Standard Deviation
SDS	Sodium dodecyl sulfate

SN-38	7-Ethyl-10-hydroxy-camptothecin
SN-38G	7-Ethyl-10-hydroxy-camptothecin glucuronide
SULTs	Sulfotransferases
TBST	Tris-Buffered Saline and Tween 20
TEMED	Tetramethylethylenediamine
TMAO	Trimethylamine-N-oxide
TMA	Trimethylamine
TOF	Time of Flight
TRIS	2-Amino-2-hydroxymethyl-propane-1,3-diol
UGTs	Uridine diphosphate glucuronosyltransferases
UPLC	Ultra Performance Liquid Chromatography
WGS	Whole genome shotgun (WGS)

Contents

Chapter 1 Review of the literature

1.1. Introduction to the project

Red ginseng extract (RGE) as a chemoprevention agent has been examined over the past 30 years (Yun TK 1983; Yun TK 1987; SH 1988; Yun 1991; Yun TK 1995; Yan, Wang et al. 2006), and accumulated evidence supports the notion that RGE is a potent agent for chemoprevention of lung cancer. The primary, or most abundant, naturally occurring ginsenosides present in RGE include Rb1 and Rd (Li, Lee et al. 2008; Kong, Wang et al. 2009). However, secondary , or less abundant, ginsenosides such as Rh2, but not primary ginsenosides, demonstrate excellent inhibitory activities in both lung adenocarcinoma cells (Cheng, Yang et al. 2005) and in A/J mice (Yun, Lee et al. 2001). Mammalian cells do not express enzymes that hydrolyze ginsenosides (Henrissat 1991; Henrissat and Bairoch 1993; Henrissat and Bairoch 1996; Park, Yoo et al. 2010) suggesting that the *in vivo* transformation of secondary ginsenosides occurs via the action of the intestinal microbiota.

In the present thesis, we unveiled the reason of the *in vitro* and *in vivo* discrepancy of bioactivities of RGE and characterized the kinetics of microbiota mediated biotransformation of ginesnosides (**Aim I**). To elucidate the contribution of gut microbiota to the chemoprevention efficacy of RGE, 16s rRNA pyrosequencing together with ginsenoside hydrolysis assay were employed (**Aim II**). Lastly, a classic chromatographic "bottom up" approach was applied to enrich and identify bacterial glycosidases from A/J mouse feces, aiming at identifying bacteria species as agents for probiotic intervention (**Aim III**). The major purposes of the studies in the present thesis are to investigate

contribution of gut microbiota to the chemoprevention efficacy of RGE and identify bacteria species as lead probiotic intervention agents for chemoprevention of lung cancer using RGE.

1.2. Chemoprevention of lung cancer

Lung cancer remains the major cause of new cancer cases and deaths for both men and women in United States (Siegel, Ma et al. 2014). Epidemiological and laboratory animal model studies have demonstrated that smoking and environmental exposures to carcinogens are closely linked to increased lung cancer risk (Fielding 1985; Herzog, Lubet et al. 1997; Witschi, Espiritu et al. 1997; Witschi, Espiritu et al. 1997). For examples, active smoking accounts for the majority (85 - 90%) of lung cancer deaths (Centers for Disease, Control Prevention 2005). Despite the increased risk, many people are unable or unwilling to stop smoking. Additionally, advanced stage of disease at time of diagnosis, a poor response to traditional approaches against lung cancer, such as chemotherapy and surgery, and the aggressive biologic nature lead to high fatality rate. Therefore, there is an increasing interest in strategies for lung cancer chemoprevention to reduce the large number of smoking-caused cancer deaths, especially for former smokers. Chemoprevention involves the chronic administration of medication to reduce or delay the occurrence of malignancy and has recently drawn public attention. Chemoprevention should begin early, preferably before carcinoma in situ because lung cancer development is progressive, involving increasing genetic mutations over time that result in progressively higher levels of cell abnormalities. Intervention needs to be continuous due to residual or continued exposure to harmful substances, and hence

chemoprevention agents should have an outstanding safety profile and be amenable to oral administration.

1.2.1. Animal Model

The safety and efficacy of putative lung cancer chemoprevention agents are currently evaluated first using mouse lung tumor models (Malkinson 1992; Herzog, Lubet et al. 1997), because of similarities in the histopathology and tumor progression stages between mouse and human lung adenocarcinomas. In addition, several genetic changes including hypermethylation of p16, p53 mutation, and deletion of 3p, 9p, and 17p, which are frequently detected in lung hyperplasia and dysplasia (Sozzi, Miozzo et al. 1992; Bennett, Colby et al. 1993; Kishimoto, Sugio et al. 1995; Merlo, Herman et al. 1995; Belinsky, Nikula et al. 1998), are also found in both mouse and human lung cancer cells. These genetic changes are also considered primary candidates for intermediate biomarkers in clinical trials for lung cancer chemoprevention studies.

1.2.2. Chemoprevention agents

Among the more than 50 different chemoprevention agents tested, several groups of chemicals have shown significant efficacy against mouse lung tumor development. For example, RGE (and certain ginsenosides), polyphenols, soy isoflavones, isothiocyanates, glucocorticoids, and non-steroidal anti-inflammatory drugs (NSAIDs), are among the most effective compounds.

Recently, Dr. Yun's group at Korea Cancer Center Hospital performed a series of rodent studies using ginseng and showed that ginseng is a potent inhibitor of lung

carcinogenesis in a mouse lung tumor model using various carcinogens (Yun TK 1983; Yun, Kim et al. 1995; Shibata 2001; Yun, Lee et al. 2001). In addition, Dr. Yun's group also reported that two specific ginsenosides, Rg3 and Rg5, showed a statistically significant reduction of lung tumor incidence in mice (Yun, Lee et al. 2001). More recently, administration of RGE enriched water as the sole drinking source was shown to significantly reduce lung tumor multiplicity by 36% and tumor volume by 70% in A/J mice (Yan, Wang et al. 2006), and this result was confirmed in a follow-up animal bioassay. Furthermore, ginsenosides are approved for cancer treatment in several countries outside United States. For example, pure Rg3r is marketed as an anticancer drug called "Shen-Yi Capsule" in China (Yue, Wong et al. 2006) and Panagin Pharmaceuticals (http://www.panagin.com) are marketing proprietary ginsenoside formulations as approved anticancer drugs in Republic of Georgia and Republic of Uzbekistan.

Tea polyphenols, when given orally, were shown to inhibit lung carcinogenesis in rats and mice (Mimoto, Kiura et al. 2000; Liao, Yang et al. 2004). For example, 0.6% tea solution significantly reduced 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) induced lung tumor multiplicity (the number of tumors per mouse) in female A/J mice (Liao, Yang et al. 2004). Tumor multiplicity was also significantly reduced by adding EGCG in tap water (1 mg/mL) to cisplatin-treated mice (Mimoto, Kiura et al. 2000). Interestingly, the chemoprevention efficacy requires large consumption of tea polyphenols, possibly due to the limited bioavailability of tea catechins ((Zhu, Chen et al. 2000)).

Isoflavones, the major components present in soy beans, have been reported to inhibit lung cancer progression (Lee, Seo et al. 1991). Biochanin A, one of the well studied major isoflavones, showed a significant inhibitory effect on the incidence of tumor-bearing mice (12.5%, $P < 0.01$), as well as the mean number of tumors (0.13, $P < 0.001$), compared with the group treated with benzo(a)pyrene alone (Lee, Seo et al. 1991). Similar to tea polyphenols, isoflavones also suffer from poor bioavailability (Yang, Kulkarni et al. 2012).

Isothiocyanates such as phenethyl isothiocyanate (PEITC) are also effective against lung carcinogenesis induced by certain carcinogens such as NNK in rodent models (Morse, Amin et al. 1989), and clinical trials for PEITC is ongoing (Hecht, Kassie et al. 2009).

Glucocorticoids were found to be a strong inhibitor of carcinogenesis in lung in rodents (Wattenberg and Estensen 1996), but the use of glucocorticoids as chemopreventives is not currently feasible due to the commonality of systemic toxic effects.

NSAIDs may reduce lung cancer incidence and mortality in the general human population when used on a long-term basis (Jalbert and Castonguay 1992; Malkinson 1992; Duperron and Castonguay 1997). In A/J mouse model, it was reported that aspirin, sulindac, ibuprofen and piroxicam all reduced NNK-induced lung tumorigenesis, and the rates of reduction were 62%, 60%, 58%, 38% and 32%, respectively (Jalbert and Castonguay 1992; Duperron and Castonguay 1997). More recently, sulindac sulfone, a sulfone derivative of sulindac, was found to be a potent inhibitor of lung

tumorigenesis in mice, reducing tumor multiplicity by about 90% (Malkinson 1992). However, the clinical utility of NSAIDs is questionable since long-term use of these agents often leads to severe side effects in large percentages of the targeted population (Thun and Blackard 2009).

1.2.3. Efficacy of RGE in the Animal Model

Researchers have consistently observed a significant efficacy of RGE and selected ginsenosides in mouse models of lung cancer. In a recent report (Yan, Wang et al. 2006), 6-week-old A/J mice were randomized into 3 groups with 25 mice per group. Mice were given RGE in drinking water at dose levels of 2 mg/mL and 10 mg/mL.

One week after the ginseng treatment, all mice were given a single i.p. injection of benzo(a)pyrene (100 mg/kg body weight) in 0.2 mL tricaprylin. RGE treatment (in drinking water) was continued for 20 weeks after the i.p. injection of benzo(a)pyrene. The control group received deionized water. All animals were maintained on an AIN-76A purified diet. Fluids and food were available ad libitum. The experiment was terminated 20 weeks after exposure to benzo(a)pyrene and the lungs harvested. The results showed that administration of RGE water as the sole drinking source significantly reduced lung tumor total lung tumor volume by >70% in A/J mice. Tumor volume was 2.49 mm3 for control animals treated with benzo(a)pyrene, which was decreased to 0.74 mm3 with RGE treatment at 10 mg/mL in drinking water. These data represent a significant reduction in tumor volume in A/J mice. The observation is consistent with the finding of Panwar et al. who showed that the ginseng extract exhibited a 72% inhibition

of benzo(a)pyrene induced lung adenoma incidence in Swiss Albino mice (Panwar, Samarth et al. 2005).

1.2.4. Phytochemical Characterization of RGE

The major pharmacologically active components of RGE are triterpene saponins termed as ginsenosides. More than 40 ginsenosides have been isolated from ginseng, most of which belong to dammarane-type triterpene saponins with (20S)-protopanaxadiol (PPD) and (20S)-protopanaxatriol (PPT) as the aglycones (Figure 1). A HPLC-ELSD study of the RGE showed that most ginsenosides present in nature have multiple sugars (as many as 3 each) attached at C-3 or C-20 position of the correponding aglycones (Figure 2). Sugar moieties could be removed by chemical processing (heat, acid, etc) or by intestinal microflora (Niu, Smith et al. 2013). The resulting (secondary) ginsenosides such as Rh2, F2, Compound K etc, with smaller number of sugars attached, are less hydrophilic. These secondary ginsenosides (e.g., C-K and Rh2s) and their aglycones are desirable for cancer chemoprevention because they exhibit stronger anticancer activities (Yun, Lee et al. 2001; Cheng, Yang et al. 2005). Presumably, these active components are derived *in vivo* via the action of microbiota, since mammalian cells typically do not express glycosidases that will hydrolyze ginsenosides (Henrissat 1991; Henrissat and Bairoch 1993; Henrissat and Bairoch 1996; Park, Yoo et al. 2010).

Ginsenoside	R_1	R_2	MW
Rb_1	Glu^2-Glu	Glu^6-Glu	1108
Rd	Glu^2-Glu	Glu	946
F_2	Glu	Glu	784
C-K	H	Glu	622
Rg_3	Glu^2-Glu	H	784
Rh_2	Glu	H	622
PPD	H	H	460

Ginsenoside	R_1	R_2	R_3	MW
Re	H	Glu	Glu^2-Rha	946
Rg_1	H	Glu	Glu	800
Rh_1	H	H	Glu	638
F_1	H	Glu	H	638
PPT	H	H	H	476

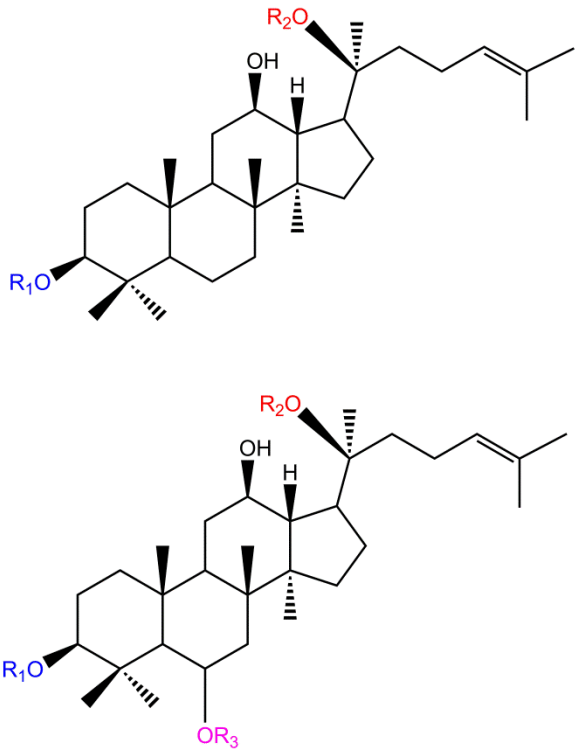


Figure 1. Structures of ginsenosides in PPD and PPT series

Glu: glucose, Rha: rhamnose. The superscript denotes the position of the hydroxyl group attached to the adjacent glucose.

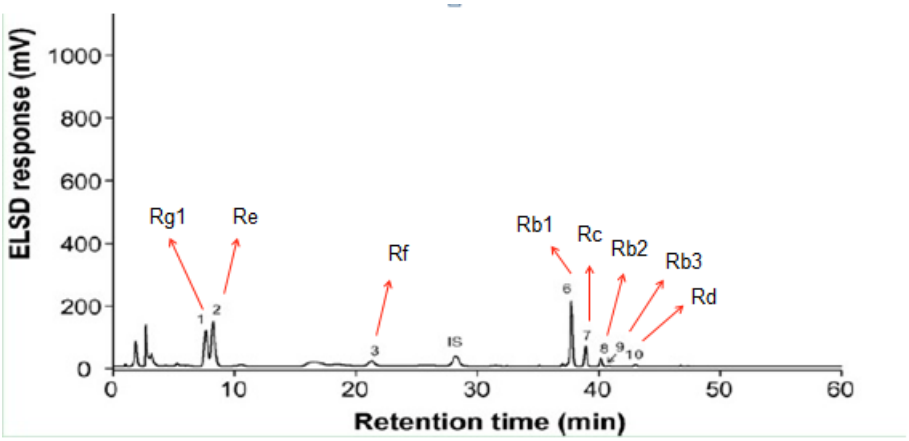


Figure 2. Representative HPLC fingerprint of red ginseng extract

Adapted from reference Kim et al. 2007

1.3. Microbiota, human health and cancer prevention

1.3.1. Introduction to the microbiota

Our human body harbors 100 trillion (10^{14}) of microbial cells, which outnumber human somatic and germ cells by one order of magnitude (Whitman, Coleman et al. 1998; Turnbaugh, Ley et al. 2007). In terms of gene number, the figure is even more surprising. The bacteria genome contains 3.3 million genes, dwarfing our own genome's 23 000 genes (Qin, Li et al. 2010).The microbes that reside in and on our body are termed as microbiota, and their genes are known as microbiome. The majority of the community remains unstudied since only a small amount of them can be cultured. Due to the recent availability the state-of-the-art culture-independent and inexpensive high throughput pyrosequencing technology, closer scrutiny of the microbiota has become reality (Riesenfeld, Schloss et al. 2004; Streit and Schmitz 2004). Studies in this area have been expanding exponentially and in 2008, National Institutes of Health (NIH) of United States initiated the Human Microbiome Project (HMP) with the aim of unveiling the association between human microbiome and health/disease. These studies have discovered that the human microbiota has a profound effect on the state of human health/disease (de Vos and de Vos 2012; Holmes, Li et al. 2012; Howitt and Garrett 2012; Vipperla and O'Keefe 2012; Eloe-Fadrosh and Rasko 2013).

The vast number and diversity of microbial functions, sometime similar to a multi-celled organ, perform a host of useful functions, for example, contribution to energy harvest from diet (Turnbaugh, Ley et al. 2006), prevention of the growth of harmful, pathogenic bacteria (Guarner and Malagelada 2003), and promotion of the development of immune

cells (Maynard, Elson et al. 2012). Furthermore, imbalance of gut microbiota is associated with several disease states including inflammatory bowel diseases and obesity (Clemente, Ursell et al. 2012; Damman, Miller et al. 2012; Kallus and Brandt 2012). It is also suggested that microbiota plays an important role in chemoprevention of colon cancer as difference in microbiota impact the production of short chain fatty acids (SCFAs), genotoxic hydrogen sulfide and carcinogenic secondary bile acids (Ridlon, Kang et al. 2006; Compare and Nardone 2011; Akaza 2012; Vipperla and O'Keefe 2012). Previously, intestinal microbiota was shown to impact the bioactivation of isoflavones, especially the production of more active compounds. Similarly, flavonoid glycosidases are often activated via the microbial glycosidases. More recently, microbial glycosidases have been shown to be involved in the activation of ginsenosides by removing sugars to produce more bioavailable and active compounds (Niu, Smith et al. 2013).

Gut microbiota may serve as a new target for developing noninvasive diagnostics and innovative treatments of diseases (Damman, Miller et al. 2012; Haiser and Turnbaugh 2012), although there is a lack of demonstration of causal relationship between microbiota and disease states. The detailed mechanism of microbiota perturbation in disease states also remains unclear. Therefore, the course of the current microbiota studies needs to be adjusted and refined, if they were to have direct impact on maintenance of human health and treatment of diseases.

Because the bulk of bacterial functions are related to metabolism and disposition of chemicals ingested or endogenous compounds, we will first review the process of

metabolism and disposition in humans enabled by human proteins (enzymes and transporters). The later part of this section will focus on microbial metabolism and then interaction between microbial metabolism and chemoprevention agents.

1.3.2. Drug metabolism and disposition in human

The drug metabolizing enzymes (DMEs) are a diverse group of proteins that are responsible for metabolizing a vast array of xenobiotic compounds including nutrients, micronutrients, drugs and xenobiotics, and various endogenous compounds such as steroids and prostaglandins. Drug metabolizing enzymes (DMEs) have been studied extensively in the last several decades (Guengerich 2008; Oakley 2011; Dong, Ako et al. 2012; Dong, Ako et al. 2012).

The drug metabolizing enzymes can be categorized into phase I enzymes such as cytochrome P450s and flavin-containing monooxygenase (FMO) that catalyze oxidation, reduction reactions, and etc. Phase II enzymes (conjugative enzymes) include uridine diphosphate glucuronosyltransferases (UGTs), sulfotransferases (SULTs), glutathione S-transferases (GSTs) and N-acetyltransferases (NATs). Phase II enzymes catalyze the transfer of a conjugation moiety from an endogenous donor molecule (i.e. uridine 5'-diphospho-glucuronic acid, adenosine 3'- phosphate 5'-phosphosulfate, or glutathione) to its substrate, which usually result in the production of more hydrophilic (water soluble) compounds, which are then eliminated through urine or bile (Yang, Zhu et al. 2012).

The extensive study of drug metabolizing enzymes greatly helps us understand the fate of xenobiotics in vivo (absorption, distribution, metabolism and excretion). However, the

pharmacokinetics of certain xenobiotics could not be explicitly interpreted without the involvement of intestinal microflora. For example, Liu et al proposed that the poor bioavailability of genistein was largely due to its enteric recycling (Liu and Hu 2002). Genistein was glucuronidated by UGTs in enterocytes and pumped back to the intestinal lumen via efflux transporters such as breast cancer resistance proteins (BCRPs) or Multidrug resistance proteins (MRPs) (Yang, Zhu et al. 2012). The microbial glycosidases then reconverted genistein glucuronide to its aglycone which underwent further metabolism in lower intestine. This study shows that the intestinal microflora play a critical role in the disposition of genistein, an important chemoprevention agent in prostate cancer prevention.

Because of the critical role played by the enzyme in drug disposition, it is generally acknowledged that the genetic polymorphisms of drug metabolizing enzymes and transporters contribute to the inter-subject variations with regard to the responses of drugs (Daly 2012). Kiyotani summarized that the individual differences of tamoxifen in clinical studies were largely due to the polymorphisms of CYP 2D6 (Kiyotani, Mushiroda et al. 2012). These studies diverge another research area that is known as pharmacogenomics since the last decade (http://www.nature.com/tpj/journal/v1/n1/index.html#ed). Recent studies suggest that the genotype of UGT1A is associated with the pharmacokinetics and toxicity of irinotecan. FDA also approved a genetic test to identify the genetic polymorphism of UGT1A (http://www.oscook.org/articles/fdacamptosar.pdf). However, subsequent studies did not support the correlation of UGT1A genotype and irinotecan pharmacokinetics and toxicity

(Nagar and Blanchard 2006). Thus, the inter-individual variability in drug responses is far less well understood.

1.3.3. Classification of Microbiome

Traditionally, humans are divided into four different groups by the major blood type (O, A, B and AB). Recently, however, some researchers have proposed to classify humans based on their enterotypes. Arumugam et al proposed that there were three distinct clusters of bacteria in the gut at the genus level, Bacteroides, Prevotella and Ruminococcus (Figure 3). In terms of metabolic functions, they speculated that the enterotypes had substrate preferences: Bacteroides to carbohydrates and proteins, Prevotella to mucin glycoproteins, and Ruminococcus to mucins and sugars (Arumugam, Raes et al. 2011). Thus, enterotypes among different people may respond differently to diet and drug intake, which contributes to the individual differences in the responses of drugs that cannot defined by pharmacogenomics. However, there are controversies in this proposed classification, and some researchers suggest that the work of Arumugam et al over-simplify the problem (Wu, Chen et al. 2011). Perhaps, more enterotypes will be established in the future to better classify humans based on their enterotypes.

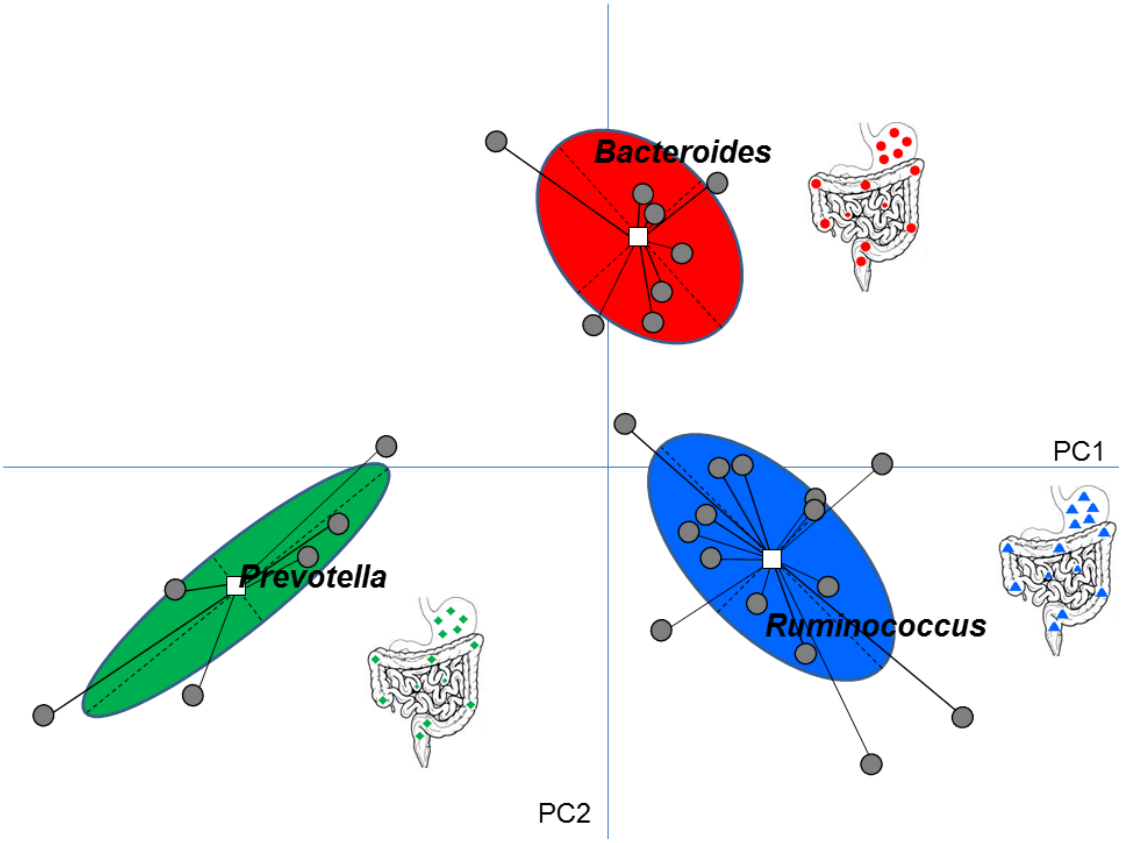


Figure 3. Three robust clusters that are termed as enterotypes from principle coordinate analysis and clustering*. Each of the enterotypes corresponds to a dominant genus (Bacteroides, Prevotella, and Ruminococcus). A cartoon to the right of the cluster in the same color denotes the enterotype.

Adapted from reference Arumugam, Raes et al. 2011.

1.3.4. Characterization of the Microbiome

Defining the core composition of microbiome in human and the inter-individual variability across the population are keys to the characterization of the microbiota. It appears that the intestinal microbiome composition reaches a general agreement at the phylum level (Figure 4) (Eckburg, Bik et al. 2005; Arumugam, Raes et al. 2011). However, large inter-individual variability at the species level of intestinal microbiome has been repeatedly reported. Less than 50% of species-level bacterial phylotypes are shared between twins (Turnbaugh, Quince et al. 2010). Moreover, there is an approximately ten fold difference in estimated microbial richness among different body sites, not to mention the considerable inter-individual variability at each of these sites (Costello, Lauber et al. 2009; Huse, Ye et al. 2012). These variations further complicate the characterization of the microbiota structure. It has been suggested that three human gut enterotypes are identified as indicated previously (Arumugam, Raes et al. 2011). However, the statement remains controversial. Two enterotypes (clusters of Bacteroides and Prevotella) rather than three were reported in a cohort study with 98 individuals in which Ruminococcus type was less well differentiated from Bacteroides, therefore, they were combined as one enterotype. This study also suggested that formation of these enterotypes was associated with long-term diet (Wu, Chen et al. 2011). Another 200 cohort study showed that most subjects fell in the two-enterotype profile at the genus level while others did not display distinct clustering. At the OTU (operational taxonomic unit, species distinction in microbiology) level, little or no segregation was observed (Huse, Ye et al. 2012). Therefore, it remains to be seen how valuable the enterotypes are in understanding the overall community structure and functions. With more data to come in

the near future, we will understand if gut microbiome fall into a defined number of community clusters or community gradients, and if so what are the core microbiome of different enterotypes.

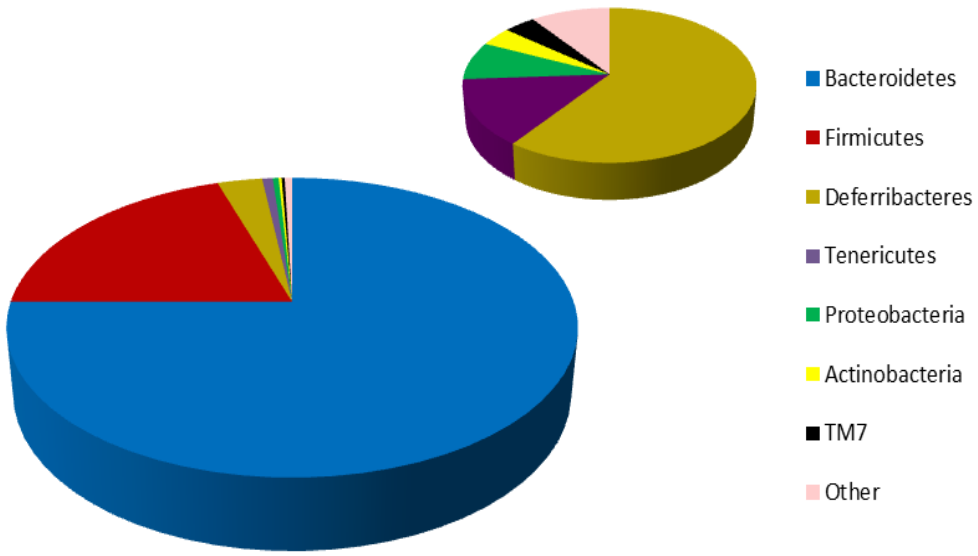


Figure 4. Intesitinal microbiota composition at the phylum level in A/J mouse, a representation of the core microbiota structure in the intestine.

The inset shows the microbiota composition excluding Bacteroidetes and Firmicutes.

1.3.5. Definition of a Healthy Microbiome

What is a healthy microbiome and what does the microbiome look like in disease states? Before answering this question, we may need to first scrutinize the microbiome from the ecological perspectives.

The first ecological perspective of microbiome is microbial diversity. Whittaker introduced three measurements of diversity: total species diversity of communities from different landscape or geographic locations (gamma diversity) is determined by alpha diversity and beta diversity. Alpha diversity measures the mean species diversity in sites or habitats at a more local scale (to quantify the richness of the species in a niche); while beta diversity assesses the differentiation among those habitats (to compare the diversity between different niches) (Whittaker 1972). Recently, microbial diversity studies started to examine the diversity by grouping sequences into OTUs (Bohannan and Hughes 2003). A thorough understanding of microbial diversity is crucial because it appears to enhance the resilience of the ecosystem (Elmqvist, Folke et al. 2003). Moreover, certain microbial diversity has been found to be associated with a variety of diseases. For example, increased alpha diversity is associated with a decreased risk of neonatal necrotizing enterocolitis (Wang, Hoenig et al. 2009); beta diversity is inversely correlated with the presence of Pseudomonas aeruginosa, which is a major pathogen of cystic fibrosis (Govan and Deretic 1996; Klepac-Ceraj, Lemon et al. 2010).

Driving forces for forming a defined community structure is the second ecological perspective used to describe the microbiome. Various factors could reshape microbiome, for example, invasive species, diets, host antibiotics microbiota interaction,

inter-community interaction. Invasive microbes are exogenous species that colonize dominantly of a habitat. The invaders could increase or reduce the abundance and repopulate or eliminate species of the community, thereby favorably (probiotics, live microorganisms that may confer a health benefit on the host) or adversely (pathogens) affect the habitats in the host. For example, pediatric antibiotic associated diarrhea could be alleviated by repopulating with probiotics (Johnston, Supina et al. 2007; Johnston, Goldenberg et al. 2011). On the other hand, chronic colonization of Pseudomonas aeruginosa in the lung plays a key role in the pathogenesis of cystic fibrosis (Govan and Deretic 1996).

The measurements used to assess microbiome state refer to resistance (the degree to which an ecosystem is altered as a result of an outside perturbation) and resilience (the capacity to recover after perturbation). However, static composition as the metric of health state of a community may simplify the microbiota definition. Alternatively, a series of complex and dynamic interactions (i.e. metabolic pathways) may serve as a better metric than static composition in defining the microbiota (Nicholson, Holmes et al. 2012).

1.3.6. Host Microbiota Interactions

Host microbiota interactions have a major impact on the microbiota structure and function, which will in turn affect the functions of the microbiota and most probably the host health. Germ-free rodents are the ideal model to study the function of microbiota and the impact of gut microbiota on host physiology. These animals have not been naturally colonized by microorganisms, and their microbiota can be controlled artificially by inoculating the animals with microorganisms of interest.

One of the major functions of microbiota is its contribution to energy harvest. Microbiota in the colon can metabolize food fiber into intermediate saccharides (oligosaccharides and monosaccharides) and SCFAs as the end products, which human cells cannot perform. This process yields energy and carbon source for microbial growth. SCFAs, primarily butyrate, propionate and acetate have profound effects on gut health, including serving as nutrient for colonic epithelial cells, host microbe signaling through G protein–coupled receptors, colonic pH control, gut motility, and etc. (Nicholson, Holmes et al. 2012; Tremaroli and Backhed 2012). Interestingly, it is suggested that the composition of the gut microbiota affects the fermentation of diets and thereby the amount and types of SCFAs (Samuel and Gordon 2006; Samuel, Shaito et al. 2008). Amount and types of SCFAs could impact human health. For example, acetate has been shown to increase cholesterol synthesis while propionate suppresses cholesterol synthesis. Therefore, a decreased ratio of acetate and propionate may reduce the risk of cholesterol related cardiovascular diseases (Wong, de Souza et al. 2006). Butyrate has been shown to exhibit protective effects against inflammation-mediated colon cancer via inhibition of histone deacetylase (HDAC) (Zimmerman, Singh et al. 2012).

Bile acids are steroid acids that are produced in the liver via the cytochrome P450-mediated oxidation of cholesterol. When bile acids are secreted into the lumen of the duodenum via the bile duct, they facilitate the formation of micelles so that dietary fats, fat soluble vitamins and cholesterol become soluble and absorbable. Primary bile acids complete an enterohepatic recycling circle by forming bile salts via conjugation with taurine and glycine, with conjugated bile acids being reabsorbed in the ileum and

transported back into the liver. However, bacteria in the ileum can hydrolyze the conjugated bile salts. The free bile acids can be further metabolized by bacteria into secondary bile acids. The enzymes responsible for the deconjugation are known as bile salt hydrolases (EC 3.5.1.24). Bacterial species, such as Bacteroides, Clostridium, and Lactobacillus have been found to express these enzymes (Ridlon, Kang et al. 2006). High level of secondary bile acids, produced solely by intestinal microbiota, may trigger gastrointestinal diseases such as colon cancer (Ridlon, Kang et al. 2006). This disease may be mediated via signaling through various nuclear receptors (Hylemon, Zhou et al. 2009).

Choline is another nutrient that must be consumed through diet. It is a key component of cell membranes and primarily comes from food such as eggs and fatty meats. Although choline is mainly metabolized in the liver, anaerobic microorganisms can also convert choline to its metabolite, trimethylamine (TMA). TMA is then further metabolized to trimethylamine-N-oxide (TMAO) by FMO in the liver. This metabolic activity is associated with multiple diseases, such as atherosclerosis, cardiovascular disease (Wang, Klipfell et al. 2011), and nonalcoholic fatty liver disease (Abu-Shanab and Quigley 2010; Mouzaki and Allard 2012). Because of a recently discovered gene cluster responsible for the microbial choline degradation (Craciun and Balskus 2012), modulation of the microbiota structure could have impact on host health through alteration of choline metabolism.

1.3.7. Association of Microbiome and Diseases

Changes in the composition of the gut microbiota have been associated with diverse, complex diseases including inflammatory bowel disease (IBD), metabolic disease and cancer.

IBD appears to be identified with inflammatory conditions at the colon and/or small intestine. The two principal disorders are Crohn's disease and ulcerative colitis. Gut microbiota is the key driving force of ulcerative colitis in murine inflammatory disease models (Strober, Fuss et al. 2002). Changes of Bifidobacterium and Bacteriodes and a reduction of SCFAs have been observed in human ulcerative colitis. Interestingly, increased level of SCFAs appears to be clinically favorable in the treatment of ulcerative colitis. Confirmative results in germ free mice have also been reported (Maslowski, Vieira et al. 2009). The composition and function of the microbiota in Crohn's disease are significantly altered, indicated by depletion of bacteria in phyla Firmicutes and Bacteroidetes (Frank, St Amand et al. 2007). Therefore, depletion of protective microorganisms and accumulation of disease driving bacteria may lead to IBD.

With the advancement of molecular biology, scientists have been hunting genetic origins of metabolic diseases for several decades. However, causal relationship between the detailed molecular mechanisms and development of metabolic diseases has yet to be established. Fortunately, the microbiome studies shed some light into this black box. Various studies demonstrated that the gut microbiota served as a signature of the metabolic phenotype (Ley, Backhed et al. 2005; Ley, Peterson et al. 2006). The alteration of the gut microbiota originating from the environment and diet could impact

microbiota composition as well as the host genome (Burcelin, Crivelli et al. 2002; Ley, Backhed et al. 2005). For example, half of the C57BL/6J mice became obese and diabetic when they were fed high-fat, carbohydrate-free diets for 9 month (Burcelin, Crivelli et al. 2002). Germ free mice do not develop diet-induced obesity since they do not possess a fasting-induced adipose factor that is suppressed by gut microbiota in normal conditions (Mandard, Zandbergen et al. 2006; Backhed, Manchester et al. 2007). Inoculating germ free mice with Bacteroides thetaiotaomicron, a bacterium that possess saccharolytic activity, allowed the mice fed with high fat diet to regain body weight and develop obesity (Samuel, Shaito et al. 2008). In addition, feeding with Western diets increases the ratio of Firmicutes to Bacteroidetes, which enhances polysaccharide fermentation and SCFAs production. This type of microbiome is termed as obesity-associated gut microbiome (Turnbaugh, Ley et al. 2006).

Cancer is often associated with unregulated cell growth. Genetic and environmental factors are the main reasons of carcinogenesis (Kamangar, Dores et al. 2006). Since the microbiota impacts numerous physiological functions related to cancer risk including control of food fermentation, nutrient production, epithelial cell proliferation and prevention of pathogenic invasion, recent studies had sought the correlation between the microbiota and cancer, especially colon cancer. Studies have shown that colon cancer patient showed an increased ratio of Bacteroides to Prevotella population in the gut (Sobhani, Tap et al. 2011). Enterotoxigenic Bacteroides fragilis may reshape the gut microbiome and induce colon tumors in intestinal neoplasia mice (Sears and Pardoll 2011). In addition, inflammation has been found to participate in all stages of cancer

development and progression. For instance, Helicobacter pylori, a carcinogen classified by the International Agency for Research on Cancer (IARC) has been shown to promote gastric carcinoma by the induction of gastric inflammation (Boleij and Tjalsma 2012).

In pursuit of microbiota-targeted early diagnosis and prevention of cancer, a key challenge is to establish causal relationships between alterations in the microbiome or function of the microbiota and diseases or etiology. Correlative studies are often not convincing enough to demonstrate the feasibility of microbiota-targeted early diagnosis and prevention strategy.

1.3.8. Microbiota-Targeted Early Diagnosis and Therapies

Information richness and ease of collection and analysis are the two key features of the microbiota, making it an outstanding origin of candidate for early diagnostics. In addition, bacterial cells respond to environment quickly and meticulously, which can be conveniently collected and analyzed in a large scale (Lemon, Armitage et al. 2012).

Microbiota has been shown to play a vital role in the occurrence and development of various diseases including cancer. Therefore, microbiota targeted therapies may open up a brand new approach in the treatment and/or chemoprevention of cancer.

Wallace et. al. provided an excellent example of microbiota targeted therapy (Wallace, Wang et al. 2010). Irinotecan is a commonly used chemotherapeutic agent against metastatic colorectal cancer with pronounced side effects (diarrhea and neutropenia). Irinotecan is converted to its active metabolite, SN-38, by enzymatic cleavage of the carbamate bond via carboxylesterases in vivo. SN-38 is absorbed in the intestine and

glucoronidated by UGT in the liver. The inactive glucuronide (SN-38G) is then excreted via bile ducts back into the intestine where the bacterial β-glucuronidases produced by the intestinal microbiota cleave the glucuronosyl bond and release SN-38 (Figure 5).

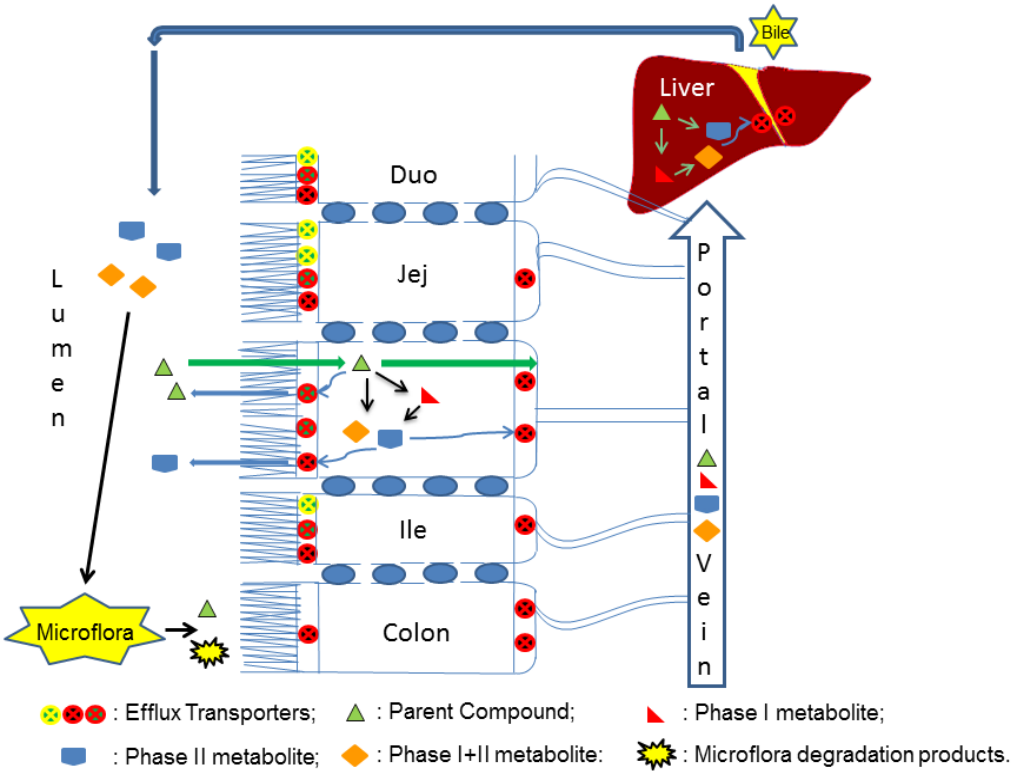


Figure 5. Scheme of disposition of xenobiotics in vivo (enterohepatic recycling, enteric recycling, microflora metabolism etc.)*

**Typical substrates could be SN-38, ginsenoside Rb1, EGCG, daidzein etc.*

The regenerated SN-38 is reabsorbed, and the reabsorbed SN-38 causes delayed diarrhea in the intestine since SN-38 damages the intestinal epithelium structurally and functionally. In Wallace's paper, selected bacterial β-glucuronidase inhibitor that blocked the conversion of SN-38G to SN-38 significantly alleviated irinotecan toxicity both in vitro and in vivo. Therefore, microbiota targeted therapy may serve as a new approach to alleviate the toxicity of drugs. Furthermore, suppression of intestinal microflora by an antibiotic cocktail pretreatment in atherosclerosis-prone mice inhibited dietary choline induced atherosclerosis as indicated by the significant reduction of aortic lesion area (Wang, Klipfell et al. 2011). In a recent study of ginsenoside bioactivation, we showed that ginsenosides with three or four sugars are poorly permeable and needs the action of bacterial glycosidases to form ginsenosides with 1 or 2 sugars, which are not only more active but also more permeable, aiding their absorption (Niu, Smith et al. 2013). Other studies have shown the impact of manipulating intestinal microflora on the metabolism of soy isoflavones to form more active compounds such as equol (Davis and Milner 2009).

In a cohort study of 32 healthy volunteers and 390 patients with or other gastrointestinal diseases, Swidsinski et. al. reported significant differences in microbial structure in patients with Crohn's disease, ulcerative colitis, and healthy and disease controls. Specifically, a depletion of Faecalibacterium prausnitzii in Crohn's disease patients and high Faecalibacterium prausnitzii in patients with ulcerative colitis was shown (0.9875 sensitivity and 0.98 specificity) (Swidsinski, Loening-Baucke et al. 2008). Therefore, diagnosing IBD can be performed by analyzing microbiota structure of the subject.

1.3.9. Chemoprevention and Probiotics/Prebiotics

Cancer remains the leading cause of cancer deaths for both men and women in the United States (Siegel, Ma et al. 2014). The administration of medication or alteration of diet to prevent or delay the development of cancer is termed as “chemoprevention” and has recently drawn public attention due to the low cure rates of many types of cancer using traditional approaches, such as chemotherapy and surgery (Hecht 2000; Clark and You 2006; Yan, Wang et al. 2006; Tan and Spivack 2009).

Multiple chemoprevention agents, such as ginsenosides (Yun 1999; Yan, Wang et al. 2006), polyphenols (Baumeister, Reiter et al. 2012; Cimino, Sortino et al. 2012; Stagos, Amoutzias et al. 2012), and isothiocyanates (Hecht 2000; Hecht, Kassie et al. 2009) have been studied extensively worldwide. Interestingly, gut microbiota has been shown to play a key role in the action of these chemoprevention agents (Chow and Hakim 2011; Gao, Basu et al. 2012; Niu, Smith et al. 2013).

Ginsenosides are a class of compounds with different sugar moieties attached to the backbone, and the sugars can be cleaved by the intestinal microflora to produce secondary glycosides or aglycones with different biological activities. The metabolic pathways of microbial metabolism of ginsenosides are distinctive among different strains (Figure 7). We have previously reported that primary ginsenoside Rb1 was hydrolyzed to ginsenoside Rd, F2, and Compound K by A/J mouse fecal lysate in a stepwise manner. In vitro cell line studies showed that instead of primary ginsenoside Rb1, compound-K possesses the anti-cancer activities (Yang, Gao et al. 2011; Niu, Smith et al. 2013) (Yang, Wang et al. 2012). However, It has also been reported that Rb1 can be

metabolized to ginsenoside Rh2, Rg3, Compound K and gypenoside LXXV by different microorganisms (Figure 6) (Bae, Han et al. 2002; Chi and Ji 2005; Hou, Xue et al. 2012; Kim, Kim et al. 2012). Due to the diverse nature of the gut microbiome, prediction of hydrolysis pattern of saponins in a particular individual may be difficult. As a result, the parent as well as the metabolite level in the systemic circulation could be highly variable among individual patients, which could represent a major challenge to the clinical testing of certain ginsenosides in vivo when given orally. Therefore, the chemoprevention efficacy of saponins may rely on the proper microbiota structure of the subject.

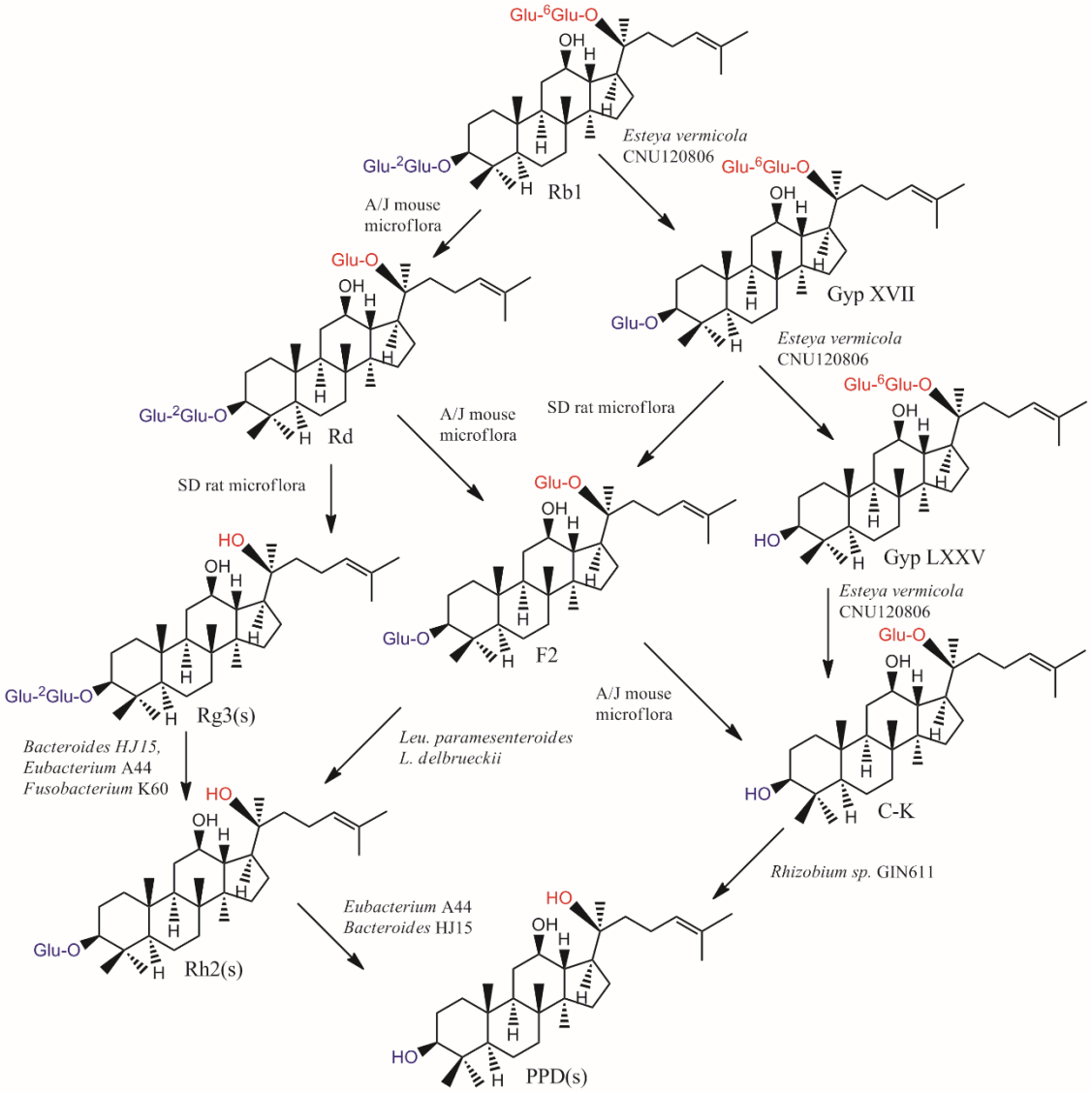


Figure 6. Distinctive metabolic pathways of ginsenoside Rb1 in different microorganisms.

Microorganisms shown in the scheme are examples from reference Chi and Ji 2005; Chen, Yang et al. 2008; Hou, Xue et al. 2012; Kim, Kim et al. 2012; Yang, Wang et al. 2012, but not exclusive examples.

Tea is an aromatic beverage widely consumed wordwide. Flavanols, commonly referred to as tea catechins, including epicatechin, epicatechin-3-gallate, epigallocatechin and epigallocatechin-3-gallate (EGCG), represents the major polyphenolic components present in green tea (91). Tea polyphenols can be directly metabolized by phase II DMEs (UGTs, SULTs, and catechol-O-methylthransferases) inside the body. Additionally, intestinal microflora produces ring fission metabolites (Yang, Sang et al. 2008). All of these metabolites are biologically active against cancer cells in *vitro*, but none of them are as active as EGCG (Yang, Wang et al. 2009). Oral administration of tea polyphenols inhibited NNK-induced lung carcinogenesis in rats and mice (Mimoto, Kiura et al. 2000; Liao, Yang et al. 2004). However, high doses of tea polyphenols are required to produce a chemoprevention effect, possibly due to the limited bioavailability of EGCG. Various epidemiological studies investigated the association between tea consumption and human cancer risks in multiple sites, but the results were not always convincing (Ju, Lu et al. 2007). In an intervention study of tea components, daily consumption of green tea for one year prevented the incidence of prostate intraepithelial neoplasia as only one among the 30 healthy subject developed tumor whereas 9 among the 30 subjects in the placebo group were diagnosed of cancer (Bettuzzi, Brausi et al. 2006).

Isoflavones are the major components present in soy beans. Daidzein, one of the well studied major isoflavones, is metabolized to equol and O-desmethylangolensin (O-DMA) by gut microflora in humans (Atkinson, Frankenfeld et al. 2005; Decroos, Vanhemmens et al. 2005). Equol and DMA have been shown to possess higher binding affinity for

human estrogen receptors α and β than the parent compound daidzein does (Muthyala, Ju et al. 2004). Therefore, *in vitro* studies suggest that equol is more biologically active than daidzein since this process is an important step in estrogen-induced biological effects. Since the production of equol is dependent of the gut microflora (NATTS Bacteria), subjects in clinical studies are classified as equol producers and non-producers (Akaza 2012). In a large-scale clinical pilot study of isoflavone preparation on subjects at high risk of prostate cancer for one year, the prostate cancer positive ratio among equol non-producers was higher than the equol producers. By combining both the equol producer and non-producer groups, the isoflavone-treatment group showed a significant decrease in the diagnosis rate for prostate cancer (Miyanaga, Akaza et al. 2012).

Clinical trials of isothiocyanates are ongoing to answer whether isothiocyanates could inhibit NNK induced lung carcinogenesis as they do in animal models (Hecht, Kassie et al. 2009). The major mechanism appears to be selective inhibition of phase I DMEs (majorly cytochrome P450) that are responsible for conversion of procarcinogens to highly reactive electrophilic species (Hecht 2000). Interestingly, isothiocyanates have also been shown to have broad anti-bacterial activities on pathogenic bacteria, including Escherichia coli, Pseudomonas aeruginosa, Listeria monocytogenes and Staphylococcus aureus (Luciano and Holley 2009; Saavedra, Borges et al. 2010). Given these results, isothiocyanates may also restructure the human gut microbial community which in turn could impact the gut health.

Multiple pathways may be involved in the chemoprevention of colon cancer, such as production of specific bacterial enzymes, reduction of reactive oxygen species (ROS) and cell apoptosis. Bifidobacterium adolescentis SPM0212 was shown to inhibit bacterial β-glucuronidase and β-glucosidase that are positively correlated with colon cancer risk (Kim, Lee et al. 2008). Impact of probiotic intervention in animal models of colon cancer also suggests that beneficial effects of probiotics depend on the species and strains. In addition, prebiotics such as polyphenols can serve as chemoprevention agents by reducing oxidative damage imposed by ROS generated during bacterial metabolism and/or catabolism.

Clinical trials of chemoprevention of colon cancer by probiotics are limited despite the large amount studies in animals (Azcarate-Peril, Sikes et al. 2011). In one study, a decrease in colorectal proliferation and genotoxins was reported in a clinical trial that uses Lactobacillus rhamnosus GG and Bifidobacterium lactis Bb12 on 37 colon cancer patients and 43 polypectomized patients (Rafter, Bennett et al. 2007). Clinical trials of prebiotics in the chemoprevention of colon cancer have also been reported. For exmaple, a randomized, phase II chemoprevention trial involving subjects 40 years or older, with previously resected colon cancer or multiple/advanced colorectal adenomas was reported (Limburg, Mahoney et al. 2011). Limburg PJ et. al. conducted a randomized, phase II chemoprevention trial involving current or former cigarette smokers (>30 pack-years) (Limburg, Mandrekar et al. 2013). However, none of these studies have produced conclusive results.

The precise mechanisms involved in the chemoprevention of cancer via probiotic/prebiotic intervention are not fully understood. Probiotics may pose transient (i.e. species of Lactobacillus and Bifidobacterium) or long-term impact (i.e. species of Bacteroides and Clostridium) on the host. Selection of probiotics is crucial since transient inoculation may have fewer undesired consequences while long-term colonization may shape the microbiome and alter host functions more effectively. Another key question of probiotics is the number of bacteria species and cells (Lemon, Armitage et al. 2012). In general, no general consensus has been reached in designing effective probiotics for consumption by humans. However, the enthusiasm surround the research of microbiome suggest that this strategy should be found in the next 5 to 10 years.

1.4. Summary

This suvey of the literature reveals that RGE is a potent agent for chemoprevention of lung cancer. The primary ginsenosides in RGE are inactive, while the secondary glycosides or aglycones possess different anticarcinogenic activities. The gut microbiota is believed to correlate with human health and most notably, the bioactivation of chemoprevention agents including RGE. However, there is no direct evidence yet that certain bacteria population is critical for the bioactivation of RGE in A/J mouse. Additionally, the impact of consumption of chemoprevetion agents to bacteria population in the gut is lacking. Therefore, unraveling the contribution of gut microbiota to the chemoprevention efficacy of RGE and the interaction between RGE and microbiota is of considerable interest in this context.

Chapter 2 Hypotheses and specific aims

2.1. Central hypothesis

RGE, a mixture of ginsenosides, has been shown as a potent agent for chemoprevention of lung cancer. The sugar moieties of ginsenosides can be cleaved by the intestinal microflora to produce secondary glycosides or aglycones with different anticarcinogenic activities. Investigation of the microbial metabolism of ginesnosides is critical to fully understand RGE's chemoprevention potential. We hypothesize that lung cancer chemoprevention of ginsenosides is mediated by the glycosidase activities of the mouse intestinal microbiota defined by pyrosequencing.

2.2. Specific aims

2.2.1. Aim I (Chapter 3)

To characterize microbiota dependent metabolism of RGE in A/J mouse fecal lysate. It is hypothesized that the biotransformation of ginsenosides possesses distinctive kinetics. We will examine the metabolic pathways of ginsenoside Rb1 in the gut of A/J mouse and determine the rate limiting step in the production of ginsenoside Compound K.

2.2.2. Aim II (Chapter 4)

To determine the impact of RGE administration on bacterial glycosidase activity and diversity of the intestinal microbiome in A/J mouse. In this study, we will focus on the changes of bacterial glycosidase activity and intestinal microbiome in A/J mouse feces after oral administration of RGE for 7 days. It is hypothesized that glycosidase activity increase and certain bacteria repopulate favorably with oral administration of RGE.

2.2.3. Aim III (Chapter 5)

To enrich and identify bacterial glycosidase(s) from A/J mouse feces that catalyze the rate limiting step in the production of ginesnoside Compound K. There are 140 Protein Data Bank entries in enzyme class E.C.3.2.1.21 (Beta-glucosidase), the major class of enzymes in the biotranformation of ginsenosides. It is hypothesized that only certain numbers of glycosidases are responsible for RGE metabolism, particularly the formation of ginsenoside F2. We will purifiy and identify bacterial glycosidases from A/J mouse feces using a classic chromagraphic approach.

2.2.4. General strategy

Due to the in vitro and in vivo discrepancy of bioactivities of RGE, we proposed an intestinal microbiota based approach to elucidate the mechanism of bioactivation of RGE in A/J mouse. In **Aim I**, the metabolic pathways of ginsenoside Rb1, one of the major ginsenoside present in RGE, was examined experimentally. There are multiple bacteria and bacterial glycosidases involved in the metabolism of ginsenoside Rb1 as shown in Figure 7. The findings that Rb1 was hydrolyzed by bacterial glycosidases in a stepwise manner and the formation of ginsenoside F2 was the rate limiting step in the production the ginesnoside Compound K narrowed down the corresponding bacteria and bacterial glycosidases involved. We tried to further narrow the scope of bacteria by studying the bacteria population in response to RGE treatment by 16s rRNA pyrosequencing in **Aim II**. The classic protein chemistry approach in **Aim III** helped us identify bacterial glycosidases and bacteria species that potentially have impacts on chemoprevention of RGE *in vivo*.

Chapter 3 Characterization of microbiota dependent metabolism of RGE in A/J mouse fecal lysate

3.1. Abstract

RGE shows promise in preventing lung cance. However, the ginsenosides responsible for the activity of RGE are poorly defined as primary ginsenosides are inactive. The objective of this study is to investigate the role of intestinal microbiota in activating primary ginsenosides in RGE using A/J mouse fecal lysate. The kinetics of microbiota mediated biotransformation of ginesnosides was characterized and kinetic parameters (metabolite formation rates) were determined. Additionally, the anti-proliferative activity of ginsenosides was tested using the mouse lung cancer LM1 cells. Permeabilities of ginesnosides were also evaluated in Caco-2 cell monolayers. Systemic exposure of secondary ginsenosides was determined in A/J mice. Secondary ginsenoside Compound K exhibited higher anti-proliferative activity and permeability than primary ginsenosides, and significant amounts of secondary ginsenosides (F2 and Compound K) were found in the blood of A/J mice following oral administration of the primary ginsenoside Rb1. Because mammalian cells did not hydrolyze ginsenoside, we determined the ability of bacteria to hydrolyze ginsenosides and found that the primary ginsenoside Rb1 underwent stepwise hydrolysis to Rd, F2, and then Compound K. Formation of F2 from Rd was the rate-limiting step in the biotransformation of Rb1 to Compound K. In conclusion, this is the first study to characterize the A/J mouse intestinal microbiome and reveal the presence of certain bacterial families capable of

efficiently converting inactive primary ginsenosides to active secondary ginsenosides in vivo.

3.2. Introduction

Lung cancer remains the leading cause of cancer deaths for both men and women in the United States (Siegel, Ma et al. 2014). The administration of medication or alteration of diet to prevent or delay the development of cancer is coined “chemoprevention” and has recently drawn public attention due to the low cure rates of advanced lung cancer using traditional approaches, such as chemotherapy and surgery (Hecht 2000; Clark and You 2006; Yan, Wang et al. 2006; Khan, Afaq et al. 2008; Kelly, Kittelson et al. 2009; Ohashi, Takigawa et al. 2009; Scott, Gescher et al. 2009; Tan and Spivack 2009; Keith, Karoor et al. 2010; Qi, Wang et al. 2010; Johnson, Hermanson et al. 2011; Sun, Qi et al. 2011).

The efficacy of red ginseng extract (RGE) as a chemoprevention agent has been examined over the past 30 years (Yun TK 1983; Yun TK 1987; SH 1988; Yun 1991; Yun TK 1995; Yan, Wang et al. 2006), and accumulated evidence supports the notion that RGE is a potent agent for the prevention of lung cancer. A recent study by Yan, et al. indicated that RGE (10 mg/ml in drinking water) prevented benzopyrene induced lung carcinogenesis with a significant reduction (70%) of tumor load (Yan, Wang et al. 2006).

The primary, or most abundant, naturally occurring ginsenosides present in RGE include Rb1 and Rd (Li, Lee et al. 2008; Kong, Wang et al. 2009). However, secondary, less abundant, ginsenosides such as Rh2 demonstrate excellent inhibitory activities in both lung adenocarcinoma cells (Cheng, Yang et al. 2005) and in A/J mice (Yun, Lee et al.

2001). Additionally, secondary ginsenosides appear to have better permeability in Caco-2 cells and hence better bioavailability (see later). Mammalian cells do not express enzymes that hydrolyze ginsenosides (Henrissat 1991; Henrissat and Bairoch 1993; Henrissat and Bairoch 1996; Park, Yoo et al. 2010) suggesting that the in vivo transformation of secondary ginsenosides occurs via the action of the intestinal microbiota.

The intestinal microbiome produces different types of glycosidases, including β-glucosidases, the predominant enzymes responsible for the hydrolysis of ginsenosides. Bacterial β-glucosidases are hydrolytic enzymes that release terminal glucose residues successively (Bhatia, Mishra et al. 2002). Primary ginsenosides are hydrolyzed stepwise to produce secondary ginsenosides and finally the aglycone. Secondary ginsenosides exhibit the much higher anti-proliferative activity against lung cancer cells and possess higher permeability across membranes of mammalian cells. Therefore, enhancing the production of secondary ginsenosides will likely benefit the lung cancer chemoprevention efficacy of RGE.

Recent investigations showed that the primary ginsenoside Rb1 is hydrolyzed in a stepwise manner to different secondary ginsenosides, such as Rg3, Rh2, F2, and Compound K, by fecal extracts and specific microorganisms (Hasegawa, Sung et al. 1997; Akao T 1998; Bae, Choo et al. 2002; Chi, Kim et al. 2005; Cheng, Na et al. 2007; Chen, Yang et al. 2008; Kim, Kim et al. 2008; Son, Kim et al. 2008; Zhou, Zhou et al. 2008). However, these studies failed to investigate the kinetics of RGE stepwise hydrolysis necessary to elucidate the rate-limiting step in the production of active

ginsenosides. Identification of the rate-limiting step and the microorganism that catalyzes it could allow us to manipulate the rate and production of desired ginsenosides with anti-cancer activity. Furthermore, the microorganisms studied do not represent the whole spectrum of the intestinal microbiome. Studies using fecal specimens failed to define the gut microbiota capable of ginsenoside hydrolysis (Relman, Dethlefsen et al. 2008; Castagnini, Luceri et al. 2009; Bolam, Sonnenburg et al. 2010). Therefore, these studies leave open the potential to discover probiotic candidates that could be administered concurrently with RGE to improve its efficacy.

In the present study, we (1) determine the rate-limiting step in the kinetics of stepwise metabolism from ginsenoside Rb1 to Compound K in A/J mouse fecal matrix; and (2) use 16s rRNA pyrosequencing to define the A/J mouse intestinal microbiome to aid future studies investigating bacterial metabolism of ginsenosides from this and other laboratories.

3.3. Materials and methods

3.3.1. Chemicals and reagents

Ginsenosides Rb1, Rd and F2 (>99% pure) were purchased from LKT Laboratories (St. Paul, MN). Ginsenoside Compound K (>95% pure) was kindly provided by Dr. Zhi-Hong Jiang from Hong Kong Baptist University. Red Ginseng powder and LM1 cells were provided by Dr. Ming You from Medical College of Wisconsin. Testosterone was purchased from Sigma-Aldrich (St. Louis, MO). Simulated intestinal fluid was purchased

from VWR (Houston, TX). BCA protein assay kit was purchased from Thermo Scientific (Rockford, IL). Other chemicals (analytical grade or better) were used as received.

3.3.2. LM1 cell culture model and anti-proliferative assays

LM1 cells were grown in plastic tissue culture flasks in minimal essential medium (MEM) mixed with 2 mM glutamine, 50 units/ml penicillin, 50 pg/ml streptomycin, 1% nonessential amino acids from Cellgro (Catalog number: 10-010-CV), and supplemented with 10% fetal bovine serum from HyClone (Catalog number: SH30088-03). Cells were incubated at 37°C and 5% CO_2, with media changes every two days. Cells were passaged 10 times, and one split (1:20) was used each week.

LM1 cells were seeded onto 96-well plates at a density of 5×10^3 per well and incubated in MEM without red phenol from Cellgro (Catalog number: 17-305-CV). After 24 hr, ginsenosides, RGE, 5-flurouracil (positive control) or 1% DMSO (negative control) was added to the medium for 48 hr. Cells were then incubated with MTT (0.5 mg/ml) for 4 hr. The formazan precipitate was dissolved in 100 µL DMSO, and the absorbance at 570 nm was detected with a Benchmark Microplate Reader (Bio-Rad, California). Cell survival was calculated by the following formula: % cell survival = (mean absorbency in test wells) / (mean absorbency in control wells) × 100. The effective dose to inhibit 50% growth (ED_{50}) was calculated for the ginsenosides and the positive control, using non-linear regression analysis. Each test was performed in triplicate.

3.3.3. Caco-2 cell culture and transcellular transport experiments

Caco-2 cells were cultivated as described previously (Yang, Gao et al. 2010). Porous polycarbonate cell culture inserts (3 μm pore size) from Corning (Catalog No: 3414) were used to seed cells at a density of 100,000 cells/cm^2. After the cell monolayers reach maturity and become ready to use in 19-21 days, they were tested for integrity (minimal transepithelial electrical resistance value of 465 Ω/cm^2), and then used for the transcellular transport as described previously (Yang, Gao et al. 2010). Briefly, 2.5 ml of ginsenoside solution (2 or 10 μM) in Hanks's balanced salt solution or HBSS was loaded on apical side of the cell monolayer and 2.5 ml of blank HBSS onto the basolateral side. Five sequential samples (0.5 mL) were taken at time points 0, 1, 2, 3 and 4 hr from both sides of the cell monolayer. Ginsenoside solution and HBSS media was added to donor or receiver side immediately to compensate for sampling volume lost. The pH of HBSS in both the apical and basolateral side was 7.4. A volume of 125 μL of an internal standard (1 μM formononetin in 100% acetonitrile) was added to the samples right after sampling. Caco-2 samples were blown dried by purified air and reconstituted with 200 μL of 100% methanol for UPLC-MS/MS analysis.

The apparent unidirectional permeability, from apical to basolateral side ($P_{a\text{-}b}$), was obtained according to the following equation (Eq.1):

$$P_{app} = \frac{dC}{dt} \times \frac{V}{SC_0} \qquad (1)$$

where $\frac{dC}{dt}$ is the rate of concentration change in the receiver chamber (equals to the slope of the regression line derived for the amount transported vs. time profile), V is the chamber volume (2.5 mL), S is the surface area of the monolayer (4.65 cm^2), and C_0 is the initial concentration in the donor side. Permeability from apical to basolateral side (P_{a-b}) was calculated according to the above equation.

3.3.4. Oral pharmacokinetic dosing studies using Rb1 in A/J mice

The study was approved by the Institutional Animal Care and Use Committee at the University of Houston. Male A/J mice (20-25 g) were purchased from Harlan Laboratory (Indianapolis, IN) at 6-8 weeks of age. They were housed individually in an environmentally controlled room (temperature: 25±2 ℃, humidity: 50±5 %, 12-hr light-dark cycle) for one week before the experiments. Drinking water and diet were supplied *ad libitum*. The body weight of mice was measured every other day for the duration of the study.

Ginsenoside Rb1 dispersed in oral suspending vehicle was given by oral gavage to 4 A/J mice at a dose of 20 mg/kg. Blood samples (20-25 µL per sample) for each mouse were collected in heparinized tubes at 0.25, 0.5, 1, 2, 3, 4, 6, 8, 12 and 24 hr by snipping the tail and stored at -20 ℃ until analysis. The internal standard solution (200 µL of 1 µM formononetin in methanol) was added to each 20 µL aliquot of blood. The samples were vortexed for 30 sec and centrifuged at 15,000 rpm for 15 min. The supernatant was

blown dried by purified air and reconstituted in 100 µL of 100% methanol, and 10 µL of the sample was injected into the UPLC-MS/MS for analysis.

3.3.5. Hydrolysis of ginsenosides in fecal lysate

3.3.5.1 Fecal lysate preparation

Feces of nine A/J mice were collected and stored at -80°C until use. Feces (500 mg) were mixed with 5 ml ice-cold 0.1 mM phosphate buffer solution or PBS (pH 7.4) and vortexed for 5 sec followed by centrifugation at 1,000 rpm and 4°C for 15 min. The pellet was further washed twice using 0.1 mM PBS, followed by resuspension in 10 ml ice-cold PBS, sonicated in ice water bath for 45 min and centrifuged at 15,000 rpm and 4°C for 30 min. The supernatant was aliqoted and stored at -80°C. Protein concentrations were determined by the BCA protein assay kit using bovine serum albumin as the standard.

3.3.5.2 Stepwise metabolism of ginsenoside Rb1

Thawed fecal lysate (200 µL) was transferred to disposable glass vials from VWR (Houston, TX) and diluted with ice-cold 0.1 mM PBS to 2 ml. Ginsenoside Rb1 was added to the samples to a final concentration of 45 µM. The mixture was incubated at 37°C and 120 rpm in an orbital shaker from Thermo Scientific (Asheville, NC) for 24 hr. Samples (100 µL) were collected with low adhesion surface tips from VWR (Houston, TX) to minimize binding to the micro-centrifuge tubes from Corning (Pittston, PA) at 0, 0.5, 1, 2, 4, 8, 12 and 24 hr. The reaction was stopped with the addition of 500 µL of 2.5 µM testosterone (internal standard) in 100% acetonitrile. The samples were vortexed for 15 sec and centrifuged at 15,000 rpm for 15 min. A 540 µL portion of supernatant was

air dried and reconstituted with 100 µL of 30% acetonitrile, and 10 µL of the sample was injected into the UPLC for analysis. The experiments were performed in triplicate.

3.3.5.3 Determination of the rate-limiting step in ginsenoside Rb1 metabolism.

Fecal lysate protein concentration, incubation conditions, and sampling time were optimized to ensure less than 30% of the metabolite would appear such that the calculated metabolic rate closely approximates the initial rate. Ginsenosides Rb1, Rd and F2 and Compound K were added to the diluted fecal lysate described in previous section, making the final substrate concentration of 20 or 5 µM. The sample processing procedures were performed as described above, in triplicates. Amounts of metabolite formed were determined by UPLC and protein concentrations were determined using BCA protein assay. The appearance rates of Rb1, Rd, F2, and Compound K were normalized by protein concentration and reaction time. Recovery of total ginsenosides was determined to ascertain that the mass balance fell in the range of 80% to 120% (i.e., no other metabolites were present in significant quantities).

3.3.6. Quantitation of ginsenosides in biological matrices

For Caco-2 and mouse blood samples, a UPLC-MS/MS method was used to quantify ginsenosides. The LC conditions for the analysis were the same as those for the quantitation of ginsenoside Rh2 (Yang Z 2011). A triple quadruple mass spectrometer (API 3200 Qtrap, Applied Biosystems, Foster City, CA) was used to perform the analysis of the eluent from the UPLC. The ion spray voltage and ion source temperature were set to 5500 kV and 600°C, respectively. Nebulizer gas, turbo gas, and curtain gas were

optimized to 40, 40, and 20 psi, respectively. Multiple reactions monitoring (MRM) mode was used to monitor ginsenoside Rb1, Rd, F2, Compound K and formononetin (internal standard). The compound dependent parameters were listed in Table 1.

Table 1. Compound dependent parameters of ginsenoside Rb1, Rd, F2, C-K and internal standard testosterone in MRM mode LC/MS analysis

Compound	Q1	Q3	Time (msec)	DP	CEP	CE	CXP
Rb1	1110.0	325.3	100	36	62	38	5
Rd	946.7	161.0	100	-54	-42	-64	-1
F2	807.6	627.5	100	125	50	53	5
C-K	645.4	202.9	100	96	34	45	3
IS	269.2	197.1	100	10	14	49	3

DP, decluster potential; CEP, collision cell entrance potential; CE, collision enery; CXP, collision cell exit potential; IS (internall standard), testosterone

Processed samples were injected into an Acquity UPLC BEH C18 column (50 mm×2.1 mm, 1.7 µm, Waters, Milford, MA) and run for 6.5 min with a flow rate of 0.5 ml/min. The elution gradient was as follows: initial, 80% A (100% water) and 20% B (100% acetonitrile), 0–0.5 min, 20% B, 0.5–1 min, 20-35% B, 1–2 min, 35% B, 2–3 min, 35-50% B, 3–4 min, 50% B, 4–4.5 min, 50-75%, 4.5-5 min, 75-90% B, 5–5.2 min, 90-100% B; 5.2-6 min 100% B, 6-6.5 min 100-20% B. UPLC analysis was performed on a Acquity UPLC (Waters, Milford, MA) with photodiode-arrayed (PDA) detector. Quantitation was performed at 200 nm wavelength. The column temperature and sample temperature were set to 45°C and 20°C, respectively. The injection volume was 10 µL. The chromatograph and UV spectrum of Rb1, Rd, F2, Compound K, and testosterone are shown in Appendix A.

The standard curves for Rb1 and Rd in fecal lysate were linear in the concentration range of 0.39-50 µg/ml with correlation coefficient values >0.999. The lower limit of quantification (LLOQ) was 0.39 µg/mLfor both Rb1 and Rd. The standard curves for F2 were linear in low concentration range of 0.306-4.9 µg/mL and in high concentration range of 4.9 - 39.2 µg/ml. As for Compound K, standard curves were linear in low concentration range of 0.243 – 3.89 µg/ml and in high concentration range of 3.89 - 31.1 µg/ml. The correlation coefficient values were 0.9994, 0.9937 for F2 and 0.9999, 0.9916 for Compound K, respectively. The LLOQ was 0.306 µg/mL and 0.243 µg/mL for F2 and Compound K, respectively. Intra-day and Inter-day precision and accuracy were well within the 15% acceptance range for all quality control (QC) samples at three concentrations levels in fecal lysate (Appendix B). The mean extraction recoveries

determined using three replicates of QC samples at three concentration levels in fecal lysate fell in the range of 51.4% to 98.1% (Appendix B). The stability of 20 µM Rb1, Rd, F2, and Compound K in PBS (37°C, 4 hr), A/J mouse liver, small intestine, and colon S9 (37°C, 24 hr), the long term stability of glycosidases in fecal lysate (-80°C, 1 month), and three cycles of freeze-thaw effect were evaluated in triplicates. All the samples displayed 90-120% recoveries in the stability tests (Appendix C).

3.3.7. Pyrosequencing of gut microbiome using 16s rRNA

Fecal samples (500 µL) were homogenized in 1.5 ml of Fecal Bead Solution (MoBio, Carlsbad, CA), centrifuged for 5 min at 2000x g, and 500 µL supernatant was transferred to PowerBead Tubes (MoBio, Carlsbad, CA). Samples were heated for 10 min at 65°C and 95°C to aid bacterial lysis. Genomic DNA was isolated using PowerSoil DNA Isolation Kit purchased from MoBio Laboratories (Carlsbad, CA) starting at step 2 of the manufacturer's protocol. DNA concentration and purity was determined by spectrometry on the NanoDrop ND-2000 (Thermo Scientific). Variable regions 3 to 5 (V3-V5) of the 16S rRNA gene were amplified using barcoded primers 357F (CTGCTGCCTCCCGTAGG) and 926R (CCGTCAATTCMTTTRAGT). Polymerase chain reaction (PCR) (total volume 20 µL) contained AccuPrime PCR Buffer II (Invitrogen, Carlsbad, CA), AccuPrime Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA), 4 nM barcoded primers, and 2 µL of template DNA. PCR was performed on Mastercycler ep gradient S (Eppendorf, Hamburg, Germany) with the following cycling conditions: 2 min at 95°C and 30 cycles of 20 sec at 95°C, 30 sec at 50°C, and 5 min at 72°C. DNA concentration of PCR products was determine by PicoGreen (Invitrogen,

Carlsbad, CA) and on an Agilent Bioanalyzer 2100 DNA 1000 chip (Agilent Technologies, Santa Clara, CA). Pyrosequencing was performed on Roche 454 FLX Titanium platform (Branford, CT) according to the manufacturer's instructions.

Sequences were pre-processed using Mothur (Schloss, Westcott et al. 2009). Reads were removed from subsequent analysis if they had a quality score of <35 over a 50 bp window, contained ambiguous bases or homopolymer repeats of >8 bp, had >1 bp mismatch from the barcode sequence, or had >2 bp mismatch from the primer sequence. Sequences were binned based on barcode followed by trimming of the barcode and primer sequences. Sequences were analyzed using the CloVR 16S pipeline (Angiuoli, Matalka et al. 2011).

3.3.8. Statistical analysis

The data in this study are presented as mean $\pm$ SD, if not specified otherwise. For enzyme function studies, significance is assessed by one way ANOVA and Student's-test. A p-value of < 0.05 was considered statistically significant. Relative abundance and mean relative abundance were determined for each taxa found in the microbial communities.

3.4. Results

3.4.1. Anti-proliferation activities of ginsenosides

We determined the anti-proliferative activities of related primary and secondary ginsenosides in LM1 cells, a metastatic lung cancer cell line derived from the A/J mouse (McDoniels-Silvers, Herzog et al. 2001). While primary ginsenosides are well

represented in RGE, secondary ginsenosides are derived from the primary ginsenosides by the metabolic action of bacterial glycosidases (Figure 11). The MTT assay showed that secondary ginsenosides, but not primary ginsenosides, significantly inhibit proliferation of LM1 cells (Figure 7). No anti-proliferation activity was observed for RGE, where the major species are primary ginsenosides. The IC_{50} value of ginsenoside F2 was more than 100 μg/mL in the LM1 cell line, while the activity of Compound K was the highest among the tested ginsenosides, with the IC_{50} value of approximately 13 μg/mL. These results show that the anti-proliferation activities of ginsenosides may be closely correlated to the number of sugars attached.

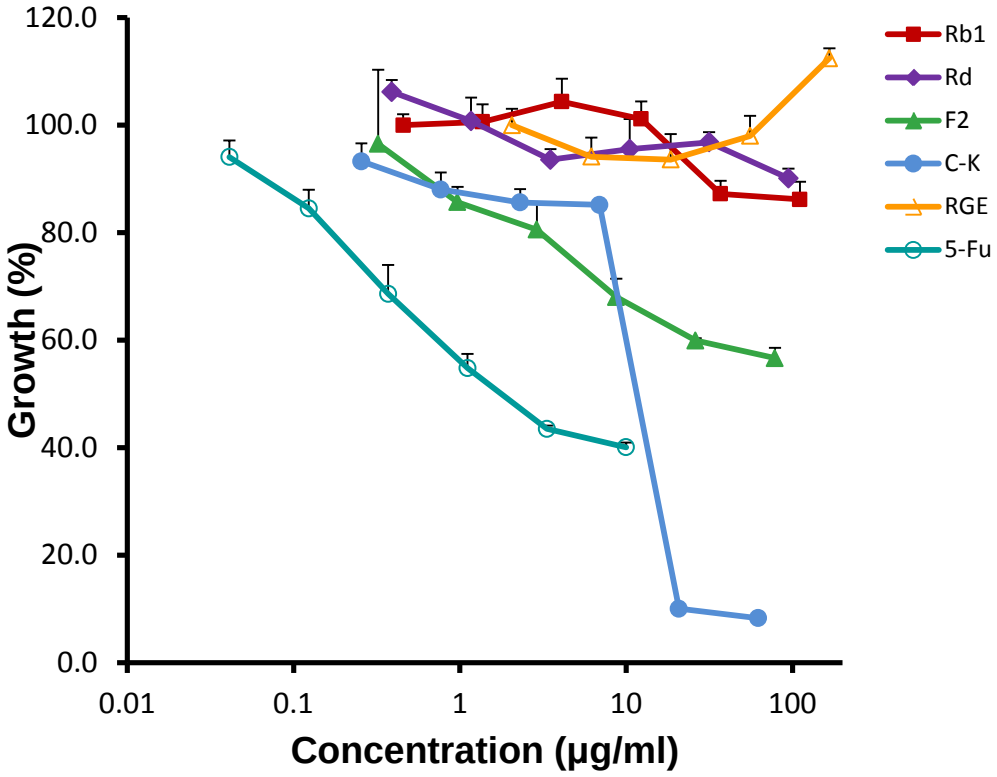


Figure 7. Anti-proliferation activity of ginsenosides Rb1, Rd, F2, C-K, and RGE in the lung cancer LM1 cell line.

5-fluorouracil was used as the positive control and 1% DMSO was used as the negative control. The concentration range of tested agents was 0.46-110.80 µg/mL for Rb1, 0.39-94.6 µg/mL for Rd, 0.32-78.4 µg/mL for F2, 0.26-62.2 µg/mL for C-K, 2.06-500 µg/mL for RGE and 0.041-10 µg/mL for 5-fluorouracil, respectively.

3.4.2. Transcellular transport of ginsenosides across Caco-2 cell monolayers

We determined the permeability of ginsenosides Rb1, Rd, F2, and Compound K in Caco-2 cell monolayers, a model employed to mimic human intestinal absorption characteristics (Artursson and Karlsson 1991; Hu, Chen et al. 1994). Transport of 10 μM Rb1 and Rd, 2 μM F2 and Compound K from apical side to basolateral side (A-B) in Caco-2 cells were studied at pH 7.4. The results (Figure 8) show that the permeability of ginsenosides Rb1, Rd and F2 were less than 1×10^{-6} cm/sec (corresponding to incomplete absorption in humans (Artursson and Karlsson 1991), indicating that ginsenosides with more than one glucose were poorly permeable. Whereas ginsenoside Compound K, with only one glucose attached, exhibited moderate permeability (P_{app} between 1×10^{-6} cm/s and 10×10^{-6} cm/s, corresponding to good absorption in humans (Artursson and Karlsson 1991). This pattern of distinctive permeabilities between primary and second ginsenosides was similar to the correlation of ginsenoside anti-proliferation activity and the number of sugar moieties.

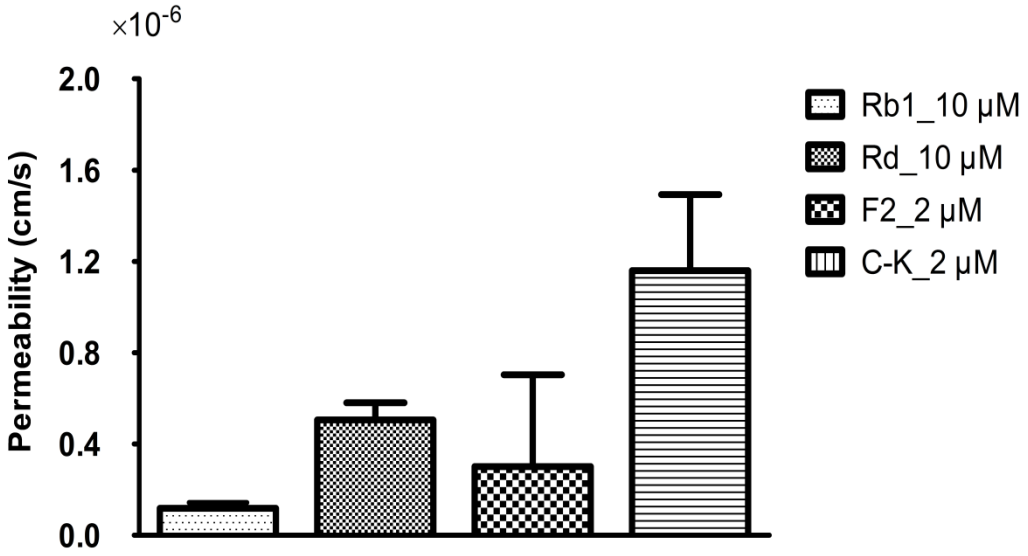


Figure 8. Transcellular transport of ginsenosides Rb1, Rd, F2, and C-K from apical side to basolateral side in Caco-2 cell monolayers.

Data are presented as mean ± S.D.; n=3.

3.4.3. Oral pharmacokinetics of Rb1 in A/J mice

The plasma concentrations of ginsenoside Rb1 and its metabolites in A/J mice were determined following oral administration of Rb1 at 20 mg/kg. The results (Figure 9) show that significant amounts of secondary ginsenosides F2 and Compound K were observed in blood, reaching peak concentrations higher than 1 μM. A/J mouse liver, small intestine, and colon S9 fraction (which contains epithelium-derived enzymes) did not hydrolyze Rb1 (Appendix C). This is consistent with the fact that β-glucosidases are responsible for hydrolysis of ginsenosides Rb1 and none of the β-glucosidases identified so far are from mammalian cells (Henrissat 1991; Henrissat and Bairoch 1993; Henrissat and Bairoch 1996; Park, Yoo et al. 2010). Therefore, the presence of secondary ginsenosides F2 and Compound K were attributed to the action of bacterial glycosidases in the intestinal microbiome.

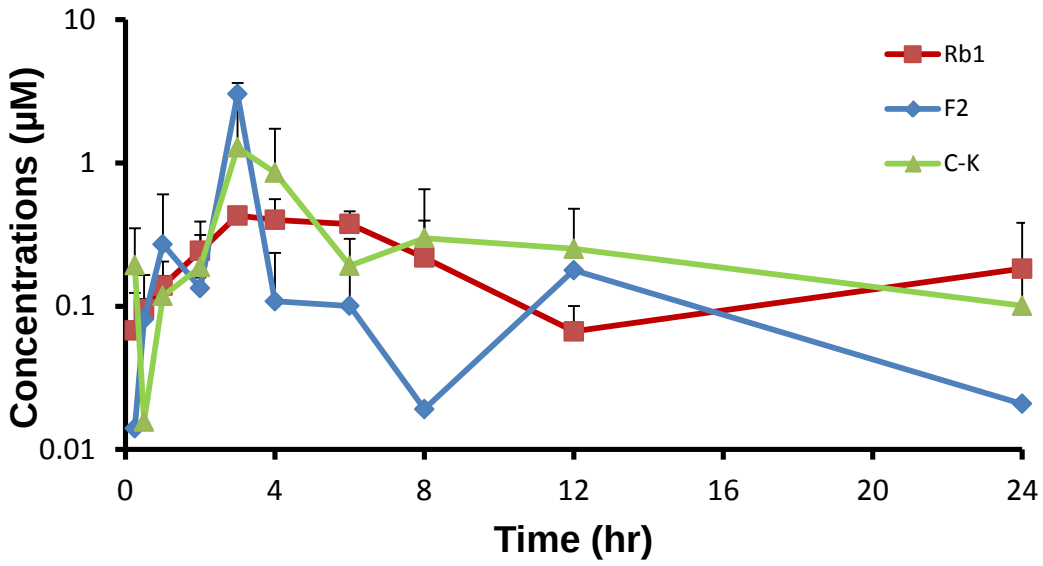


Figure 9. Pharmacokinetic profiles of Rb1 and its metabolites F2 and C-K following Rb1 oral administration (20mg/kg) in male A/J mice.

Data are presented as mean ± S.D.; n=4.

3.4.4. Stepwise metabolism of ginsenoside Rb1

An *in vitro* hydrolysis study of ginsenoside Rb1 by A/J mouse fecal lysate was performed to correlate with the pharmacokinetic profiles seen in Figure 10. Primary ginsenoside Rb1 was hydrolyzed to ginsenoside Rd, as indicated by the rapid disappearance of Rb1 and appearance of Rd within 1 hr, while F2 and Compound K were rarely found at this time point (Figure 10). Concentration of Rd plateaued around 2 hr and then began to drop. However, Rd's metabolite F2 accumulated only slightly during the course of incubation, suggesting that the formation of F2 from Rd was the slowest step. Compound K's concentration paralleled F2 concentration initially but increased significantly after 8 hr, compensating for the dramatic loss of Rd. Also, this pattern of hydrolysis indicated that the formation of Compound K from F2 was rapid. Taken together, primary ginsenoside Rb1 was hydrolyzed stepwise to Rd, F2 and finally Compound K as shown in Figure 11.

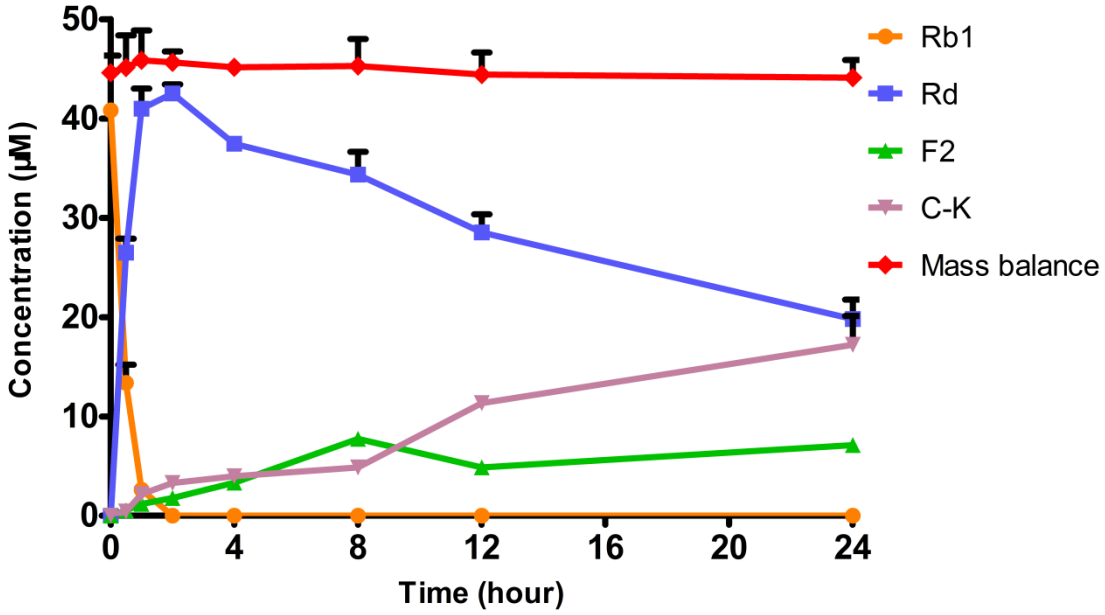


Figure 10. Metabolic profile of ginsenoside Rb1 (45 μM) in A/J mouse fecal lysate. Mass balance stands for the total amount of ginsenosides recovered.

Data are presented as mean ± S.D.; n=3.

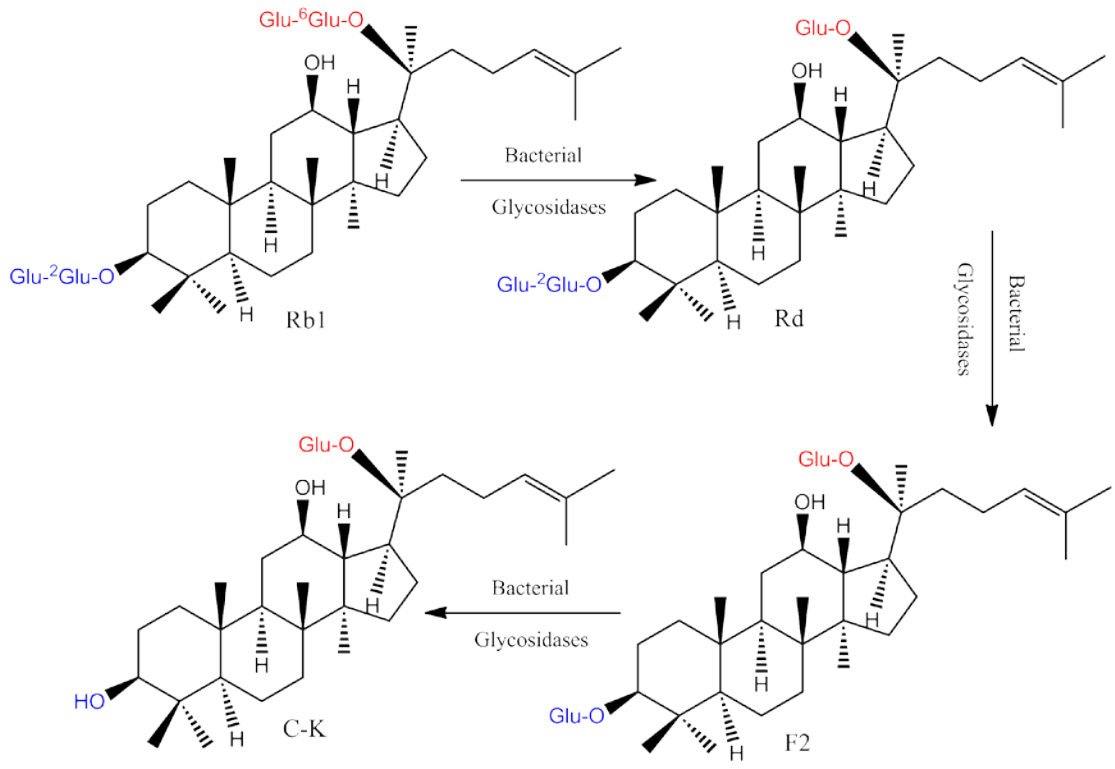


Figure 11. The proposed metabolic pathway for production of active ginsenoside C-K in A/J mouse fecal lysate.

Data are presented as mean ± S.D.; n=3. The superscript denotes the position of the hydroxyl group attached to the adjacent glucose.

3.4.5. Rate-limiting step in Rb1 stepwise hydrolysis

Metabolite formation rates of ginsenosides normalized for protein concentration are presented in Figure 12. The hydrolysis rates of Rb1 to Rd, Rd to F2, and F_2 to Compound K were 1.13 ± 0.04, 0.09 ± 0.003, 1.36 ± 0.27 nmol/min/mg at 20 μM concentration, and 0.60 ± 0.01, 0.09 ± 0.01, 0.52 ± 0.01 nmol/min/mg at 5 μM concentration, respectively. The hydrolysis rate of Rd to F2 was approximately 11 fold and 14 fold lower than the formation of Rd and Compound K (from their corresponding substrates) at 20 μM, respectively. At 5 μM concentration, a similar hydrolysis pattern was observed. The formation rate of F2 was around 5 fold less than the formation of Rd or Compound K. Formation of aglycone protopanaxadiol (PPD) from Compound K was not observed during incubation, which could be attributed to the low solubility of the aglycone in aqueous solution (3 μg/mL) (Han, Ma et al. 2015). However, the amount of aglycone formed must be small, since the recovery was within 80-110% at two concentrations.

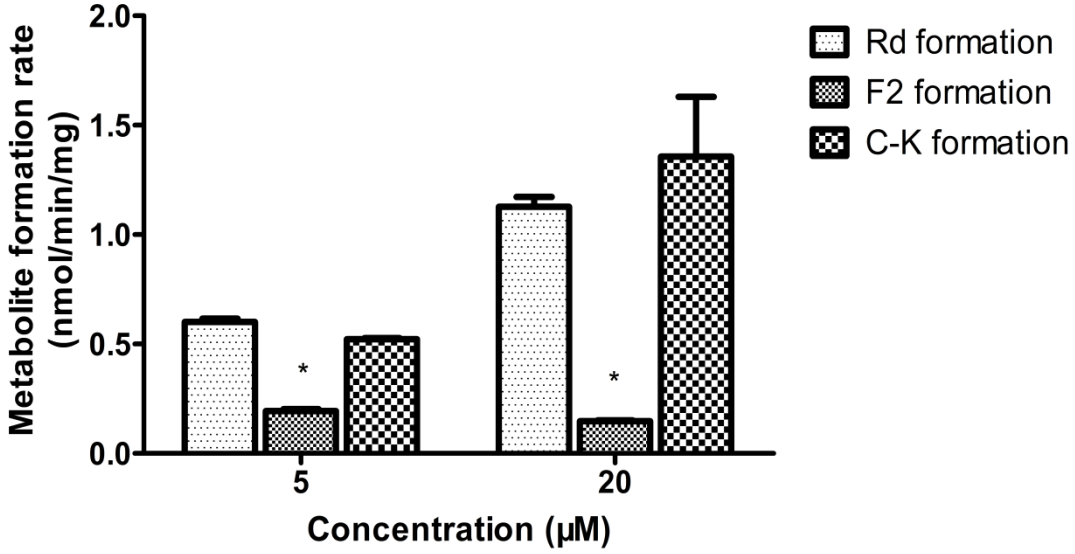


Figure 12. Stepwise metabolite formation rates of ginsenosides Rb1, Rd, and F2 in A/J mouse fecal matrix.

Data are presented as mean ± S.D.; n=3. 20 or 5 μM ginsenosides as substrates, fecal lysates were incubated for 30 min. Metabolite formation rates were calculated as the metabolite concentration divided by protein concentration of the fecal lysate and reaction time. The "" symbol indicated p<0.05, analyzed by one-way ANOVA and Student's t-test.*

3.4.6. Membership and relative abundance of the A/J mouse intestinal microbiota

Presence of bacteria capable of ginsenoside hydrolysis, we performed 16S rRNA pyrosequencing to determine membership and relative abundance of the A/J mouse intestinal microbiota and identify bacterial species capable of hydrolyzing ginsenosides. Relative abundance at the family level of 9 A/J mice demonstrates the low inter-animal variability of the intestinal microbiome (Figure 13A). The mean relative abundance in Figure 13B demonstrates that 5 major families (“Lachnopiraceae”, “Ruminococcaceae”, Bacteroidaceae, Porphyromonadaceae, and Prevotellaceae) comprise approximately 80 percent of the A/J mouse intestinal microbiota. An additional 15 minor families account for 10% of the intestinal microbiome.

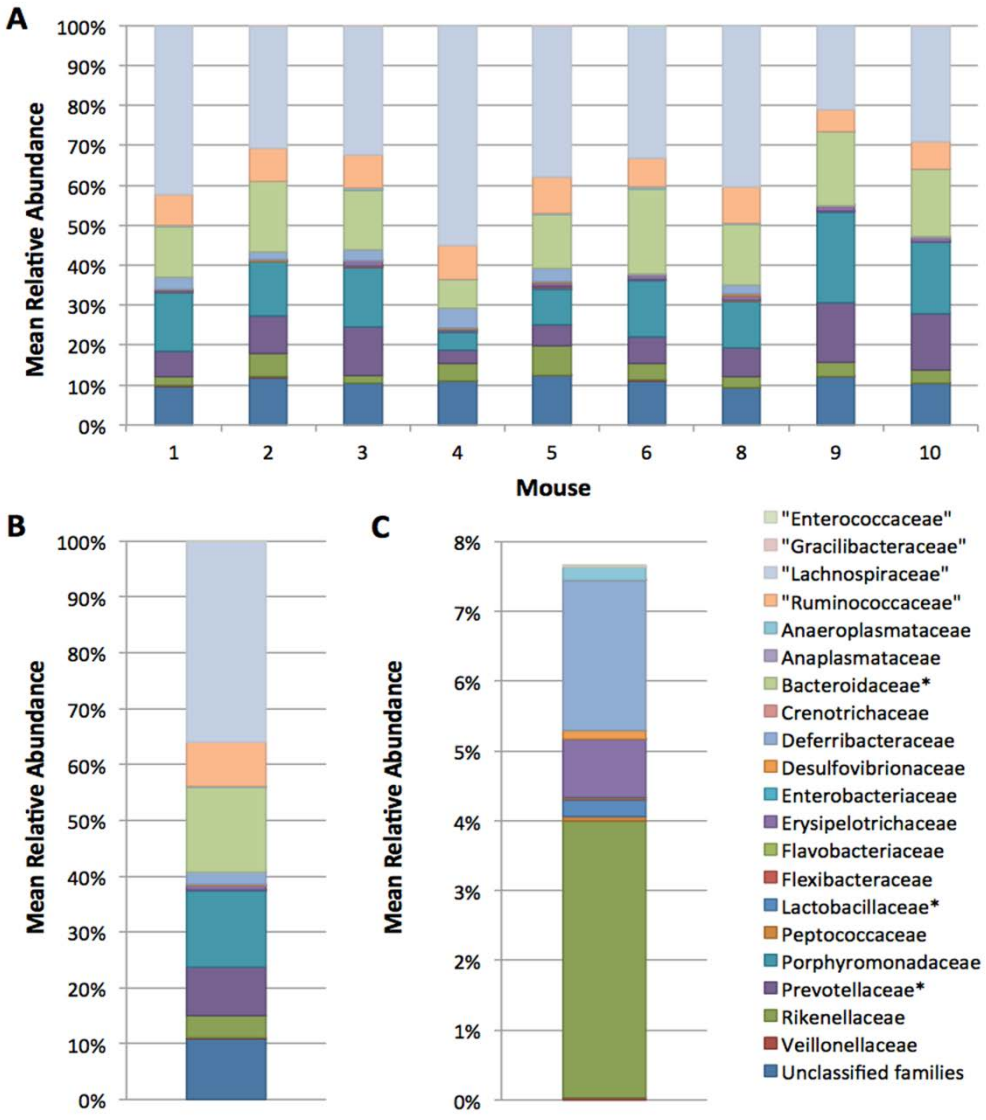


Figure 13. Relative abundance of families in the A/J mouse intestinal microbiome. A: Relative abundance for individual mice at the family level. B: Mean relative abundance of families in the intestinal microbiome of 9 mice. C: Mean relative abundance after the 5 most abundant families are removed.

The "" symbol denotes families containing species capable of hydrolyzing ginsenosides.*

3.5. Discussion

RGE shows promise in preventing lung cancer (Yan, Wang et al. 2006). However, the ginsenosides responsible for the activity of RGE are poorly defined as primary ginsenosides are inactive (Figure 7). In this thesis work, we have systematically investigated the role of intestinal microbiome in activating primary ginsenosides in RGE using A/J mouse fecal lysate. To our knowledge, this is the first comprehensive study demonstrating the formation of ginsenoside F2 from Rd is the rate limiting step in ginsenoside Rb1 stepwise hydrolysis in a mouse intestinal microbiota. Furthermore, we have begun identifying the intestinal bacteria potentially responsible for this pattern of stepwise hydrolysis, which should facilitate future studies of this type.

Our results showed that secondary ginsenosides exhibited high anti-proliferation potency in LM1 cells, while primary ginsenosides and RGE did not possess this activity (Figure 7). In addition, secondary ginsenosides exhibited better intestinal permeability in the Caco-2 cell monolayers (Figure 8). In contrast to their higher activity and better permeability, secondary ginsenosides are normally present in very low abundance in RGE (less than 1%). To explain this apparent discrepancy, we performed a hydrolysis study of primary ginsenoside Rb1 using A/J mouse fecal lysate, since A/J mice were used to demonstrate the efficacy of RGE in vivo (Yan, Wang et al. 2006). Ginsenoside hydrolyzing enzymes are β-glucosidase, β-xylosidase, α-L-arabinofurano-sidase, and α-L-rhamnosidase since the sugars attached to ginsenosides are glucose, L-arabinopyranoside, L-arabinofuranoside, D-xylose, and/or L-rhamnose (Park, Yoo et al. 2010). Mammalian cells do not express these enzymes (Henrissat 1991; Henrissat and

Bairoch 1993; Henrissat and Bairoch 1996), which is consistent with our result that ginsenoside Rb1 was stable in freshly prepared A/J mouse liver, small intestine, and colon S9 fraction (Appendix C). The results of in vivo pharmacokinetic study of ginsenoside Rb1 validated the concept that microbial hydrolysis of primary ginsenosides to bioactive secondary ginsenoside were occurring *in vivo*, since oral administration of Rb1 resulted in clear blood level of Compound K (Figure 9). Therefore, the efficacy of RGE demonstrated previously in A/J mice (Yan, Wang et al. 2006) is largely explained by enzymatic activity of the intestinal microbiota.

The peak concentration of Compound K in vivo (1.5 μM in Figure 9) following Rb1 oral administration at 20 mg/kg is around 7% of its in vitro IC_{50} value (13 μg/mlL in Figure 7, equivalent to 20.9 μM, MW is 622 g/mol) in LM1 cells. Yang et. al. also reported that the peak concentration of Compound K could reach as high as 1 μM following oral administration of Compound K at 10 mg/kg in FVB mice (Yang, Wang et al. 2012). Compound K's concentration *in vivo* is expected to increase more than proportion at higher doses since it is a substrate of P-gp and it serves as a P-gp inhibitor itself (Yang, Wang et al. 2012). Therefore, it is possible to enhance the efficacy of RGE by increasing the bioavailability of Compound K.

To enhance the therapeutic efficacy of RGE, recent investigations focused on maximizing the production of active ginsenosides by microorganisms (Chi, Kim et al. 2005; Cheng, Na et al. 2007; Son, Kim et al. 2008; Zhou, Zhou et al. 2008). However, these microorganisms may not be present in the intestinal microbiome, rendering these results less physiologically relevant. Furthermore, two key questions need to be

addressed before trying to enhance the efficacy of RGE: (1) what is the rate limiting step of RGE hydrolysis and (2) which bacterial β-glucosidases catalyze this reaction. In a sequential reaction system, the rate-limiting step is defined as the slowest step. Here, we determined the metabolite formation rates in an in vitro hydrolysis assay by optimizing protein concentration, incubation time, and sampling time such that only one metabolite would be detected. Formation of ginsenoside F2 from Rd was found to be the rate-limiting step. It has been reported that Rb1 can be metabolized to Rh2, Rg3, Compound K and gypenoside LXXV by different microorganisms (Hou, Xue et al.), and the results of our study indicated that we should pay attention to the conversion to Compound K and especially the formation of F2 in mice The intestinal microbiome structure is provided here because this activity may change in different labs and animal models, due to differences in species, food, and etc. Hence, the publication of our microbiome structure allows others to compare their results with our results. Taken together, the formation rate of ginsenoside F2 would likely determine the rate and extent of active ginsenoside production and hence greatly impacts the efficacy of RGE. Future studies are ongoing to identify specific bacterial β-glucosidases hydrolyzing ginsenoside Rd to F2.

A viable approach to enhance therapeutic efficacy of RGE is the manipulation of bacterial β-glucosidases responsible for ginsenoside F2 formation. Recent studies have shown that the abundance of Lactobacillus and Bifidobacteria as well as activities of β-glucosidase(s) can be significantly increased by probiotic intervention (Marteau, Pochart et al. 1990; McBain and Macfarlane 1998; Goossens, Jonkers et al. 2003). Since

members of these genera have been frequently used as probiotics and are known to produce β-glucosidases (McBain and Macfarlane 1998), an appropriately designed regimen combining probiotic intervention and RGE administration may serve as a novel approach for the chemoprevention of lung cancer.

In this thesis work, we have demonstrated the metabolic pathway of ginsenoside Rb1 and the rate limiting reaction of the stepwise hydrolysis of RGE in A/J mouse fecal lysates. However, different metabolic pathways have been reported using various individual microorganisms (Cheng, Kim et al. 2006; Chen, Yang et al. 2007; Cheng, Na et al. 2007; Chen, Yang et al. 2008; Chen, Yang et al. 2008; Cheng, Na et al. 2008) with unknown in vivo relevance. Therefore, it is imperative to characterize the intestinal microbiome of relevant mouse models for pre-clinical studies involving RGE and other compounds. Here, we characterized the A/J mouse intestinal microbiome (Figure 13) to aid future studies comparing changes in the microbiome and its effect on conversion of ginsenosides.

The interest in this mouse model stems from its application as a pre-clinical model to determine the efficacy of RGE as a chemoprevention agent against lung cancer. Moreover, the mouse intestinal microbiome resembles that of the human microbiome in terms of taxa present (Ley, Peterson et al. 2006). Of the 5 families that comprise over 80% of the mean relative abundance of the A/J mouse intestinal microbiome (Figure 13B), Bacteroidaceae and Prevotellaceae contain species capable of hydrolyzing ginsenosides present in RGE (Hasegawa, Sung et al. 1997; Chi and Ji 2005). Lactobacillaceae, less than 1% of the A/J mouse intestinal microbiome, also contains

Lactobacillus species with ginsenoside hydrolysis activity (Chi and Ji 2005). Additionally, several Lactobacillus species have been identified as generally-recognized-as-safe (GRAS) food microorganisms by the US FDA (http://www.fda.gov/Food/IngredientsPackagingLabeling/GRAS/MicroorganismsMicrobialDerivedIngredients/default.htm). Identification of GRAS microorganism would be highly desirable for use as probiotics to be administered to humans in conjunction with RGE to enhance the production of active secondary ginsenosides, thereby improving efficacy.

In summary, we have demonstrated that in vivo conversion of primary ginsenosides in RGE to the secondary and bioactive ginsenoside Compound K was only mediated by microbial glycosidases. The formation of F2 from Rd was found, for the first time, to be the rate-limiting step in the biotransformation of Rb1 to Compound K. The intestinal microbiome of the A/J mouse capable of producing Compound K was characterized preliminary, which formed the basis for future studies of how changes in intestinal microbiome will impact bioactivities of RGE *in vivo* (**Aim II**).

Chapter 4 Determination of impacts of RGE administration on bacterial glycosidase activity and intestinal microbiome in A/J mouse

4.1. Abstract

In this study, the impact of RGE administration on activity of bacterial glycosidases and diversity of the intestinal microbiome in A/J mice was investigated. The dose dependent enzymatic functions were tested by giving RGE orally at low (5 mg/kg), medium (50 mg/kg) and high (500 mg/kg) dose. The impact of RGE administration on enzymatic activity was also evaluated. Glycosidase activity in response to RGE treatment exhibited a dose dependent manner. The optimal dose of RGE was found to be 50 mg/kg. Oral administration of RGE at 50 mg/kg for 7 days significantly enhanced glycosidase activity of A/J mice by exhibiting a markedly change ($p<0.0001$) in metabolite formation rate in the treatment group. Large inter-subject variability of glycosidase activity was observed among the A/J mice. While none of the mice in the Dose Response study exhibited changes in the microbiome following RGE treatment, distinct changes in microbiome composition and richness were observed after 50mg/kg RGE treatment in the RGE Interaction study. We also identified significant changes in relative abundance of the genus *Lactobacillus*, which contains species that can hydrolyze RGE. In conclusion, measurement of bacterial glycosidase activity may help differentiate potential responders and non-responders to chemoprevention activity of RGE, informed by the large inter-subject variability of glycosidase activity among A/J mice.

4.2. Introduction

Multiple chemopreventative agents, such as ginsenosides (Yun 1991; Yun 1999; Yan, Wang et al. 2006), polyphenols (Baumeister, Reiter et al. 2012; Cimino, Sortino et al. 2012; Stagos, Amoutzias et al. 2012), and isothiocyanates (Hecht 2000; Hecht, Kassie et al. 2009) have shown enormous potential and red ginseng extract (RGE) is a major one among these agents. The efficacy of RGE, a mixture of ginsenosides, has been examined extensively in the last 3 decades against lung cancer chemoprevention (Yun TK 1983; Yun TK 1987; SH 1988; Yun 1991; Yun TK 1995; Yan, Wang et al. 2006). Numerous studies have demonstrated that RGE is a potent agent for the chemoprevention of lung cancer. A recent study by Yan et al., for example, demonstrated that RGE in drinking water significantly inhibit pulmonary adenoma formation and growth in A/J mice (Yan, Wang et al. 2006).

Ginsenosides, the major components in RGE, are a class of steroid glycosides and triterpene saponins. The glycosidic bonds of ginsenosides are prone to hydrolysis by bacterial glycosidases, releasing the deglycosylated ginsenosides or aglycones with different anti-cancer activities (Yang, Wang et al. 2012; Niu, Smith et al. 2013) Bioactivation of RGE via hydrolysis by bacterial glycosidases is required for the *in vivo* efficacy as the presence of deglycosylated ginsenosides in RGE is negligible (Kim, Ha et al. 2007). In **Aim I**, formation of ginsenoside F2 from Rd was found to be the rate-limiting step in the biotransformation of Rb1 to Compound K. Therfore, manipulation of bacterial glycosidases responsible for ginsenoside F2 formation could be used as a viable approach to enhance therapeutic efficacy of RGE.

Bacterial glycosidases, or glycoside hydrolases, are classified into EC 3.2.1 superfamily (http://www.cazy.org/Glycoside-Hydrolases.html). Of the 195 classes in the EC 3.2.1 super family, β-glucosidase (EC 3.2.1.21) is of particular interest (http://www.enzyme-database.org/query.php?ec=3.2.1.21). Enzymes in this class hydrolyze terminal, non-reducing beta-D-glucosyl residues with release of beta-D-glucose, in accordance with structures of ginsenosides discussed in this thesis (Shin, Lee et al. 2015). (Kim, Ha et al. 2007). Presumably, A/J mice with higher bacterial glycosidase activity tend to produce more active ginsenosides and possibly show higher chemoprevention efficacy *in vivo*, while A/J mice with lower bacterial glycosidase activity are less likely to show chemoprevention efficacy. Additioanlly, due to the complexity of glycosides (glucose, arabinopyranose, arabinofuranose, xylopyranose etc.) (Nag, Qin et al. 2012) attached to the backbone and the breaking order of glycosidic bonds, various deglycosylated ginsenosides or aglycones could be produced. Numerous studies have demonstrated that production of these deglycosylated ginsenosides or aglycones is dependent on certain bacteria in the microbiota (Figure 7) (Chi and Ji 2005; Chen, Yang et al. 2008; Hou, Xue et al. 2012; Kim, Kim et al. 2012; Yang, Wang et al. 2012). For example, *Eubacterium sp.* A-44 hydrolyzed ginsenoside Rb1 to ginsenoside Rd, F2, and Compound K (Akao T 1998). Meanwhile, ginsenoside Rh2, but not ginsenoside Compound K, was produced via route Rb1-Rd-F2-Rh2 by *Lactobacillus delbrueckii*, and *Leuconostoc paramesenteroides* (Chi and Ji 2005). Prediction of metabolic pathways of ginsenosides in a particular individual may be difficult due to the diversity of gut microbiota. Consequently, the concentration of active ginsenosides produced by gut microbiota may vary significantly among individual subjects, posing a major challenge to

the clinical testing of certain ginsenosides when given orally. It is essential, therefore, to characterize the intestinal microbiome and determine the bacterial glycosidases activities of A/J mouse since the chemoprevention efficacy of RGE may depend on the proper microbiota structure of a particular individual.

In the present study, we (1) determine the dose response of RGE with regards to microbial metabolism of ginsenosides and the intestinal microbiome composition; and (2) employed an ginsenoside hydrolysis assay to study the activity of bacterial glycosidases in response to RGE treatment at the optimal dose; and (3) use 16s rRNA pyrosequencing to define the impact of RGE administration, when given at the optimal dose, to A/J mouse intestinal microbiome.

4.3. Materials and methods

4.3.1. Chemicals and reagents

Ginsenoside Rd and F2 (>99% pure) were purchased from LKT Laboratories (St. Paul, MN). Testosterone was purchased from Sigma-Aldrich (St. Louis, MO). Red Ginseng powder was kindly provided by Dr. Ming You from Medical College of Wisconsin. BCA protein assay kit was purchased from Thermo Scientific (Rockford, IL). Other chemicals (analytical grade or better) were used as received.

4.3.2. Animals

The study was approved by the Institutional Animal Care and Use Committee at the University of Houston. Male A/J mice (20-25 g) were purchased from Harlan Laboratory (Indianapolis, IN) at 6-8 weeks of age. They were housed individually in an

environmentally controlled room (temperature: 25 ± 2°C, humidity: 50 ± 5%, 12-hr light-dark cycle) for one week before the experiments. Drinking water and diet were supplied *ad libitum*. The body weight of mice was measured every day for the duration of the study.

4.3.3. Study Design

4.3.3.1 Dose Response Study

Male A/J mice (n=16, RGE naive) were randomly divided into four groups: low, medium, high dose group and the control group, with four mice in each group (Figure 14).

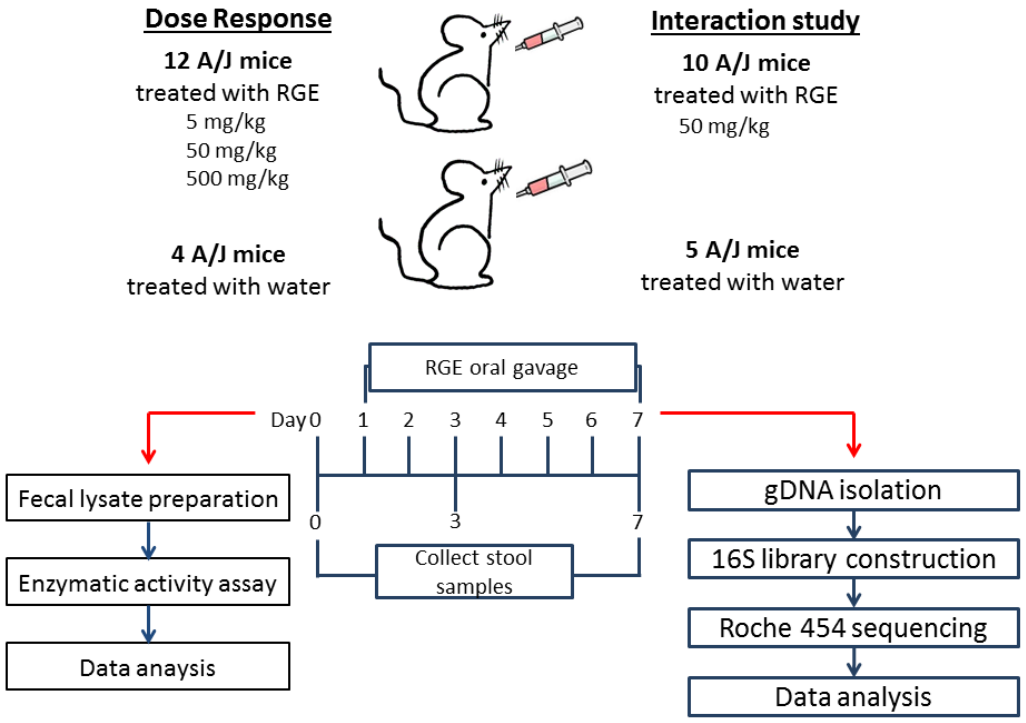


Figure 14. Study design of the Dose Response study and the RGE Interaction study.

Each A/J mouse was orally garaged with RGE or water and the stool samples were collected for further analysis (glycosidase activity assay and 16S rRNA pyrosequencing).

Mice in the low, medium and high dose group were orally gavaged with RGE daily at 5, 50, and 500 mg/kg, respectively, for seven days. Mice in the control group received the same amount of water by oral gavage for seven days. Fresh stool samples from each A/J mouse were collected at day 0, day 3, day 7 and stored at -80°C until analysis.

4.3.3.2 RGE Interaction study

Male A/J mice (n=10, RGE naive) in the treatment group received RGE daily at 50 mg/kg by oral gavage (Figure 14). Male A/J mice (n=5, RGE naive) in the control group received the same amount of water daily. Fresh stool samples of each A/J mouse were collected at day 0, day 3, day 7 and stored at -80°C until analysis.

4.3.4. Ginsenoside hydrolysis assay

The fecal lysate was prepared exactly as previously described in Section 3.3.5.1. Fecal lysates from all the A/J mice in the treatment group (RGE group) were pooled together by volume. Likewise, fecal lysates from all the A/J mice in the control group were also pooled together by volume. Activity of ginsenoside hydrolyzing enzymes was determined by measuring the metabolite formation rates of ginsenoside F2. The experiments were performed in triplicates as aforementioned in Section 3.3.5.2 with the only exception that 10 μg/mL ginsenoside Rd was used as the enzyme substrate. The same UPLC method described in Section 3.3.6 was applied for the sample analysis. The metabolite formation rates of ginsenoside F2 (the rate limiting step in the production of Compound K) were normalized by protein concentration and reaction time. Recovery of total ginsenosides was determined to ascertain that the mass balance fell in the range of 80% to 120% (i.e., no other metabolites were present in significant quantities).

4.3.5. 16S rRNA gene pyrosequencing

Fecal samples (500 µL) were homogenized in 1.5 ml of Fecal Bead Solution (MoBio, Carlsbad, CA), centrifuged for 5 min at 2000x g, and 500 µL supernatant was transferred to PowerBead Tubes (MoBio, Carlsbad, CA). Samples were heated for 10 min at 65°C and 95°C to aid bacterial lysis. Genomic DNA was isolated using PowerSoil DNA Isolation Kit purchased from MoBio Laboratories (Carlsbad, CA) starting at step 2 of the manufacturer's protocol. DNA concentration and purity were determined by spectrometry on the NanoDrop ND-2000 (Thermo Scientific). Variable regions 3 to 5 (V3-V5) of the 16S rRNA gene were amplified using barcoded primers 357F (CTGCTGCCTCCCGTAGG) and 926R (CCGTCAATTCMTTTRAGT). Polymerase chain reaction (PCR) (total volume 20 µl) contained AccuPrime PCR Buffer II (Invitrogen, Carlsbad, CA), AccuPrime Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA), 4 nM barcoded primers, and 2 µl of template DNA. PCR was performed on Mastercycler ep gradient S (Eppendorf, Hamburg, Germany) with the following cycling conditions: 2 min at 95°C and 30 cycles of 20 sec at 95°C, 30 sec at 50°C, and 5 min at 72°C. DNA concentration of PCR products was determined by PicoGreen (Invitrogen, Carlsbad, CA) and on an Agilent Bioanalyzer 2100 DNA 1000 chip (Agilent Technologies, Santa Clara, CA). Pyrosequencing was performed on Roche 454 FLX Titanium platform (Branford, CT) according to the manufacturer's instructions.

4.3.6. Glycosidase activity data analysis

The fold difference of glycosidase activity for each A/J mouse after treatment of RGE or water for 7 days, using ginsenoside Rd as the probe substrate, was obtained according to the following equation (Eq.2):

$$\text{fold difference} = \frac{glycosidase\ activity\ at\ day\ 7}{glycosidase\ activity\ at\ day\ 0} \qquad (2)$$

where glycosidase activity is the metabolite formation rate of ginsenoside F2 for a particular mouse. Metabolite formation rate is calculated as the metabolite concentration divided by protein concentration of the fecal lysate and reaction time. A/J mice with at least 3 fold difference of glycosidase activity were termed as responders (R). Non-responders (N) were A/J mice with less than 3 fold difference of glycosidase activity.

4.3.7. Statistical analysis

4.3.7.1 Glycosidase activity data

The data in this study are presented as mean$\pm$SD, if not specified otherwise. For the activty of ginsenoside hydrolyzing enzymes, significance is assessed by Student's-test, one way ANOVA with and without Tukey's post hoc test. A p-value of < 0.05 is considered statistically significant.

4.3.7.2 16S rRNA gene compositional analysis

Read pairs were de-multiplexed by the unique molecular barcodes and merged using Usearch[1] v7.0.1001.16S rRNA gene sequences were binned into Operational Taxonomic Units (OTUs) at 97% similarity through the UPARSE[3] algorithm.

Representative sequences from each OTU were aligned to the SILVA rRNA sequence database [2]. Sequences were normalized to 3747 and 1000 reads/sample for the Interaction and Dose Response studies, respectively. Alpha-diversity, beta-diversity, and taxa analyses were generated using an in-house R[4] pipeline.

4.4. Results

4.4.1. Changes in activity of ginsenoside hydrolyzing enzymes

4.4.1.1 Activity of ginsenoside hydrolyzing enzymes in the Dose Response study

Glycosidase activity of each A/J mouse in the dose response study is presented in Figure 15 and Table 2. In control group, glycosidase activity of mouse #1 and #2 did not change significantly from day 7 to day 0, while mouse #3 and mouse #4 showed significant decrease on day 7 compared to day 0 ($p<0.05$). In the low dose (5 mg/kg oral gavage once daily) group, a different pattern was observed. RGE administration diminished the glycosidase activity of mouse #1 ($p<0.0001$) and #2 ($p<0.05$) significantly, while mouse #3's glycosidase activity remained unchanged after RGE treatment. Glycosidase activity of mouse #3, however, increased more than 7 fold on day 7 compared to day 0 ($p<0.0001$). In contrast to the large variations of glycosidase activity from mouse to mouse in the control and the low dose group, all the A/J mice in the medium (50 mg/kg) and high dose (500 mg/kg) group showed a significant enhancement of glycosidase activity when treated with RGE for 7 days (Figure 15C Figure 15D). Additionally, differences of glycosidase activity of all the A/J mice in the medium and the high dose group were close to, if not more than 3 fold (Table 2), a cut-

off value chosen to distinguish RGE responders (R) and non-responders (NR). None of the A/J mice in control group was RGE responders. Interestingly, only mouse #4 in the low dose group was RGE responders. To determine the differences of glycosidase activity among all the dose groups, one way ANOVA with Tukey's post hoc test was performed. The mean fold difference of glycosidase activity in the medium dose group is significantly higher than control group ($p<0.05$), as illustrated in the dose response relationship (Figure 16) and no statistical difference was observed among all the other groups. Taken together, glycosidase activity in response to RGE treatment exhibited a dose dependent manner and oral administration of RGE at 50 mg/kg (the optimal dose) for 7 days significantly enhanced glycosidase activities derived from microflora of A/J mice.

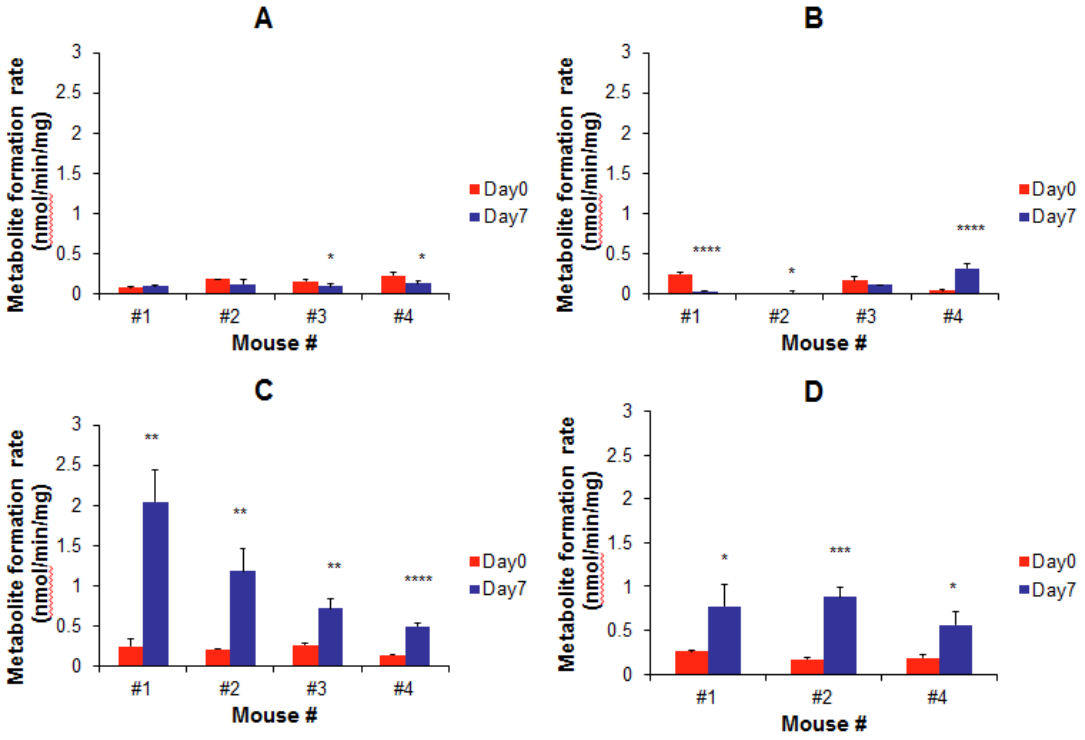


Figure 15. Glycosidase activity of each A/J mouse in the Dose Response study.

Glycosidase activity is represented by metabolite formation rate of ginsenoside F2. Metabolite formation rate is calculated as the metabolite concentration divided by protein concentration of the fecal lysate and reaction time. Panel A, B, C, and D stands for the glycosidase activity of A/J mice in the control, 5 mg/kg, 50 mg/kg and 500 mg/kg dosing group at day 0 and day 7, respectively. Data are presented as mean ± S.D.; n=3. The "" symbol indicates p<0.05, "**" p<0.01, "***" P<0.001, "****" p<0.0001, analyzed by Student's t-test. Mouse #3 in the 500 mg/kg group is excluded because it did not produce enough feces on day 7.*

Table 2. Fold differences of glycosidase activity of each A/J mouse in the Dose Response study.

Group	Mouse #			
	1	2	3	4
Control	1.21/N	0.69/N	0.59/N	0.57/N
5 mg/kg	0.15/N	0.72/N	0.67/N	7.31/R
50 mg/kg	8.78/R	6.09/R	2.94/N	4.94/R
500 mg/kg	2.96/N	5.21/R	--	3.05/R

A/J mice with at least 3 fold difference of glycosidase activity were termed as responders (R); A/J mice with less than 3 fold difference of glycosidase activity were coined as non-responders (N). Mouse #3 in the 500 mg/kg group was excluded from the study because it did not produce enough feces on day 7.

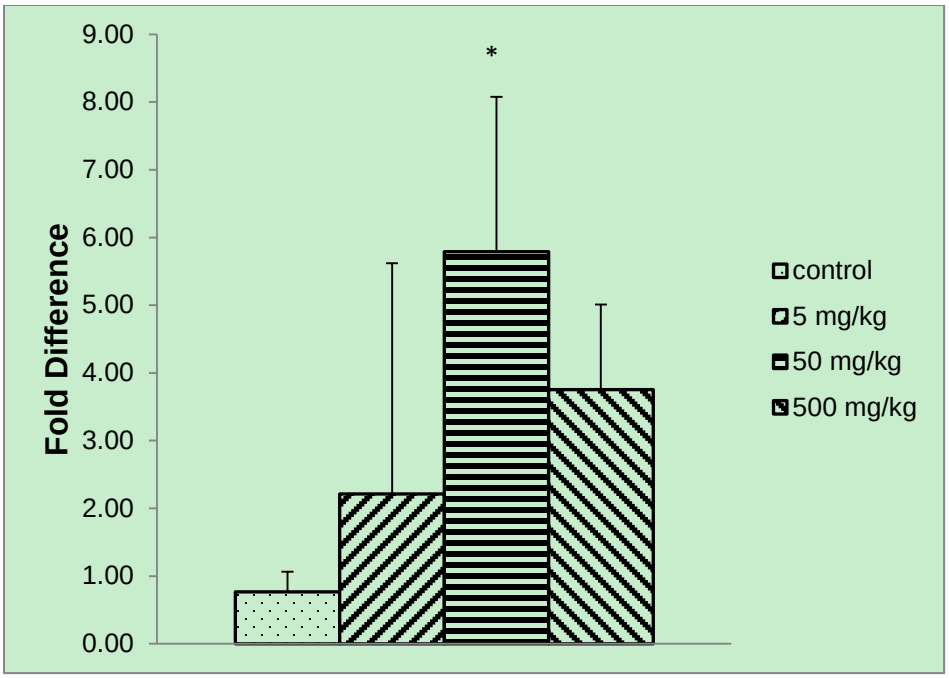


Figure 16. Dose dependent glycosidase activity of A/J mouse in response to RGE treatment.

The mean fold difference of glycosidase activity in the 50 mg/kg dose group is significantly different from the control group, analyzed by one-way ANOVA and Tukey's post hoc test, $p<0.05$.

4.4.1.2 Activity of ginsenoside hydrolyzing enzymes in the RGE Interaction study

The RGE Interaction study (n=10) was designed to determine the impact of RGE administration to glycosidase activity since the sample size in the Dose Response study was relatively small (n=4). Remarkable enhancement of glycosidase activity of all the A/J mice was observed as shown in Figure 17A. Notably, fold differences of glycosidase activity ranged from 1.1 to 27.57, suggesting large inter-subject variations among the A/J mice (Table 3). A closer look at the data revealed that 5 A/J mice (#1, 3, 6, 7, and 9) were RGE responders, whereas the rest 3 A/J mice (#2, 4, and 8) were non-responders (Table 3). The overall glycosidase activity of control and RGE treatment group was provided by pooling fecal lysate from 5 A/J mice in the control group and 8 A/J mice in the RGE treatment group respectively and the glycosidase activity was determined accordingly. No correlation of glycosidase activity and water administration was seen in control group (Figure 17B). Conversely, positive correlation of glycosidase activity and RGE treatment was demonstrated by a substantial change ($p<0.0001$) in metabolite formation rate in RGE treatment group (Figure 17B). Given the fact that glycosidases are required to produce the active metabolites, such as Compund K, these results suggest that activation of RGE against lung cancer depend on the proper microflora structure in the gut.

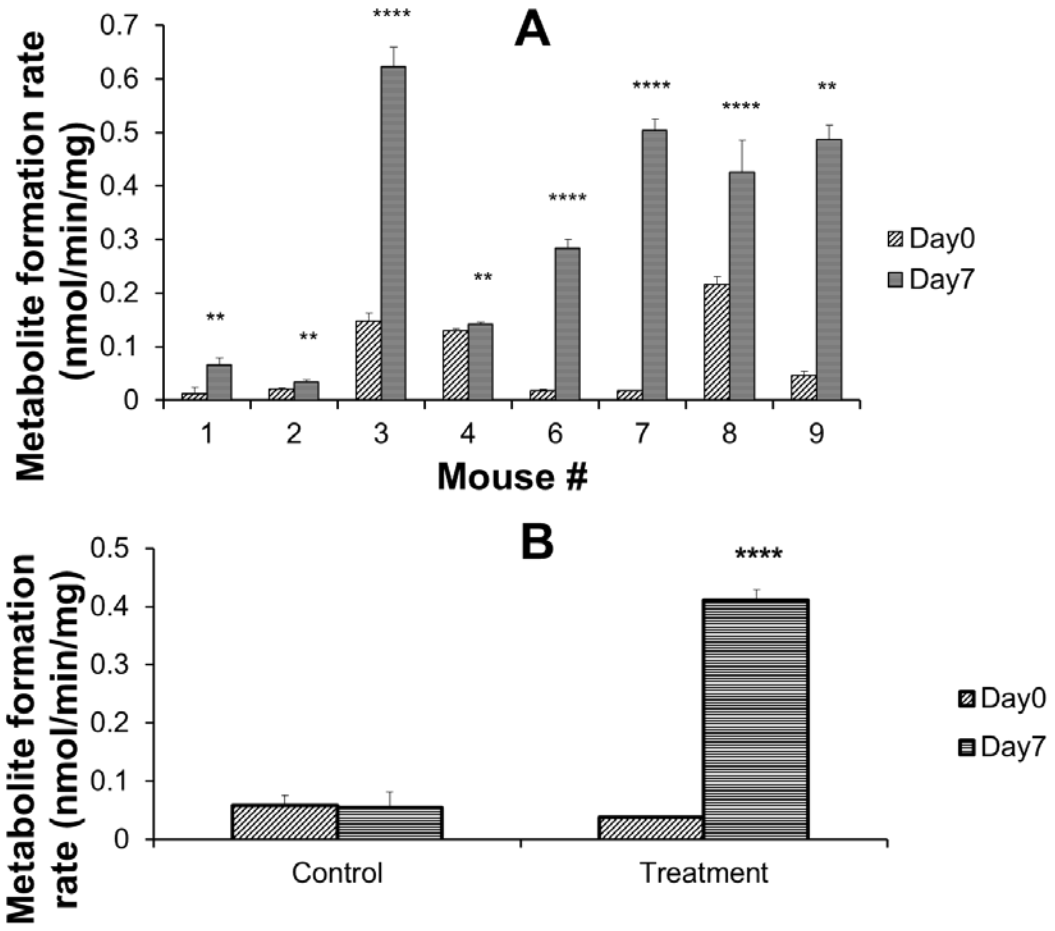


Figure 17. A: Glycosidase activity of each A/J mouse in the RGE Interaction study. B: Glycosidase activity of pooled fecal lysate in the RGE Interaction study.

Data are presented as mean ± S.D.; n=3. The "" symbol indicates p<0.05, "**" p<0.01, "***" P<0.001, "****" p<0.0001, analyzed by Student's t-test. Fecal lysate of each mouse in the treatment (RGE) group or the control group was pooled together. Mouse #5 and #10 did not produce enough feces on day 7; therefore, they were excluded from the study.*

Table 3. Fold differences of the glycosidase activity of each A/J mouse in the RGE Interaction study

	Mouse #									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Fold Difference	3.49	1.61	4.16	1.10	--	15.08	27.57	1.96	10.45	--
RGE Response	R	N	R	N	--	R	R	N	R	--

The definition of fold difference, R and N are the same as those in Table 2.

"--"indicates that the mouse was excluded from the study because no enough feces were produced on day 7.

4.4.2. Membership and relative abundance of the A/J mouse intestinal microbiota

4.4.2.1 RGE had little to no impact on the intestinal microbiome of A/J mice in the Dose Response study

We performed 16S rRNA sequencing to assess changes in the microbiome following treatement with RGE at 3 dosing levels. No significant differences were observed in alpha-diversity in any treatment group between day 0 and day 7 (Figure 18). While no significant differences were observed in normalized weighted UniFrac ordination, the samples collected from the medium dose treatment group on day 0 and day 7 clustered apart along PC2 (Figure 19). Furthermore, we did not observe any significant differences in abundance in the phlyum or genus level (Figure 20 and 21). This data suggested that treatment with RGE had little to no impact on the intestinal microbiome of A/J mice.

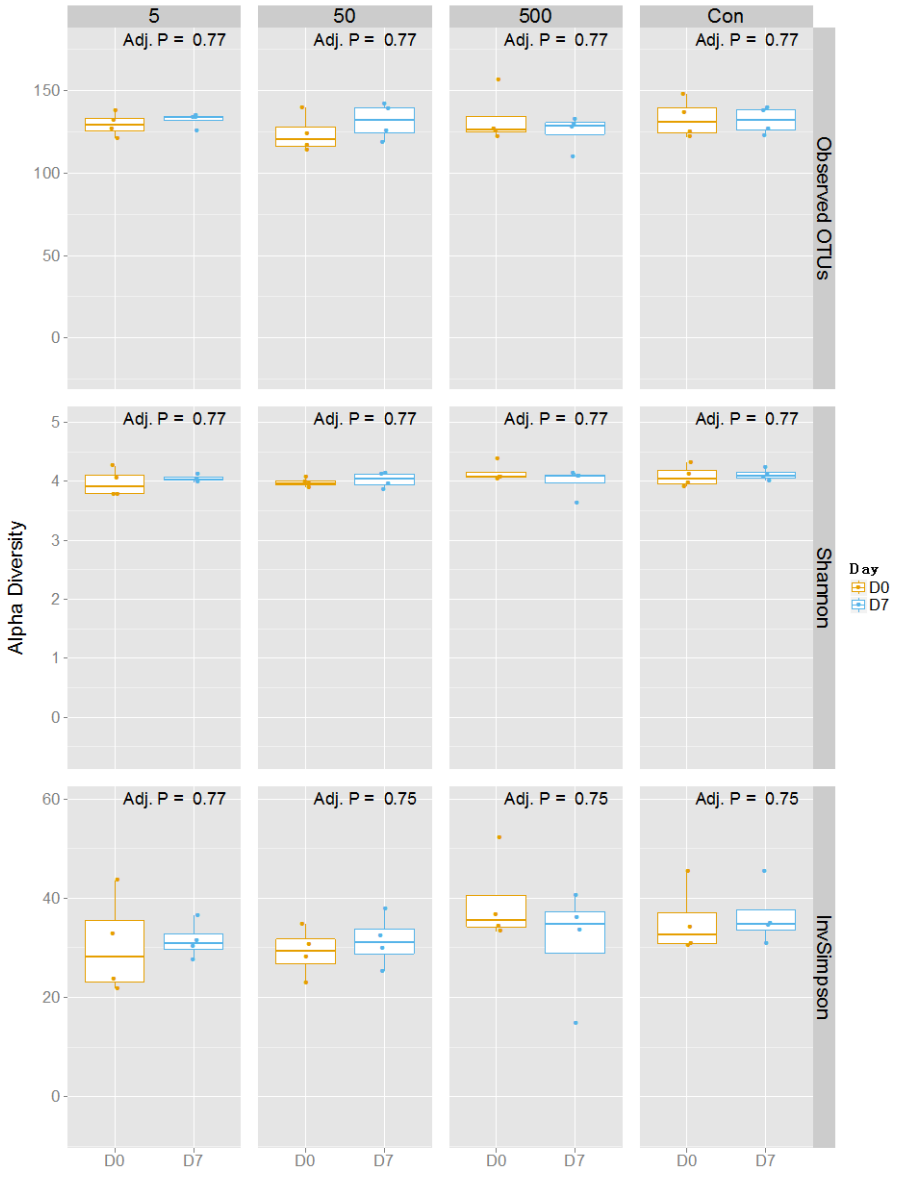


Figure 18. Alpha diversity indices plotted for each A/J mouse and dosing group in the Dose Response study.

Significant increases in richness (observed OTUs) and diversity (Shannon and inverse Simpson) are observed from day 0 to day 7. 5, 50, 500 and con are 5 mg/kg, 50 mg/kg, 500 mg/kg and control dosing groups, respectively.

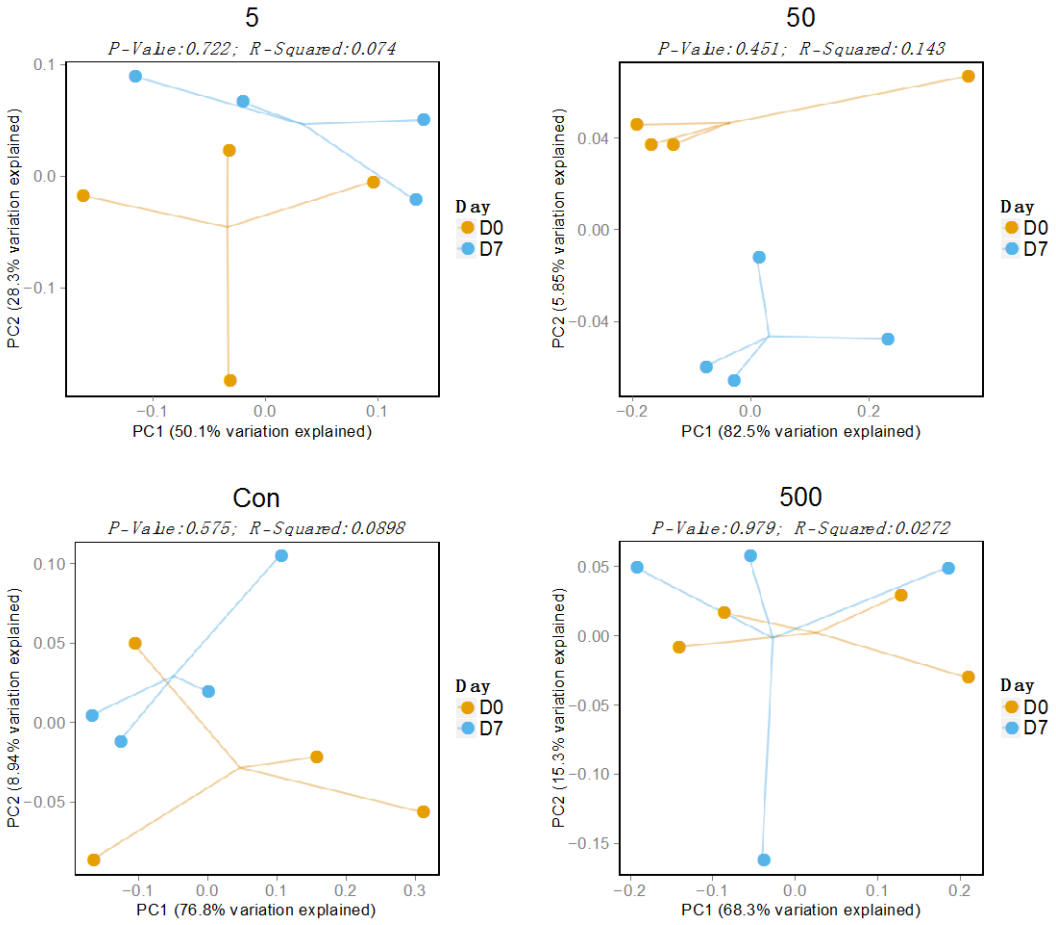


Figure 19. Normalized weighted UniFrac analysis of stool samples collected from mice pre- and post-treatment with 50 mg/kg of RGE.

5, 50, 500 and con are 5 mg/kg, 50 mg/kg, 500 mg/kg and control dosing groups, respectively.

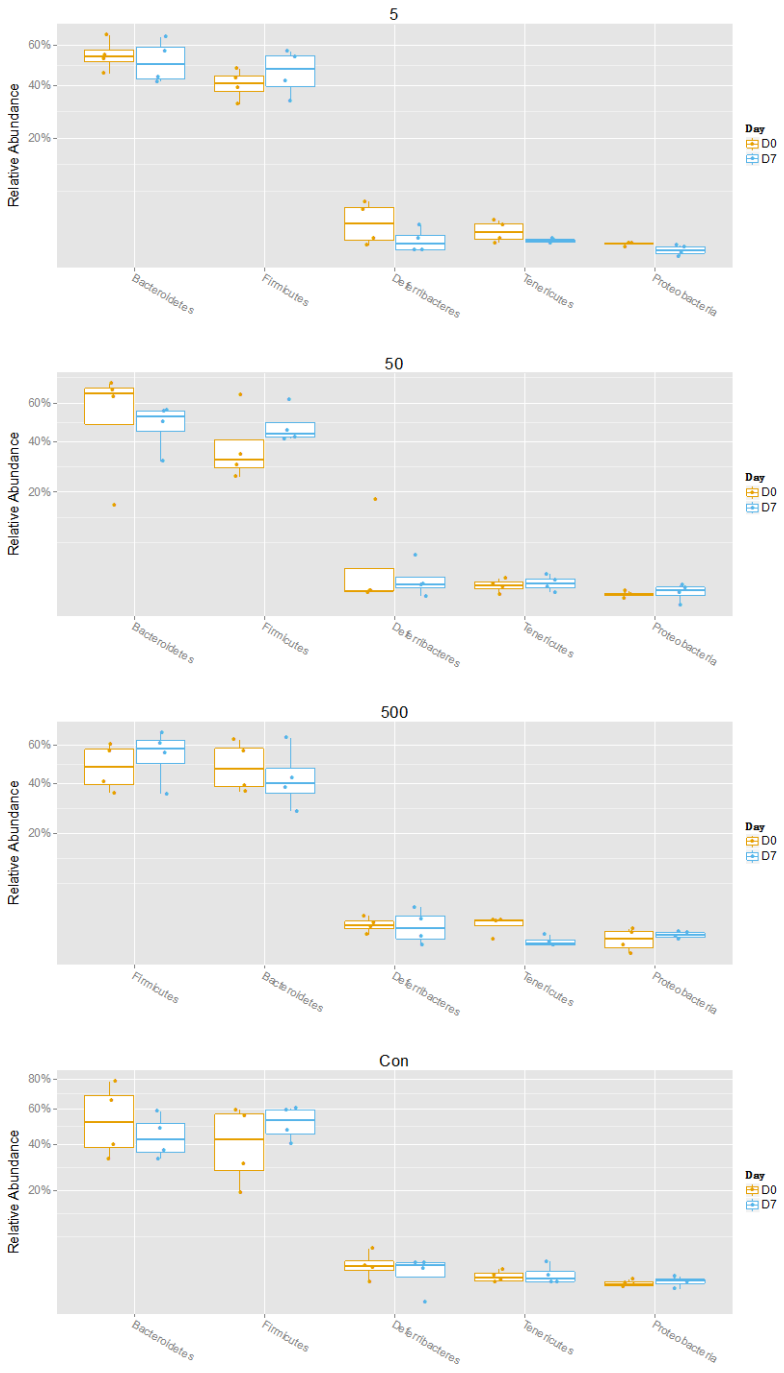


Figure 20. Relative abundance at the phylum level showed significant differences between top 5 most abundant phyla on day 0 and day 7 stratified by dosing group.

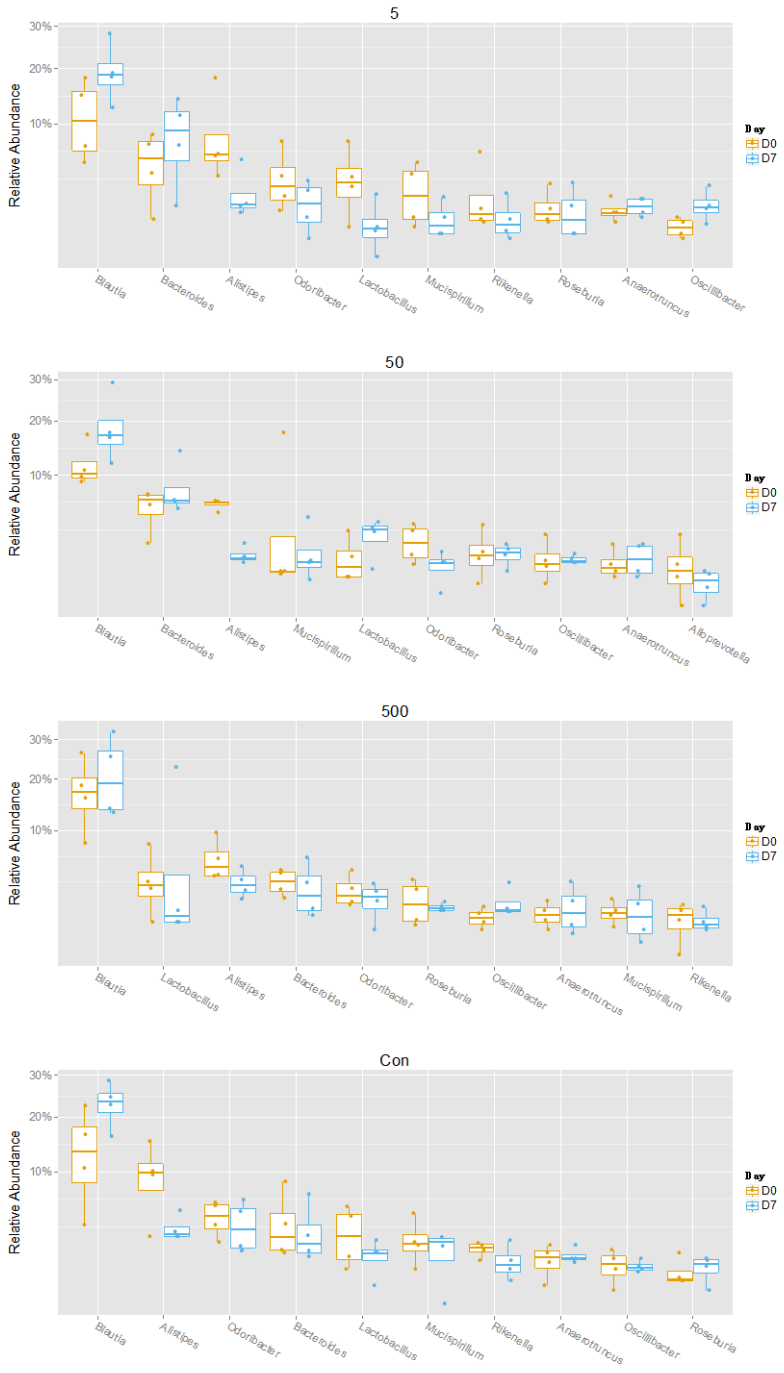


Figure 21. Relative abundance of top 10 most abundant genera at day 0 and day 7 in the Dose Response study.

4.4.2.2 RGE treatment at 50mg/kg altered the microbiome of A/J mice and increased abundance of *Lactobacillus* in the RGE Interaction study

We performed a RGE Interaction study assessing changes in the microbiome of A/J mice treated with 50 mg/kg of RGE previously. 16s ranalyses demonstrated that richness and diversity increased (Figure 22), the overall structure of the microbiota changed (Figure 23), and the Bacteroidetes:Firmicutes ratio decreased (Figure 24) following 50 mg/kg RGE treatment in contrast to our finding in the Dose Response experiment. Furthermore, analysis at genus level revealed a significant increase in *Lactobacillus*, which contains species capable of hydrolyzing ginsenosides (Figure 25). We also noted that the mice from the RGE Interaction and Dose Response studies exhibited different Bacteroidetes:Firmicutes ratios (Figure 20 and 24) at day 0 demonstrating the variability of the intestinal microbiome in mice of the same strain from the same distributor. Overall, we observed that treatment with RGE at 50mg/kg altered the microbiome of A/J, specifically increasing abundance of *Lactobacillus*, which contains species capable of hydrolyzing ginsenosides.

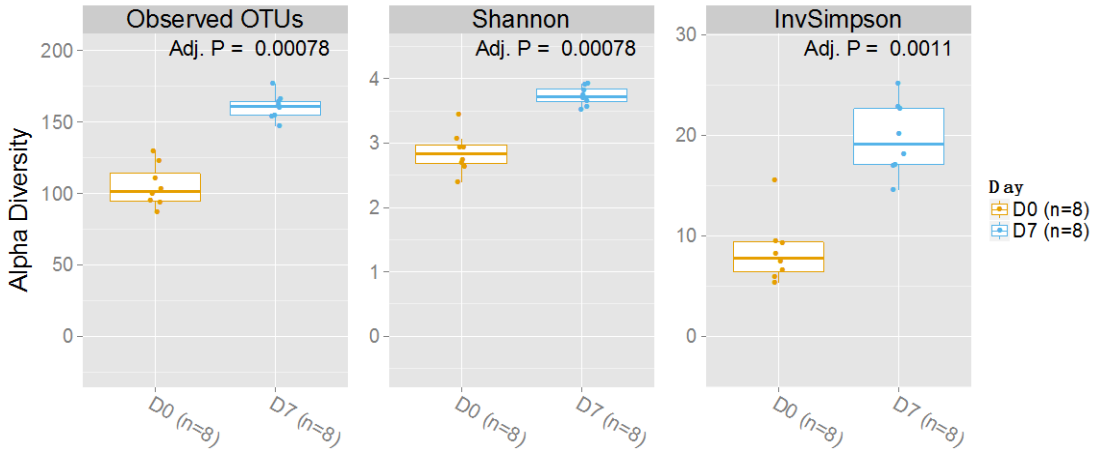


Figure 22. Alpha diversity indices plotted for each A/J mouse in the RGE Interaction study.

Significant increases in richness (observed OTUs) and diversity (Shannon and inverse Simpson) are observed from day 0 to day 7.

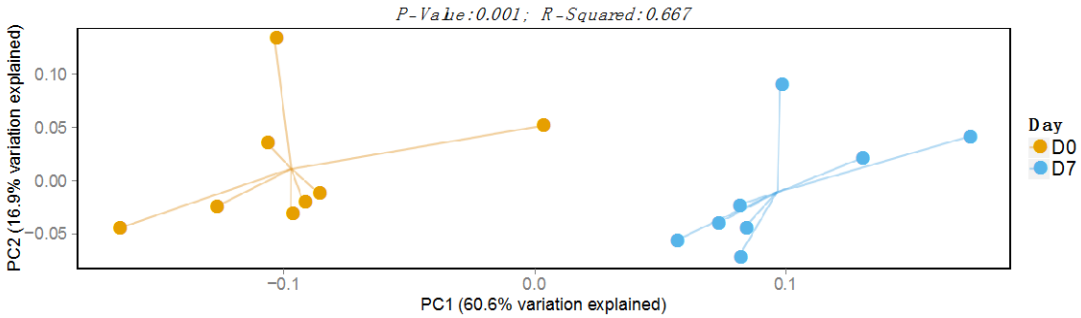


Figure 23. Normalized weighted UniFrac analysis of stool samples collected from mice pre- and post-treatment with 50mg/kg of RGE.

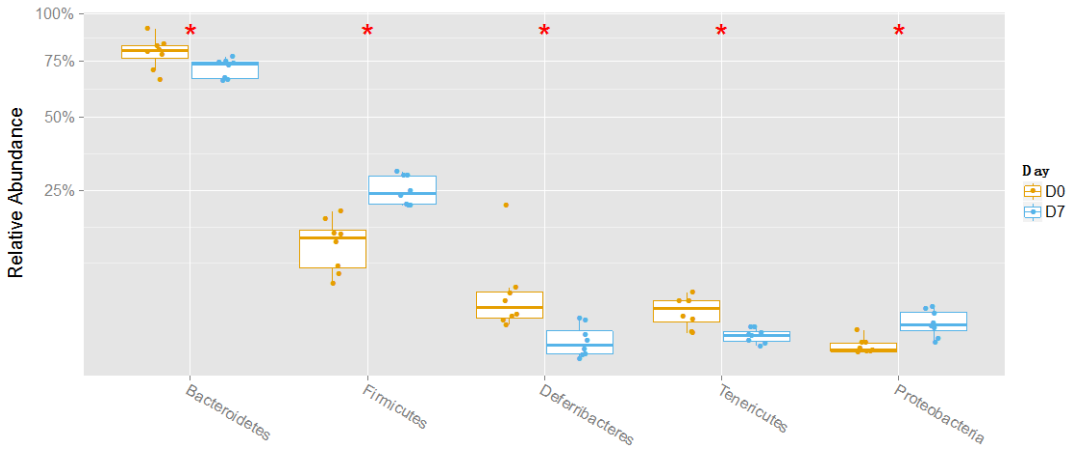


Figure 24. Relative abundance at the phylum level showed significant differences between top 5 most abundant phyla on day 0 and day 7 in the RGE Interaction study.

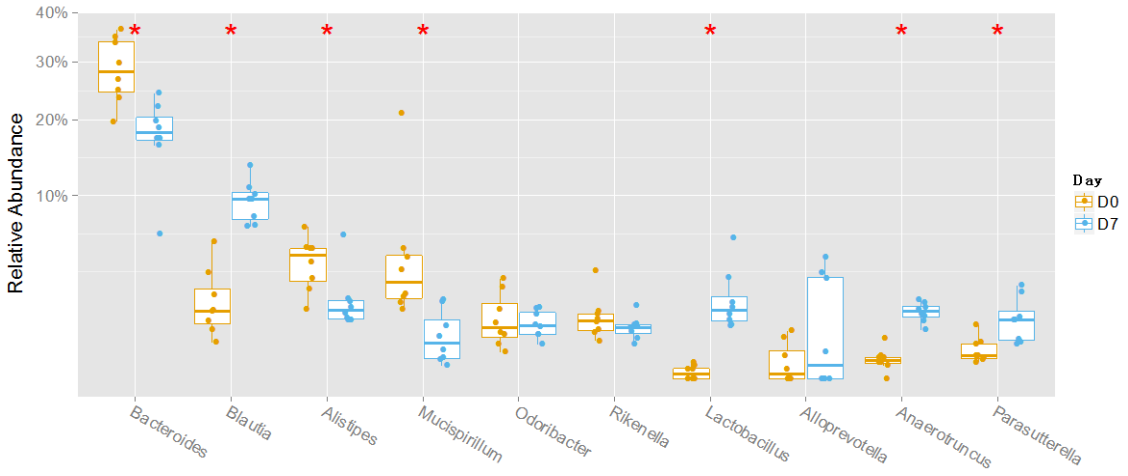


Figure 25. Relative abundance of top 10 most abundant genera at day 0 and day 7 in the RGE Interaction study.

Significant differences in relative abundance are observed between day 0 and day 7 in genera that contain species capable of ginsenoside hydrolysis including Bacteroides and Lactobacillus.

4.5. Discussion

RGE has been reported to possess non-organ-specific preventative effects against several types of cancer, including lung cancer (Yun 2001; Yun, Zheng et al. 2010). Since most of the naturally occurring ginsenosides (e.g. Rb1) require the action of bacterial glycosidases to become active ginsenosides (Niu, Smith et al. 2013), an important question is whether functional activity of bacterial glycosidases would be altered following RGE administration. A more intriguing question is whether and how RGE treatment can affect the growth of certain bacteria species in the intestinal microbiome, and whether alterations in the bacteria population would change chemoprevention efficacy of RGE. In this study, we systemically investigate the impact of RGE administration on the microbial community structure and function in the gut of A/J mice. Our results showed that activity of bacterial glycosidases was enhanced significantly following oral administration of RGE at 50 mg/kg in A/J mice. Large inter-subject variation of enzyme activity among A/J mice was also observed, suggesting that A/J mice could be classified into RGE responders and non-responders. At the microbiome level, the diversity and relative abundance of the genus *Lactobacillus*, with species capable of RGE metabolism, was also enhanced significantly. To our knowledge, this is the first comprehensive study showing that chemoprevention efficacy of RGE may rely on certain bacteria in the intestinal microbiome. Screening of the gut microbiome and functional activity of the gut microbiota in cancer patients may be necessary.

We examined how oral administration of RGE altered enzyme activity of the intestinal microbiota. A significant reduction of glycosidase activity was seen in two out of four A/J

mice in the control group of the dose response study. The 40% change, albeit statistical significant, is negligible as it would not translate in vivo. Likewise, glycosidase activity of pooled control samples in the RGE Interaction study showed no statistical change with time. Therefore, the dosing procedure had minimal effects on enzymatic functions of the intestinal microbiota. To identify *in vitro* changes that may potentially translate *in vivo*, 3 fold difference of glycosidase activity is defined as the cut-off value between RGE responders (R) and RGE non-responders (N). Selection of the empirical cut-off value is based on our *in vitro* and *in vivo* metabolism study of polyphenols, another major class of chemoprevention agents (Wang, Chen et al. 2006; Yang, Kulkarni et al. 2012).

Enzymatic functions showed a large inter-subject variability in the 5 mg/kg group of the dose response study, with a positive RGE responder, and three RGE non-responders. Medium dose of RGE (50 mg/kg), however, rendered changes of enzymatic functions in a particular pattern, indicating a positive correlation of glycosidase activity and RGE treatment at this dose. The positive correlation observed in all the mice within this dosing group is of particular interest because enhanced glycosidase activity *in vitro* may translate into higher production of Compound K *in vivo*. The reduction of enzymatic activities at the high dose (500 mg/kg) (Figure 16) is likely due to the anti-bacterial nature of RGE, although the formulations of RGE used in this study may not be the same compared to other studies (Lee, Lee et al. 1998; Bae, Han et al. 2001; Lee, Shim et al. 2006; Lee, Shim et al. 2009).

The positive correlation of glycosidase activity and RGE treatment at 50 mg/kg in the RGE Interaction study was replicated, with noted inter-subject variability of glycosidase

activity among the A/J mice. Furthermore, the RGE responders, particularly those mice with at least 10 fold enhancement of glycosidase activity in this study, have a strong potential to show *in vivo* changes of bioactivities of RGE. In contrast, the RGE non-responders may not show any *in vivo* changes of bioactivities of RGE. The inter-subject variability of enzyme activity in human could be much more substantial than animals as these mice came from the same inbred strain, were the same age and lived in the same well-controlled environment. Indeed, remarkable impacts of diets, environmental changes, and drug usage on intestinal microbiota in human have been repeatedly reported (Phillips 2009; Cotter, Stanton et al. 2012; Scott, Gratz et al. 2013). Therefore, the ginsenoside hydrolysis assay may serve as a screening tool to differentiate the potential responders of chemoprevention of RGE and non-responders.

While we did not observe significant changes at any dose in the intestinal microbiome of A/J mice in the Dose Response study, we observed changes in the intestinal microbiome of A/J mice in the RGE Interaction study, including increased diversity, changes in the overall structure of the microbiome, and increased relative abundance of *Lactobacillus*, a genera with species capable of ginsenoside hydrolysis. The differences between the 50 mg/kg groups in the Dose Response and Interaction studies could stem from the small sample size (n=4) in the Dose Response study and the difference in the composition of the microbiome at day 0. Without additional studies on the effect of RGE on the composition and structure of the microbiome, we cannot make definitive conclusion on the ability of RGE to modulate the intestinal microbiome.

In conclusion, we have demonstrated that A/J mice, an animal model of lung cancer chemoprevention studies, can be classified as RGE responders and non-responders, where the difference lies in the bacterial glycosidase activity. Screening gut microbiome and functional activity of the gut microbiota in each individual may be required for the chemoprevention studies in the future.

Chapter 5 Enrichment and identification of bacterial glycosidase(s) from A/J mouse feces that catalyze the rate-limiting step in the production of ginesnoside Compound K

5.1. Abstract

In this study, enrichment, identification and functional activity of ginsenoside hydrolyzing enzymes were investigated. Ginsenoside hydrolyzing enzymes were enriched from A/J mouse feces by a chromatographic approach. The identity of the enzymes was examined by LC-MS/MS analysis followed by gene synthesis, molecular cloning, and expression of the enzyme candidates. The functional activity of the enzymes against ginsenoside Rd was also tested. Ginsenoside hydrolyzing enzymes were enriched from A/J mouse feces by anion exchange chromatography and gel filtration chromatography. Specific activity of enriched enzymes increased from 0.757 to 27.5 μmol/mg/min. The overall enrichment fold and yield was 36 and 5.81%, respectively. The enriched enzymes were further separated by native PAGE followed by SDS-PAGE. The corresponding gel pieces were excised, digested and analyzed by LC-MS/MS. One unique peptide NGVLFPR (mass=801.4497, z=2) correlating to bacterial glycosdases

was found. Additionally, five bacterial glycosidases containing the unique peptide NGVLFPR were identified (gi: 49259571, 501268188, 147736211, 15643981, and 67696782,). Two bacterial glycosidases (gi: 501268188 and 147736211), but not the blank expression vector, hydrolyzed approximately 10 % ginsenoside Rd to F2 after 72 hour incubation at 37 °C. In conclusion, two bacterial glycosidases were enriched from A/J mouse feces and were confirmed for the first time to hydrolyze ginsenoside Rd to F2 and Compound K.

5.2. Introduction

Substantial evidences have shown that a variety of cancers can be prevented or significantly delayed. In breast cancer chemoprevention, tamoxifen and aromatase inhibitors (e.g. anastrozole, exemestane, letrozole) are effective in multiple clinical trials (Euhus and Diaz 2015). In colorectal cancer, chemoprevention efficacy of cyclooxygenase-2 (COX-2) inhibitors, nonsteroidal anti-inflammatory drugs (NSAIDs) has been established in a variety of clinical trials (Steinbach, Lynch et al. 2000; Benamouzig, Deyra et al. 2003; Hull, Sandell et al. 2013). In lung cancer, multiple chemoprevention agents, ginsenosides (Yun 1991; Yun 1999; Yan, Wang et al. 2006), polyphenols (Baumeister, Reiter et al. 2012; Cimino, Sortino et al. 2012; Stagos, Amoutzias et al. 2012), and isothiocyanates (Hecht 2000; Hecht, Kassie et al. 2009) have shown enormous potential and RGE is a major one among these agents.

Ginsenosides, the major components in RGE, are a class of compounds with different sugar moieties attached to the backbone, and the sugars can be cleaved by the

intestinal microflora to produce secondary glycosides or aglycones with different biological activities (Yun, Lee et al. 2001; Cheng, Yang et al. 2005). Since most of the naturally occurring ginsenosides (e.g., Rb1) need the action of bacterial glycosidases to become active ginsenosides, an important question is how intestinal microbiome, which is known to produce different types of glycosidases (Bae, Han et al. 2002; Kim, Lee et al. 2005), acts to produce more active ginsenosides and aglycones. In Chapter 4 (**Aim II**), we have shown that oral administration of RGE significantly enhanced bacterial glycosidases activities of A/J mice and altered the diversity, overall composition of intestinal microbiome and relative abundance of particular genera capable of RGE metabolism. However, relative abundance of bacteria community in the gut of A/J mice could only be revealed at the genus level by 16S rRNA pyrosequencing. It is therefore, imperative to define fluctuations in gut microbiome and the expression of the genes within at the species level to understand the contributions of the GI microbiota to RGE metabolism. More specifically, certain bacteria species whose relative abundance enhanced significantly after RGE treatment may serve as candidates for probiotic intervention in the chemoprevention of lung cancer. A top down approach to identify the relevant bacteria species is whole genome shotgun (WGS) pyrosequencing, where the total bacterial genomic DNA will be sequenced so that the genes encoded by the bacteria in these communities can be identified. An alternative to the top down approach is the “bottom up” approach (Figure 26), where bacterial glycosidases that catalyze the bioactivation of RGE are to be enriched and identified from A/J mouse feces.

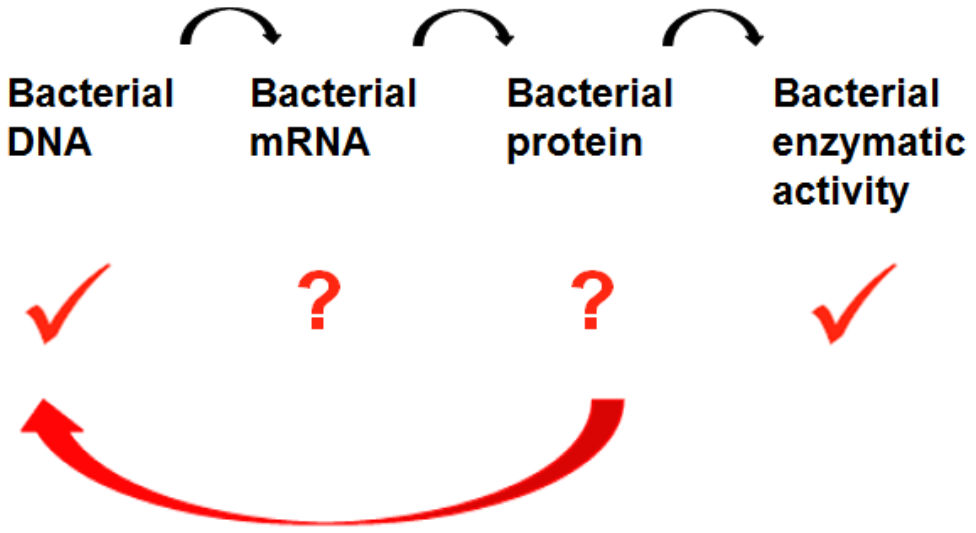


Figure 26. Bottom up approach of identification of bacteria species catalyzing bioactivation of RGE

The organisms of the identified proteins are to be determined by comparing with NCBI non-redundant database. Yan et. al. reported that a novel ginsenoside-hydrolyzing β-glucosidase that specifically transforms ginsenoside Rb1 to Compound K was successfully purified by a chromatographic approach, including ion exchange, hydrophobic interaction and CHT ceramic hydroxyapatite chromatography (Yan, Zhou et al. 2008). Similar approaches have also been applied to identify novel ginsenoside hydrolyzing glycosidases by other researchers (An, Cui et al. 2010; Quan, Wang et al. 2013). These bacteria strains, such as *Paecilomyces Bainier* sp. 229 (Yan, Zhou et al. 2008), however, are mostly isolated from ginseng plantation localities and thus have minimal physiological relevance. To find candidates for probiotic intervention in the chemoprevention of lung cancer, the bacteria strains shall be present in the A/J mouse's GI tract and are GRAS microorganisms. Therefore, A/J mouse feces would serve as the optimum source for identification of bacterial glycosidases.

In the present study, we aim at enriching and identifying ginsenoside hydrolyzing enzymes from A/J mouse feces by a classic chromatographic "bottom up" approach.

5.3. Materials and methods

5.3.1. Materials

Ginsenoside Rd, F2 and Compound K (>99% pure) were purchased from LKT Laboratories (St. Paul, MN). Testosterone, 2-Mercaptoethanol and P-nitrophenyl-β-D-pyranoside were purchased from Sigma-Aldrich (St. Louis, MO). BCA protein assay kit was purchased from Thermo Scientific (Rockford, IL). Ammonium sulfate, sodium

phosphate, dithiothreitol (DTT), sodium chloride, 4-Nitrophenyl- β-D-glucopyranoside, testosterone, sodium phosphate dibasic, and sodium phosphate monobasic were purchased from Sigma-Aldrich (St. Louis, MO). Laemmli Sample Buffer was purchased from Bio-Rad (Hercules, CA). Coomassie Brilliant Blue R250 was purchased from Bio-Rad (Hercules, California). Other chemicals (analytical grade or better) were used as received.

5.3.2. Enrichment of ginsenoside hydrolyzing enzymes from A/J mouse feces

The experimental design is shown in Figure 27. Fresh A/J mouse feces (10 g) were mixed with 25 mL ice-cold 0.1 mM PBS, vortexed for 5 min and then sonicated for 60 min in ice water. The mixture was subsequently centrifuged at 10, 000 rpm and 4°C for 30 min. The supernatant was collected (S9 fraction) and mixed with saturated ammonium sulfate (1:1 v/v) at 4°C overnight for protein precipitation. The S9 fraction was subject to centrifugation at 10, 000 rpm at 4°C. The pellets were collected and resuspended in 5 mL ice-cold sodium phosphate (10 mM, pH=7), followed by dialysis against 1, 500 mL ice-cold sodium phosphate buffer (10 mM, pH=7, with 50 mM DTT) at 4°C overnight. Upon the removal of non-dissolved fraction, the dialysate was loaded onto a Resource Q (1 mL) anion exchange column (GE Healthcare, Piscataway, NJ) and ran for 30 min with a flow rate of 1 ml/min. The elution gradient was as follows: initial, 0% A (1 M NaCl, 10 mM DTT) and 100% B (10 mM sodium phosphate, pH=8, 50 mM DTT), 0–5 min, 0% A, 5–10 min, 0-25% A, 10–15 min, 25% A, 15–20 min, 25-100% A, 20–25 min, 100% A, 25–26 min, 100-0% A, 26-30 min, 0% A. The elution fractions were collected in ice-cold glass tubes (15 sec/fraction). The active fractions against 4-

nitrophenyl β-D-glucopyranoside and ginsenoside Rd (Sigma-Aldrich, St. Louis, MO) were pooled and dialyzed against 1,500 mL ice-cold sodium phosphate buffer (10 mM, pH=7, 50 mM DTT) at 4°C overnight. The dialysate was concentrated using protein concentrators (10K MW cut-off) (Thermo Scientific, Rockford, IL) and washed with ice-cold sodium phosphate buffer (10mM, pH=7, 50 mM DTT). The mixture was then passed through a PolySep-GFC-P 4000 column (300 mm ×7.8mm) (Phenomenex, Torrance, CA) and run for 20 min with a flow rate of 1 mL/min. Proteins were eluted using sodium phosphate buffer (0.05 M Na_2HPO_4, 0.05 M NaH_2PO_4, 0.15 M NaCl, 50 mM DTT, pH=7). The elution fractions were also collected in ice-cold glass tubes (15 sec/fraction). The active fractions were again pooled and concentrated as the aforementioned procedures. The fractions were stored on ice until further analysis. Protein concentration of the enzyme fractions (IEX and GFC step only) was tested by the micro BCA protein assay kit, while all the other samples were tested by BCA protein assay kit.

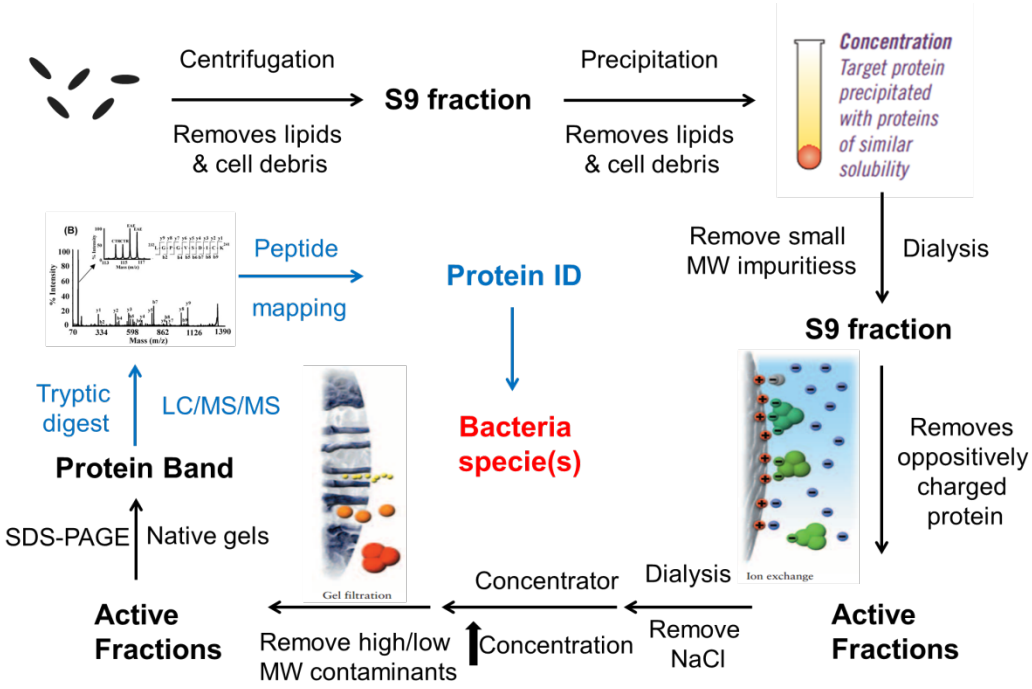


Figure 27. Study design of enrichment of the ginsenoside hydrolyzing enzymes from A/J mouce feces and back-trace the bacteria expressing these enzymes at the species level (bottom up approach).

5.3.3. Native PAGE and SDS-PAGE

Electrophoresis was carried out on a 4-10% polyacrylamide gel under native conditions. The gels were freshly prepared according to the protocol in Table 4. Fresh fractions from section 2.6 were mixed with sample buffer (40% glycerol in PBS with 1%) at a ratio of 1:3 and then loaded into two tanks of the same native gel. Running buffer was 25 mM Tris and 192 mM glycine. Upon the end of native gel electrophoresis, the two electrophoresed lanes were harvested. One lane was stained with Coomassie Brilliant Blue R250 overnight and the background color was removed by Tris-Buffered Saline and Tween 20 (TBST). The other lane was excised at 0.5 cm interval and each of those gel pieces was subject to the colorimetric activity test. Precision Plus protein™ unstained standards (10-250 kD, BioRad, Hercules, California), including three reference bands (25, 50, and 75 kD), were used as the molecular weight markers. The active gel pieces identified by the colorimetric test were minced into smaller pieces and loaded into Nanosep centrifugal tubes (100K MW cut off, VWR, Houston, TX) with ice-cold sodium phosphate buffer (10mM, pH=7, 50 mM DTT). Proteins in the gel pieces were recovered by centrifugation at 10,000 rpm at 4°C for 15 min. The recovered proteins were mixed with sample buffer (1 mL Laemmli Sample Buffer + 50 µL 2-Mercaptoethanol) at a ratio of 1:1 and heated at 95°C for 5 min. The denatured samples were subsequently loaded into a 4-10% SDS-polyacrylamide gel that was freshly prepared in accordance with Table 4. Running buffer was 25 mM Tris, 192 mM glycine and 0.1 % SDS. The same molecular weight markers were used as described above. β-glucosidase, Bacteroides fragilis NCTC 9343 (MW=95.2 kDa, Prozomix, UK) and β-1,4-Glucosidase 1A (GH1), Clostridium thermocellum (MW=52.7 kDa, NZYTech, Portugal) were used the positive

control. Upon the end of SDS-PAGE, the electrophoresed lanes were stained with Coomassie Brilliant Blue R250. Gel bands of interest were excised, and digested with trypsin for LC-MS/MS protein identification. Briefly, excised gel bands were washed with 300 µL of 100 mM NH_4HCO_3, and then 300 µL of 100 mM NH_4HCO_3/CH_3CN (50:50 v/v) repeatedly until Coomassie Brilliant Blue was completely washed off. Gel pieces were dehydrated by CH_3CN and dried in Speedvac. The gel pieces were mixed with 50 mM ammonium bicarbonate buffer (pH 7.9) and reduced with 100 mM DTT at room temperature for 30 min. The reduced cysteine residues blocked by 400 mM iodoacetamide (Sigma-Aldrich, St. Louis, MO) in dark at room temperature for 30 min. Trypsin (1 µL, sequencing grade) from Sigma-Aldrich, St. Louis, MO was added to the reaction mixtures. The reaction mixture was incubated at 37°C overnight with mild shaking. Formic acid was added to stop the digestion reaction. The tubes were speed-vac dried. The digests were re-suspended in 0.1% formic acid and 5% acetonitrile solution. The concentration of the digests was measured using NanoOrange Protein Quantitation Kit (LifeTech, Grand Island, NY).

Table 4. Standard protocol for native and SDS-PAGE gel preparation

	Native		SDS	
	Stacking	**Separation**	**Stacking**	**Separation**
Gel %	4%	10%	4%	10%
Acrylamide-bis solution (29:1)	1 mL	5 mL	1 mL	5 mL
1.5 M Tris (pH=8.8)	---	5 mL	---	5 mL
0.5 M Tris (pH=6.8, 4X)	2.5mL	---	2.5mL	---
DD H_2O	6.4 mL	9.8 mL	6.3 mL	9.6 mL
10% SDS	---	---	0.1 mL	0.2 mL
10% APS	0.1 mL	0.2 mL	0.1 mL	0.2 mL
TEMED	10 µL	20 µL	10 µL	20 µL

"---"indicates that the buffer/reagent was not used.

5.3.4. LC-MS/MS analysis

Trypsin digested samples (200 ng each) were injected onto a trap column (ChromXP C18-CL, 3 µm, 200 µm×0.5 mm, 120 Å) and washed for 5 min at a flow rate of 3 µL/min using 0.1% formic acid and acetonitrile (1:1, v/v). The samples were loaded onto a ChromXP C18-CL analytical column (3 µm, 75 µm×15 cm, 120 Å). The flow rate was 300 nL/min. Elution gradient started from 5 to 35% acetonitrile in 0.1% formic acid in 90 min. The percentage of acetonitrile was increased to 80% in 5 min and kept for an additional 5 min. The final percentage of acetonitrile was reduced to 5% from 95 to 100 min and then equilibrated for 20 min for the next sample. The samples were introduced to TripleTOF 5600 mass spectrometer (AB SCIEX, Framingham, MA) via the nanoflex link. The spray tip was the PicoTip emitter (360 µm OD, 20 µm ID, 12 cm long with 10 µm tip opening size) from New Objective. GS1 and curtain gas was set at 3, 24 units respectively. Ionization voltage and interface heater temperature was set at 2400 V, 150°C respectively. MS precursor ion selection window was 400 to 1250 amu. Cycle time was 0.25 sec and 40 most abundant parent ions were passed to Q2 for product ion scan. The MS/MS spectra were acquired at 0.1 second at a m/z range from 100 to 2000. Protein Pilot version 4.5 (AB Sciex, Framingham, MA) was used for peak generation and database search. The organism was set as bacteria and the database was NCBI non-redundant database with 10% false discovery rate. Peptide identifications were accepted if they could be established at more than 95.0% probability. Proteins with two peptides matched from the database search with 95% confidence were kept. The experiments

and data analysis were performed at mass spectrometry proteomics core facilities at Baylor College of Medicine.

5.3.5. Gene synthesis, cloning, and transformation

To confirm the identity of enzymes discovered from LC-MS/MS analysis, gene synthesis, molecular cloning, and expression of the enzyme candidates were performed as shown in Figure 28. The native gene employs tandem rare codons that can reduce the efficiency of translation or even disengage the translational machinery. Therefore, genes for glycosidase [*Thermotoga sp*. RQ2] (gi: 501268188), glycosidase PH1107-related [*Thermotoga petrophila* RKU-1] (gi: 147736211), were subject to codon optimization for expression in *Escherichia coli* from GenScript (Piscataway, NJ). The codon usage bias in *Escherichia coli* was enhanced by upgrading the Codon Adaptation Index (CAI). GC content and unfavorable peaks were optimized to prolong the half-life of the mRNA. The Stem-Loop structures, which impact ribosomal binding and stability of mRNA, were broken. In addition, the artificial genes were subject to screening and modification of negative cis-acting sites. The codon-optimized genes were synthesized, and inserted into the expression vector pET-3a at its 5' NdeI and 3' BamHI restriction sites by Genscript (Piscataway, NJ).

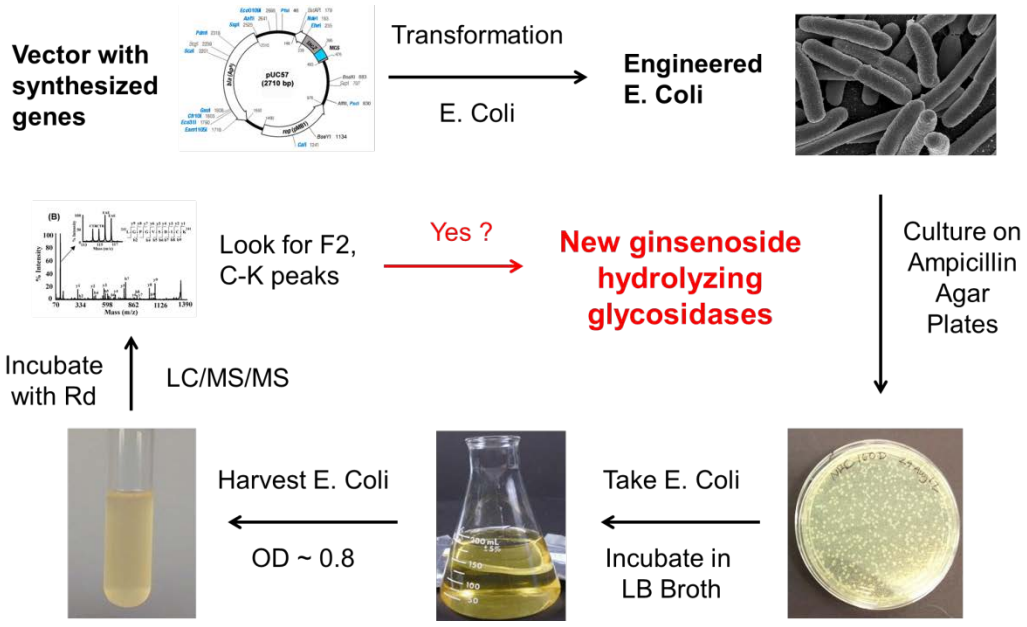


Figure 28. Study design of testing the functional activity of the potential ginsenoside hydrolyzing enzymes from A/J mouse feces

The construct, glycosidase PH1107-related/pET-3a was taken as the template and mutagenesis (insert deletion) was performed to prepare the blank pET-3a vector (mutation sequence CATATGGGATCC). The pET-3a vector with synthesized genes or the blank pET-3a vector was transformed into *Escherichia coli* DH5α by heat-shock transformation, respectively (Figure 29). Briefly, one shot of top10 competent cells (LifeTech, Grand Island, NY) was taken from -80°C freezers and thawed on ice (approximately 20-30 min). agar plates (containing the appropriate antibiotic) out of 4°C to warm up to room temperature or place in 37°C incubator. The pET-3a vector (100 ng DNA) was transferred into 50 µL of competent cells in microcentrifuge tubes and gently mixed by flicking the bottom of the tube with fingers for a few times. The competent cell/DNA mixture was placed on ice for 20-30 min. Heat-shock transformation was performed by placing the bottom half of the tubes into a 42°C water bath for 1 min. The tubes were put back on ice for 2 min followed by addition of 250-500 µL LB medium with 50 µg/mL ampicillin. The tubes incubated at 37°C in a shaking incubator for 45 min. The engineered *Escherichia coli* was then cultured on agar plates (Hardy Diagnostics, Santa Maria, CA) for overnight at 37°C prior to the experiment. The engineered *Escherichia coli* from Ampicillin-resistant colonies was picked up and inoculated into pre-autoclaved LB medium (2.5 % LB Broth in H_2O, w/v, pH=7). The inoculum was incubated in a shaking water bath at 37 °C. The bacteria was taken for the functional activity assay when the optical density (OD) value of the bacteria suspension was determined to be 0.8 based on its absorbance at 600 nm by Synergy 2 Multi-Mode Reader (BioTek, Winooski, VT).

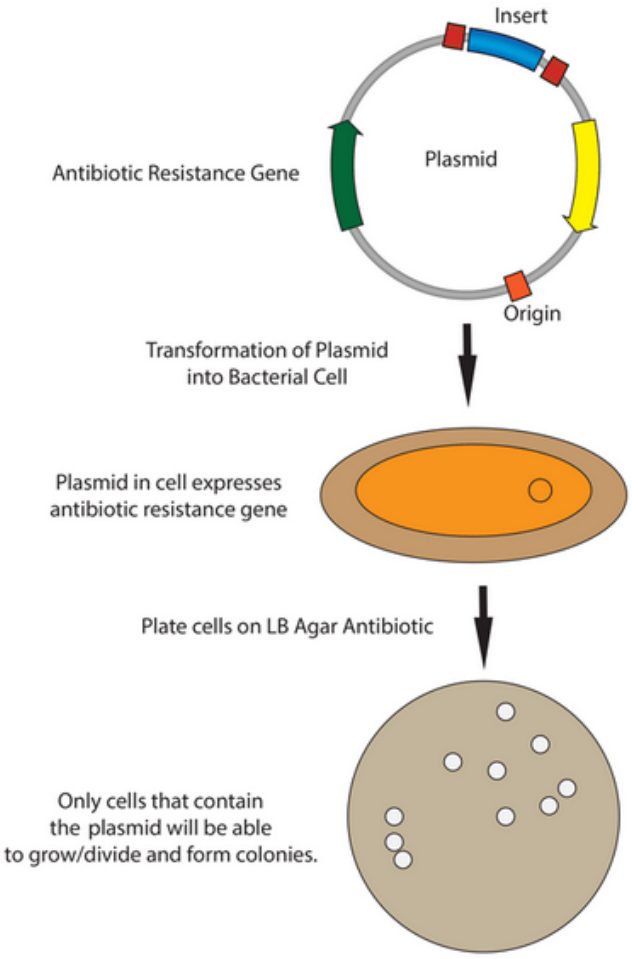


Figure 29. Scheme of transformation of pET-3a and culture on Agar plates

5.3.6. Enzymatic function assays

5.3.6.1 Colorimetric assay

A colorimetric assay was used to determine the enzymatic activity of fractions and gel pieces in section 5.3.2 and 5.3.3. A reaction mixture containing 20 µL 25 mM p-nitrophenyl-β-D-pyranoside, 40 µL PBS (pH=3.7), 200 µL enzyme fractions was incubated at 45 ℃ and 120 rpm in an orbital shaker from Thermo Scientific (Asheville, NC) for 60 min. The reaction was stopped with the addition of 20 µL 1 M sodium carbonate. To test the activity of gel pieces, 1 mL PBS (pH=3.7) and 150 µL p-nitrophenyl-β-D-pyranoside were used. The reaction was stopped by adding 100 µL 1 M sodium carbonate. The absorbance (405 nm) of the reaction mixture was determined by Synergy 2 Multi-Mode Reader (BioTek, Winooski, VT). Specific activity was defined as units per mg protein per min (Table 5). Purification fold was the ratio of specific activity for a particular enrichment step over the S9 preparation step. Yield was calculated by taking the ratio of the total activity of a particular enrichment step over the S9 preparation step.

5.3.6.2 Ginsenoside hydrolysis assay

Ginsenoside Rd was used as the substrate to confirm the ginsenoside hydrolyzing activity of enzyme fractions and gel pieces. A reaction mixture containing 10 µM ginsenoside Rd, 1 mL enzyme fraction, and 1 mL PBS (pH=3.7) was incubated at 45°C for up to 96 hr. The samples (100 µL) were collected, processed and analyzed following the same procedures in section 3.3.5.1.

Functional activity of the engineered bacteria and blank pET-3a vector was determined by examining the formation of ginsenoside F2 and Compound K. Engineered *Escherichia coli* (2 mL) was transferred to disposable glass vials from VWR (Houston, TX). Ginsenoside Rd was added to the glass vials to a final concentration of 5 μM. The mixture was incubated at 37°C and 120 rpm in an orbital shaker from Thermo Scientific (Asheville, NC) for up to 72 hr. The remaining experimental procedures were the same as in Section 4.3.4.

5.3.7. Statistical analysis

The data in this study are presented as mean ± SD, if not specified otherwise. For the ginsenoside hydrolysis assay, significance is assessed by Student's-test, one way ANOVA with and without Tukey's post hoc test. A p-value of < 0.05 was considered statistically significant.

5.4. Results

5.4.1. Enrichment of ginsenoside hydrolyzing enzymes

Ginsenoside hydrolyzing enzymes were enriched from A/J mouse feces by a classic chromatographic approach for the purpose of identifying some of the glycosidases involved in the ginsenoside deglycosylation. The crude enzyme fractions (S9 fractions) were prepared from the A/J mouse feces and precipitated with saturated ammonium sulfate to remove lipids and cell debris. The S9 fractions were then subject to removal of small molecular impurities by dialysis. The enzyme fractions were loaded onto an anion

exchange chromatographic column. The UV chromatogram of the elution profile (280 nm) was shown in Figure 30A.

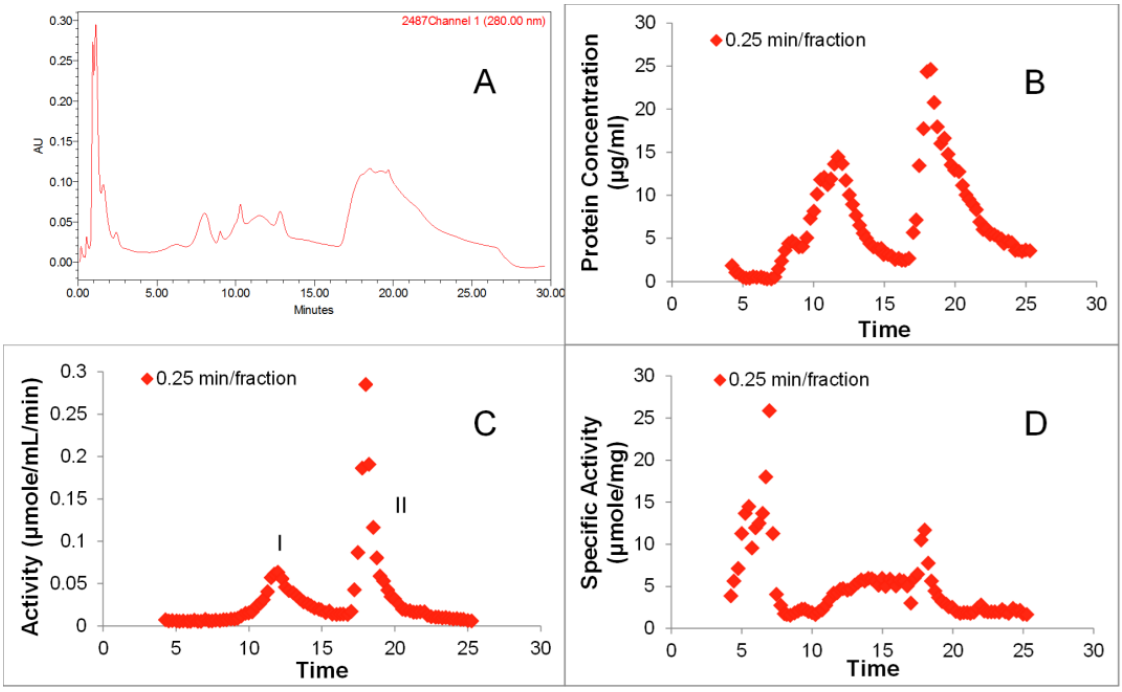


Figure 30. UV spectrum (A), protein concentration (B), activity (C) and specific activity (D) of enzyme fractions in ion exchange chromatographya

Incubation was carried out at 45℃ for 45 min using p-nitrophenyl-β-D-pyranoside as the probe substrate.

There were two major peaks identified, suggested by the protein concentration profile in Figure 30B. Both of these peaks contained enzyme fractions that were active against p-nitrophenyl-β-D-pyranoside (Figure 30C), although the maximum activity of peak II was more than 5 fold of peak I. However, the maximum specific activity of peak II was only half of peak I when normalized by protein concentration of each fraction (Figure 30D). Activities of both peaks were further examined by the ginsenoside hydrolysis assay. Peak II, but not peak I, produced ginsenoside F2 and Compound K, suggesting that only peak II had ginsenoside hydrolyzing enzymes (Figure 31). Therefore, fractions (17.5-19 min) from peak II were pooled and dialyzed to remove high concentration of sodium chloride. These fractions were concentrated by centrifugation in protein concentrators and then loaded onto a gel filtration column aiming at removal of high and low molecular weight contaminants. The UV chromatogram (280 nm) and protein concentration of the enzyme fractions was shown in Figure 32A and B, respectively.

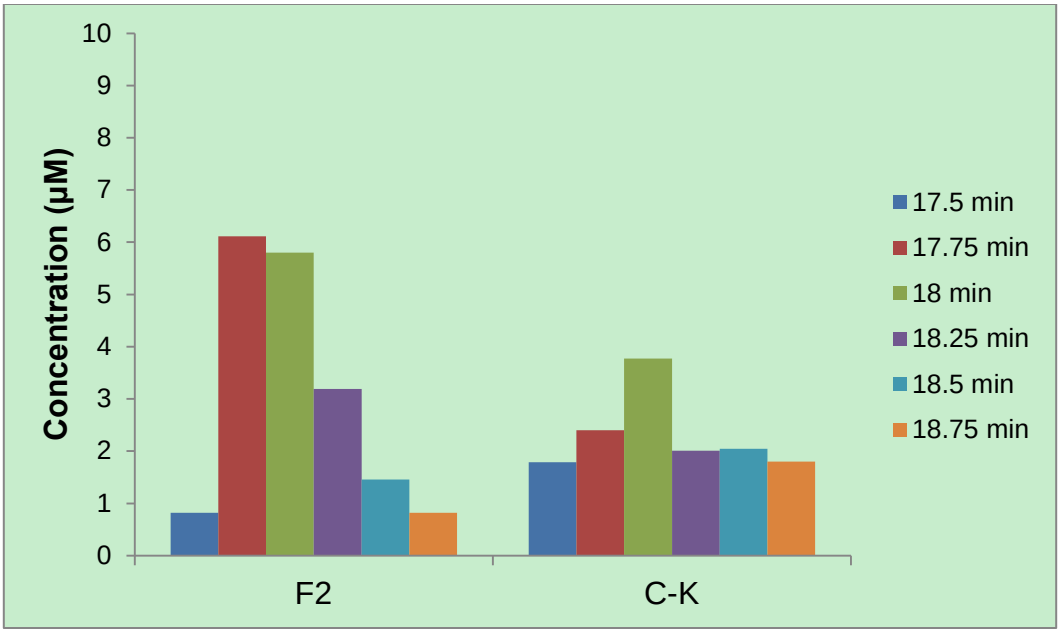


Figure 31. Activity of enzyme fractions in ion exchange chromatography, represented by metabolites formed after incubation

Incubation was carried out at 45℃ for 48 h using 10 μM ginsenoside Rd as the probe substrate. The protein concentration was 0.013-0.06 mg/ml. Fractions were labeled by the corresponding reaction time (15 sec/fraction).

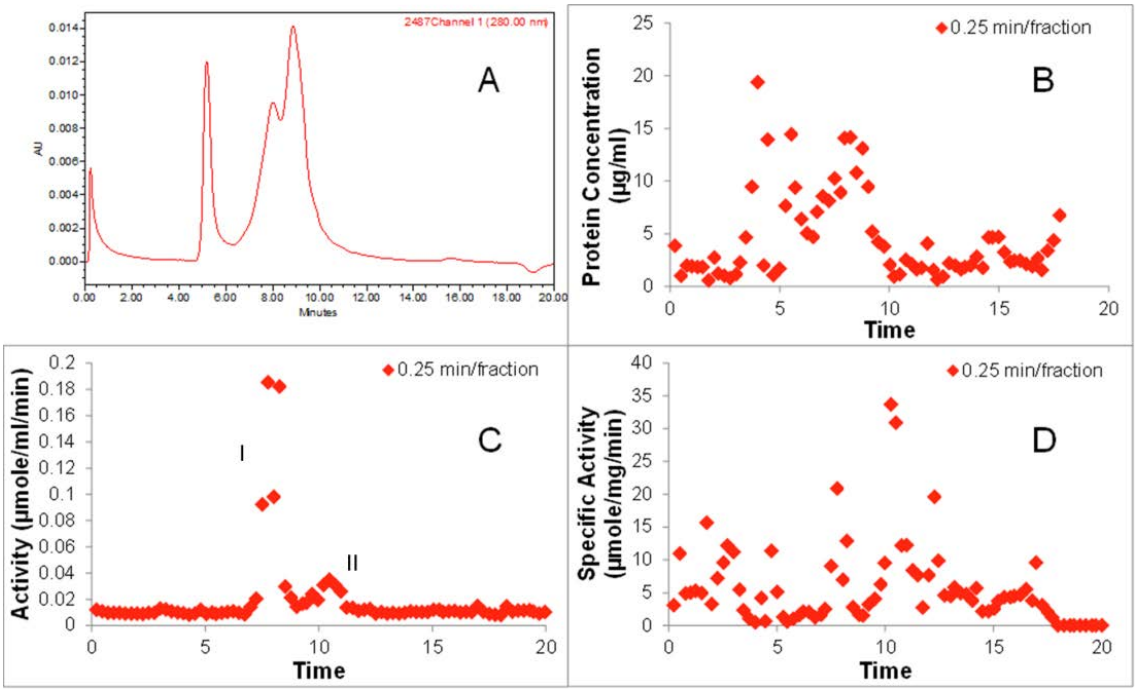


Figure 32. UV spectrum (A), protein concentration (B), activity (C) and specific activity (D) of enzyme fractions in gel filtration chromatography

Incubation was carried out at 45℃ for 45 min using p-nitrophenyl-β-D-pyranoside as the probe substrate.

Similar to ion exchange chromatography, two major peaks were identified. Enzyme fractions from both peaks I and II were active against p-nitrophenyl-β-D-pyranoside (Figure 32C). However, the maximum activities of these two peaks showed a 6-fold difference. In contrast to ion exchange chromatography, both peak I and II carried ginsenoside hydrolyzing enzymes as indicated by the formation of ginsenoside F2 and Compound K (Figure 33). Consequently, fractions (7.5-8.5 min and 10.25-10.75 min) from peak I and II were pooled, concentrated and stored on ice for further analysis. Summary of all the enrichment steps is listed in Table 5.

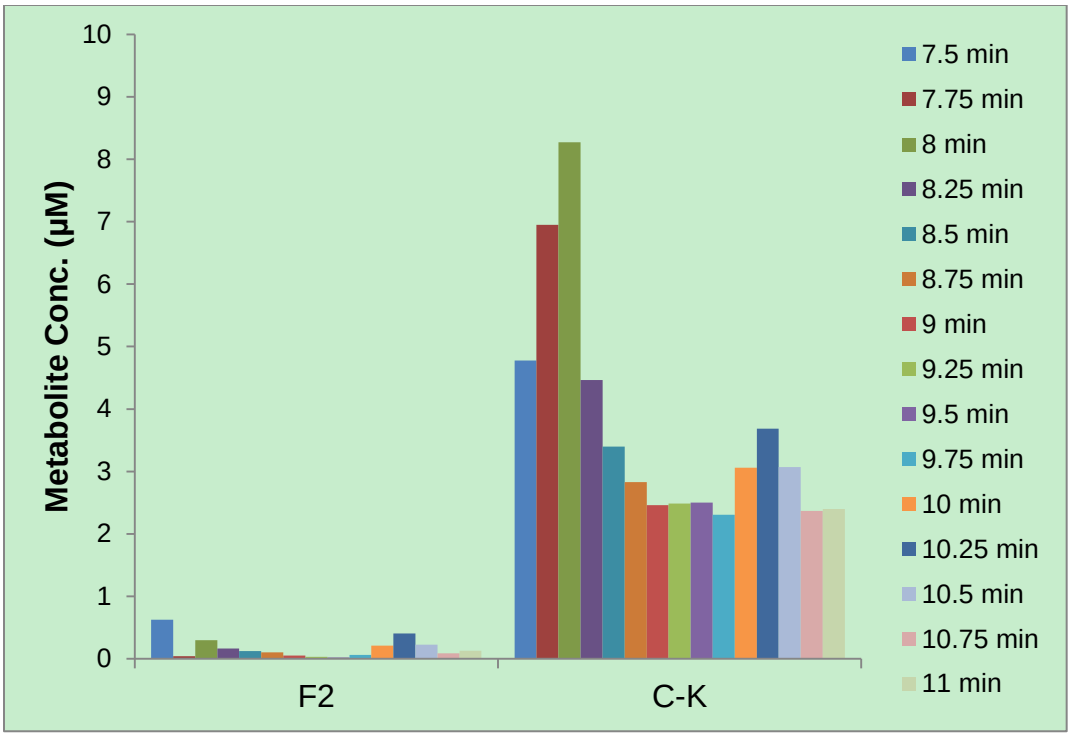


Figure 33. Activity of enzyme fractions in ion exchange chromatography, represented by metabolites formed after incubation

Incubation was carried out at 45°C for 96 h using 10 μM ginsenoside Rd as the probe substrate. The protein concentration was 0.06-9.5 μg/mL. Fractions were labeled by the corresponding reaction time (15 sec/fraction).

Table 5. Summary of enrichment of ginsenoside hydrolyzing enzymes

Step	Total Activity (U)	Specific Activity (μmol/mg/min)	Total Protein (mg)	Purification Fold	Yield (%)
S9 preparation	75.74	0.757	100.05	1.00	100.00
Protein precipitation	56.67	1.969	28.78	2.60	74.79
Dialysis	31.39	3.633	8.64	4.80	41.43
IEX	21.03	11.684	1.80	15.43	27.72
Dialysis/concentration	6.67	11.118	0.60	14.68	8.85
GFC/concentration	4.40	27.50	0.16	36.33	5.81

One unit of enzyme activity was defined as the amount of enzyme releasing 1 μmol of p-nitrophenol per min under standard conditions (pH 4.5, 45°C).

5.4.2. Gel electrophoresis

Native gel electrophoresis was performed to further enrich the ginsenoside hydrolyzing enzymes. Relative activity of each excised gel pieces was shown in Figure 34A. Gel pieces with molecular weight estimated approximately 100 kD and 37 kD (band# 2, 3, 8, 9) were active against p-nitrophenyl-β-D-pyranoside. Additionally, gel pieces (band #2, 3, 8, 9) showed protein bands at the same locations in the Coomassie Brilliant Blue stained lane in accordance with the activity. Proteins in these gel pieces were recovered and subject to SDS-PAGE (Figure 34B). There were quite a few proteins identified by SDS-PAGE. Molecular weights of the major proteins were approximately 25-37 kD, 37-50 kD and 75-100 kD. These gel pieces were cut for LC-MS/MS analysis.

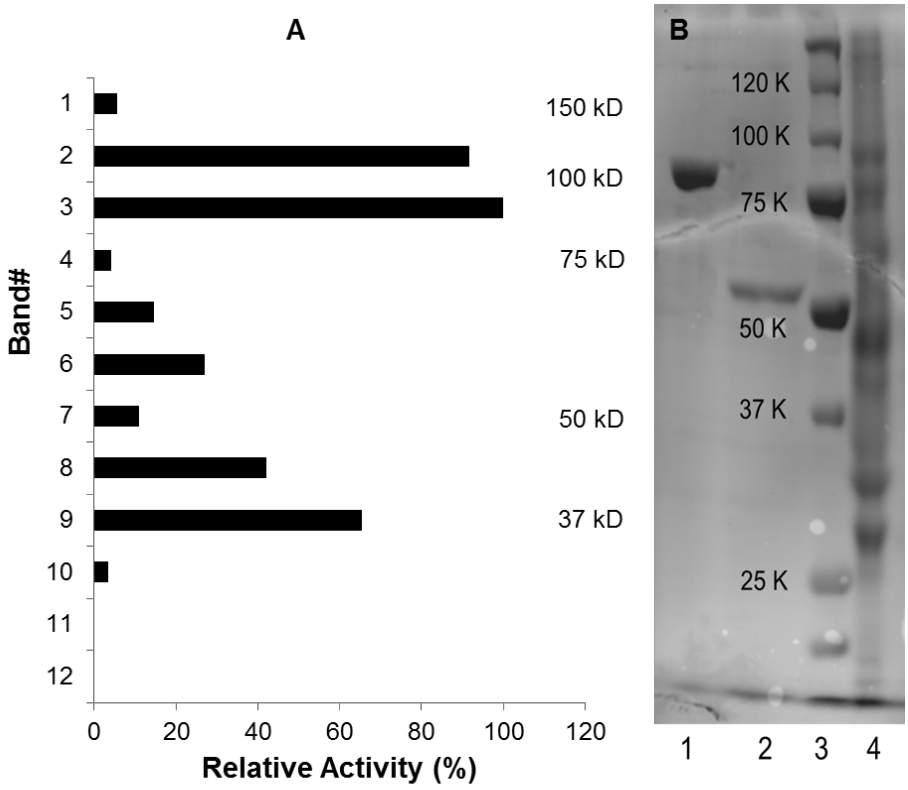


Figure 34. A: Activity of excised gel pieces in native PAGEa, B: SDS-PAGE of enzyme fractions containing the potential ginsenoside hydrolyzing enzymesb.

Activity of each band was normalized by the 3rd band. The position of molecular weight markers were aligned accordingly. Panel B, Lane 1, β-glucosidase, Bacteroides fragilis NCTC 9343 (MW=95.2 kDa); lane 2, and β-1,4-Glucosidase 1A (GH1), Clostridium thermocellum (MW=52.7 kDa); lane 3, molecular weight markers; lane 4, enzyme fractions.

5.4.3. Identification of potential ginsenoside hydrolyzing enzymes by LC-MS/MS

The identity of the ginsenoside hydrolyzing enzymes was partially revealed by analyzing tryptic peptide sequences by nanoLC-ESI-QTOF (Table 6). Only one unique peptide NGVLFPR was identified, covering 2% (7/326 or 327) of the protein sequence (Table 7). No post translational modification of the identified proteins was found. There was no bacteria glycosidase with at least two identified peptides. On the top of the identified protein list were bacteria membrane proteins, such as, OmpA family lipoprotein (22 unique peptides, 68% coverage, gi: 388000759) from *Pseudomonas fluorescens* SS101. The results are probably due to the fact that these proteins could not be differentiated from other interfering proteins in the feces.

Table 6. Characteristics of spectra of unique tryptic peptides of potential ginsenoside hydrolyzing enzymes identified by nanoLC-ESI-QTOF with 95% probability

Peptide	Unique	-10lgP	Mass	ppm	Z	PTM
NGVLFPR	Y	22.91	801.4497	-0.2	2	N
NGVLFPR	Y	22.91	801.4497	-0.2	2	N
NGVLFPR	Y	22.91	801.4497	-0.2	2	N
NGVLFPR	Y	22.91	801.4497	-0.2	2	N
NGVLFPR	Y	22.91	801.4497	-0.2	2	N
NGVLFPR	Y	22.91	801.4497	-0.2	2	N

Table 7. Characteristics of potential ginsenoside hydrolyzing enzymes identified by nanoLC-ESI-QTOF

Accession	Mass	AA	Description	Bacteria
gi|49259571	38994	326	Chain A Crystal Structure Of A Predicted Glycosidase	Thermotoga maritima Msb8
gi|501268188	37237	326	glycosidase	Thermotoga sp. RQ2
gi|147736211	37258	326	glycosidase PH1107-related	Thermotoga petrophila RKU-1
gi|15643981	37258	326	hypothetical protein TM1225	Thermotoga maritima MSB8
gi|167696782	37139	327	hypothetical protein BACSTE_03540	Bacteroides stercoris ATCC 43183

5.4.4. Optimized gene sequences of potential ginsenoside hydrolyzing enzymes

The codon optimization adjusted the GC content (average GC content 51.63%, 51.91% for glycosidase, and glycosidase PH1107-related respectively, Appendix D) to prolong mRNA half life. Codon usage was adapted to the bias of *Escherichia coli* resulting in a Codon Adaptation Index (CAI) value of 0.90, 0.90 for glycosidase, and glycosidase PH1107-related, respectively (Appendix D). The optimized genes should therefore allow high and stable expression rates in *Escherichia coli*. The optimized sequences were shown below.

glycosidase PH1107-related [Thermotoga petrophila RKU-1]

NdeI: a restriction endonuclease from Neisseria denitrificans which cleaves DNA at 5'-CA/TATG-3' sequences; BamHI: a restriction endonuclease from Bacillus amyloliquefaciens H (ATCC 49763) which cleaves DNA at 5'-G/GATCC-3' sequences; DNA codons of corresponding amino acids were determined in accordance with Appendix E.

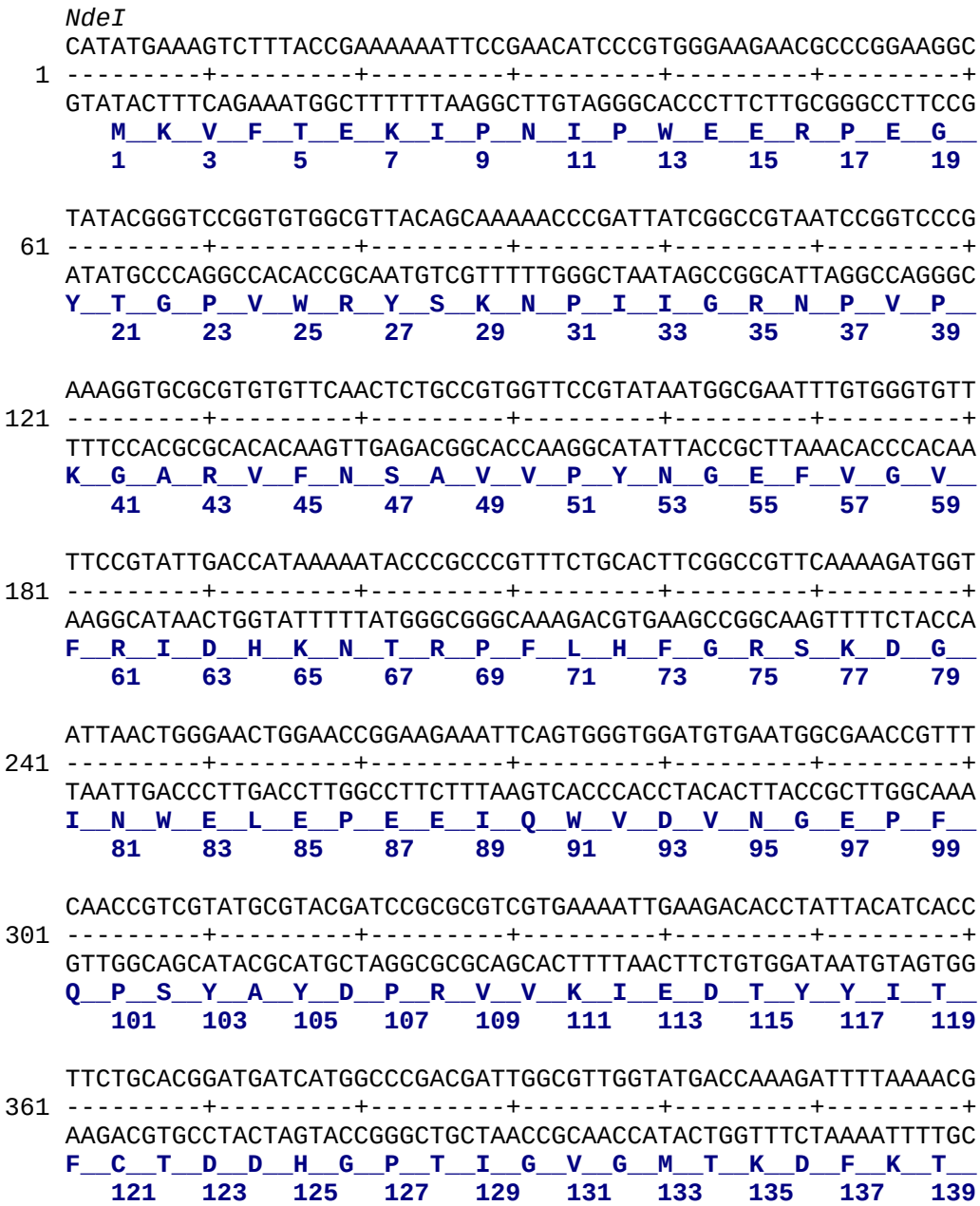
```
    NdeI
    CATATGAAAGTTATGGGCGAACGCATTCCGAACATCCCGTGGGAAGATCGTCCGGAAGGC
  1 ---------+---------+---------+---------+---------+---------+
    GTATACTTTCAATACCCGCTTGCGTAAGGCTTGTAGGGCACCCTTCTAGCAGGCCTTCCG
       M__K__V__M__G__E__R__I__P__N__I__P__W__E__D__R__P__E__G__
       1     3     5     7     9     11    13    15    17    19

    TATACCGGTCCGGTGTGGCGCTACTCAAAAAACCCGATTATCGGCCGTAATCCGGTGCCG
 61 ---------+---------+---------+---------+---------+---------+
    ATATGGCCAGGCCACACCGCGATGAGTTTTTTGGGCTAATAGCCGGCATTAGGCCACGGC
    Y__T__G__P__V__W__R__Y__S__K__N__P__I__I__G__R__N__P__V__P__
       21    23    25    27    29    31    33    35    37    39

    AAAGGTGCGCGTGTTTTTAACTCGGCCGTGGTTCCGTATAATGGCGAATTTGTGGGTGTT
121 ---------+---------+---------+---------+---------+---------+
    TTTCCACGCGCACAAAAATTGAGCCGGCACCAAGGCATATTACCGCTTAAACACCCACAA
    K__G__A__R__V__F__N__S__A__V__V__P__Y__N__G__E__F__V__G__V__
       41    43    45    47    49    51    53    55    57    59

    TTCCGTATTGACCATAAAAATACCCGCCCGTTTCTGCACTTCGGCCGTTCCGAAGATGGT
181 ---------+---------+---------+---------+---------+---------+
    AAGGCATAACTGGTATTTTTATGGGCGGGCAAAGACGTGAAGCCGGCAAGGCTTCTACCA
    F__R__I__D__H__K__N__T__R__P__F__L__H__F__G__R__S__E__D__G__
       61    63    65    67    69    71    73    75    77    79

    ATTCATTGGGAAATCGAACCGGAAGAAATCCAGTGGGTTGACCTGAACGGCAATCCGTTT
241 ---------+---------+---------+---------+---------+---------+
    TAAGTAACCCTTTAGCTTGGCCTTCTTTAGGTCACCCAACTGGACTTGCCGTTAGGCAAA
    I__H__W__E__I__E__P__E__E__I__Q__W__V__D__L__N__G__N__P__F__
       81    83    85    87    89    91    93    95    97    99

    CAACCGAGCTATGCGTACGATCCGCGTGTCGTGAAAATTGAAGACACCTATTACATCACC
301 ---------+---------+---------+---------+---------+---------+
    GTTGGCTCGATACGCATGCTAGGCGCACAGCACTTTTAACTTCTGTGGATAATGTAGTGG
    Q__P__S__Y__A__Y__D__P__R__V__V__K__I__E__D__T__Y__Y__I__T__
       101   103   105   107   109   111   113   115   117   119

    TTCTGCACGGATGACCACGGTCCGACGATTGGCGTGGGTATGACCAAAGATTTTAAAACG
361 ---------+---------+---------+---------+---------+---------+
    AAGACGTGCCTACTGGTGCCAGGCTGCTAACCGCACCCATACTGGTTTCTAAAATTTTGC
    F__C__T__D__D__H__G__P__T__I__G__V__G__M__T__K__D__F__K__T__
       121   123   125   127   129   131   133   135   137   139
```

```
   TTCGTCCGCCTGCCGAACGCCTATGTGCCGTTTAACCGCAATGGCGTTCTGTTCCCGCGT
421 ---------+---------+---------+---------+---------+---------+
   AAGCAGGCGGACGGCTTGCGGATACACGGCAAATTGGCGTTACCGCAAGACAAGGGCGCA
   F__V__R__L__P__N__A__Y__V__P__F__N__R__N__G__V__L__F__P__R__
      141   143   145   147   149   151   153   155   157   159

   AAAATCAAGGGTAAATACGTTATGCTGAACCGTCCGTCTGATAATGGCCATACCCCGTTT
481 ---------+---------+---------+---------+---------+---------+
   TTTTAGTTCCCATTTATGCAATACGACTTGGCAGGCAGACTATTACCGGTATGGGGCAAA
   K__I__K__G__K__Y__V__M__L__N__R__P__S__D__N__G__H__T__P__F__
      161   163   165   167   169   171   173   175   177   179

   GGTGACATTTTCCTGAGCGAATCTCCGGATATGATCCATTGGGGTAATCACCGTTTTGTG
541 ---------+---------+---------+---------+---------+---------+
   CCACTGTAAAAGGACTCGCTTAGAGGCCTATACTAGGTAACCCCATTAGTGGCAAAACAC
   G__D__I__F__L__S__E__S__P__D__M__I__H__W__G__N__H__R__F__V__
      181   183   185   187   189   191   193   195   197   199

   ATGGGCCGCAGCGGTTATAACTGGTGGGAAAATCTGAAAATTGGCGCGGGTCCGTATCCG
601 ---------+---------+---------+---------+---------+---------+
   TACCCGGCGTCGCCAATATTGACCACCCTTTTAGACTTTTAACCGCGCCCAGGCATAGGC
   M__G__R__S__G__Y__N__W__W__E__N__L__K__I__G__A__G__P__Y__P__
      201   203   205   207   209   211   213   215   217   219

   ATCGAAACGTCTGAAGGCTGGCTGCTGATTTACCACGGTGTCACCCTGACGTGTAACGGC
661 ---------+---------+---------+---------+---------+---------+
   TAGCTTTGCAGACTTCCGACCGACGACTAAATGGTGCCACAGTGGGACTGCACATTGCCG
   I__E__T__S__E__G__W__L__L__I__Y__H__G__V__T__L__T__C__N__G__
      221   223   225   227   229   231   233   235   237   239

   TATGTGTACAGTTTTGGTGCGGCCCTGCTGGATCTGGATGACCCGAGTAAAGTGCTGTAT
721 ---------+---------+---------+---------+---------+---------+
   ATACACATGTCAAAACCACGCCGGGACGACCTAGACCTACTGGGCTCATTTCACGACATA
   Y__V__Y__S__F__G__A__A__L__L__D__L__D__D__P__S__K__V__L__Y__
      241   243   245   247   249   251   253   255   257   259

   CGTTCCCGCTATTACCTGCTGACCCCGGAAGAAGAATACGAAACGGTCGGCTTTGTGCCG
781 ---------+---------+---------+---------+---------+---------+
   GCAAGGGCGATAATGGACGACTGGGGCCTTCTTCTTATGCTTTGCCAGCCGAAACACGGC
   R__S__R__Y__Y__L__L__T__P__E__E__E__Y__E__T__V__G__F__V__P__
      261   263   265   267   269   271   273   275   277   279

   AATGTTGTCTTCCCGTGCGCAGCTCTGTGTGATGCAGACACCGGCCGCGTTGCTATTTAT
841 ---------+---------+---------+---------+---------+---------+
   TTACAACAGAAGGGCACGCGTCGAGACACACTACGTCTGTGGCCGGCGCAACGATAAATA
   N__V__V__F__P__C__A__A__L__C__D__A__D__T__G__R__V__A__I__Y__
      281   283   285   287   289   291   293   295   297   299

   TACGGTGCGGCCGATACGCATGTCGCACTGGCTTTCGGTTACATCGATGAAATCGTTGAC
901 ---------+---------+---------+---------+---------+---------+
```

```
    ATGCCACGCCGGCTATGCGTACAGCGTGACCGAAAGCCAATGTAGCTACTTTAGCAACTG
    Y__G__A__A__D__T__H__V__A__L__A__F__G__Y__I__D__E__I__V__D__
       301   303   305   307   309   311   313   315   317   319

                          BamHI
    TTCGTCAAACGTAACAGCATGGGATCC
961 ---------+---------+-------
    AAGCAGTTTGCATTGTCGTACCCTAGG
    F__V__K__R__N__S__M__*__
       321   323   325
```

glycosidase [Thermotoga sp. RQ2]

NdeI: a restriction endonuclease from Neisseria denitrificans which cleaves DNA at 5'-CA/TATG-3' sequences; BamHI: a restriction endonuclease from Bacillus amyloliquefaciens H (ATCC 49763) which cleaves DNA at 5'-G/GATCC-3' sequences; DNA codons of corresponding amino acids were determined in accordance with Appendix E.



```
   TTCGTTCGCCTGCCGAACGCCTATGTCCCGTTTAACCGCAATGGCGTGCTGTTCCCGCGT
421 ---------+---------+---------+---------+---------+---------+
   AAGCAAGCGGACGGCTTGCGGATACAGGGCAAATTGGCGTTACCGCACGACAAGGGCGCA
   F__V__R__L__P__N__A__Y__V__P__F__N__R__N__G__V__L__F__P__R__
      141   143   145   147   149   151   153   155   157   159

   AAAATCAACGGTAAATACGTTATGCTGAACCGTCCGAGTGATAATGGCCACACCCCGTTT
481 ---------+---------+---------+---------+---------+---------+
   TTTTAGTTGCCATTTATGCAATACGACTTGGCAGGCTCACTATTACCGGTGTGGGGCAAA
   K__I__N__G__K__Y__V__M__L__N__R__P__S__D__N__G__H__T__P__F__
      161   163   165   167   169   171   173   175   177   179

   GGTGACATTTTCCTGAGTGAATCCCCGGATATGATCCATTGGGGCAACCACCGTTTTGTG
541 ---------+---------+---------+---------+---------+---------+
   CCACTGTAAAAGGACTCACTTAGGGGCCTATACTAGGTAACCCCGTTGGTGGCAAAACAC
   G__D__I__F__L__S__E__S__P__D__M__I__H__W__G__N__H__R__F__V__
      181   183   185   187   189   191   193   195   197   199

   CTGGGTCGCAGCTCTTATAACTGGTGGGAAAATCTGAAAATTGGCGCGGGTCCGTATCCG
601 ---------+---------+---------+---------+---------+---------+
   GACCCAGCGTCGAGAATATTGACCACCCTTTTAGACTTTTAACCGCGCCCAGGCATAGGC
   L__G__R__S__S__Y__N__W__W__E__N__L__K__I__G__A__G__P__Y__P__
      201   203   205   207   209   211   213   215   217   219

   ATCGAAACCAGTGAAGGCTGGCTGCTGATTTATCATGGCGTTACCCTGACGTGTAACGGC
661 ---------+---------+---------+---------+---------+---------+
   TAGCTTTGGTCACTTCCGACCGACGACTAAATAGTACCGCAATGGGACTGCACATTGCCG
   I__E__T__S__E__G__W__L__L__I__Y__H__G__V__T__L__T__C__N__G__
      221   223   225   227   229   231   233   235   237   239

   TATGTCTACTCCTTTGGTGCGGCCCTGCTGGATCTGGATGACCCGTCAAAAGTGCTGTAT
721 ---------+---------+---------+---------+---------+---------+
   ATACAGATGAGGAAACCACGCCGGGACGACCTAGACCTACTGGGCAGTTTTCACGACATA
   Y__V__Y__S__F__G__A__A__L__L__D__L__D__D__P__S__K__V__L__Y__
      241   243   245   247   249   251   253   255   257   259

   CGTTCGCGCTATTACCTGCTGACCCCGGAAGAAGAATACGAAACGGTTGGCTTTGTCCCG
781 ---------+---------+---------+---------+---------+---------+
   GCAAGCGCGATAATGGACGACTGGGGCCTTCTTCTTATGCTTTGCCAACCGAAACAGGGC
   R__S__R__Y__Y__L__L__T__P__E__E__E__Y__E__T__V__G__F__V__P__
      261   263   265   267   269   271   273   275   277   279

   AATGTTGTCTTCCCGTGCGCAGCTCTGTGTGATGCAGACACCGGCCGCGTGGCTATCTAT
841 ---------+---------+---------+---------+---------+---------+
   TTACAACAGAAGGGCACGCGTCGAGACACACTACGTCTGTGGCCGGCGCACCGATAGATA
   N__V__V__F__P__C__A__A__L__C__D__A__D__T__G__R__V__A__I__Y__
      281   283   285   287   289   291   293   295   297   299

   TACGGTGCGGCCGATACGCACGTTGCACTGGCTTTTGGTTATATTGATGAAATCGTGGAC
```

```
901 ---------+---------+---------+---------+---------+---------+
    ATGCCACGCCGGCTATGCGTGCAACGTGACCGAAAACCAATATAACTACTTTAGCACCTG
    Y__G__A__A__D__T__H__V__A__L__A__F__G__Y__I__D__E__I__V__D__
       301   303   305   307   309   311   313   315   317   319

                           BamHI
    TTCGTTAAACGTAATTCTATGGGATCC
961 ---------+---------+-------
    AAGCAATTTGCATTAAGATACCCTAGG
    F__V__K__R__N__S__M__*__
       321   323   325
```

5.4.5. Functional activity of potential ginsenoside hydrolyzing enzymes

Functional activities of the potential ginsenoside hydrolyzing enzymes were shown in Figure 35. The percentage of metabolites (ginsenoside F2 and Compound K) produced by the engineered *Escherichia coli* was approximately 10% after 72 hr incubation with ginsenoside Rd at 37°C. The chromagrams of ginsenoside Rd, F2 and Compound K during the course of incubation was shown in Figure 36 and Figure 37. Blank pET-3a vector showed no functional activity (Figure 38), suggesting that bacterial glycosidase PH1107-related (gi: 147736211) and bacterial glycosidase (gi: 501268188) were capable of hydrolyzing ginsenoside Rd to F2 and Compound K.

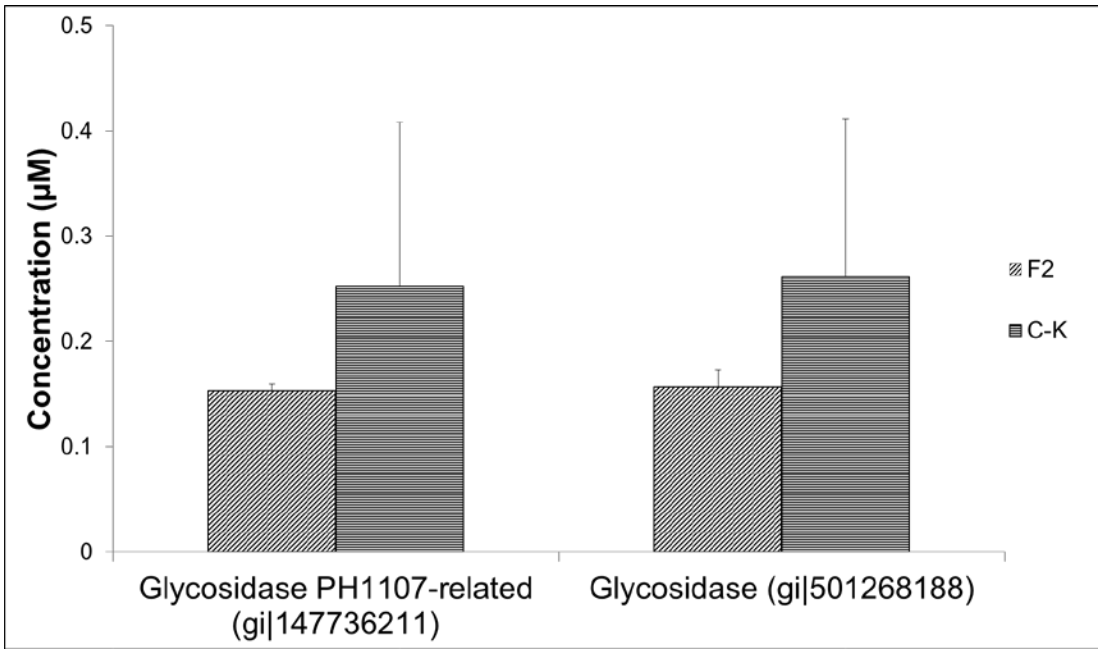


Figure 35. Functional activity of the two potential ginsenoside hydrolyzing enzymes

Incubation was carried out at 37℃ for 72 h using 5 μM ginsenoside Rd as the probe substrate.

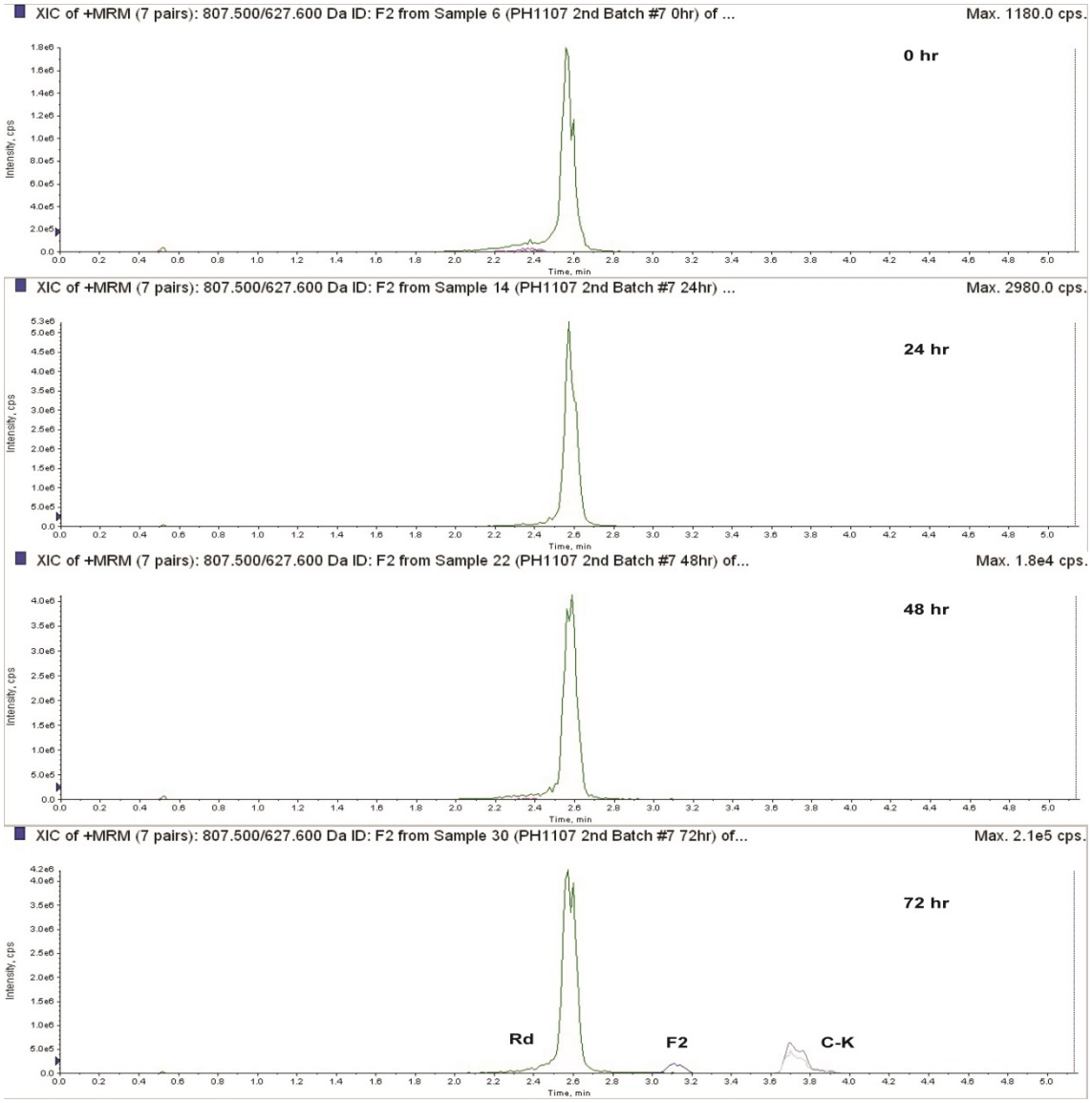


Figure 36. Chromatograms of metabolites of ginsenoside Rd by engineered *Escherichia coli* (glycosidase PH1107-related, gi: 147736211)

Incubation was carried out at 37℃ for up to 72 hr using ginsenside Rd (5 μM) as the probe substrate.

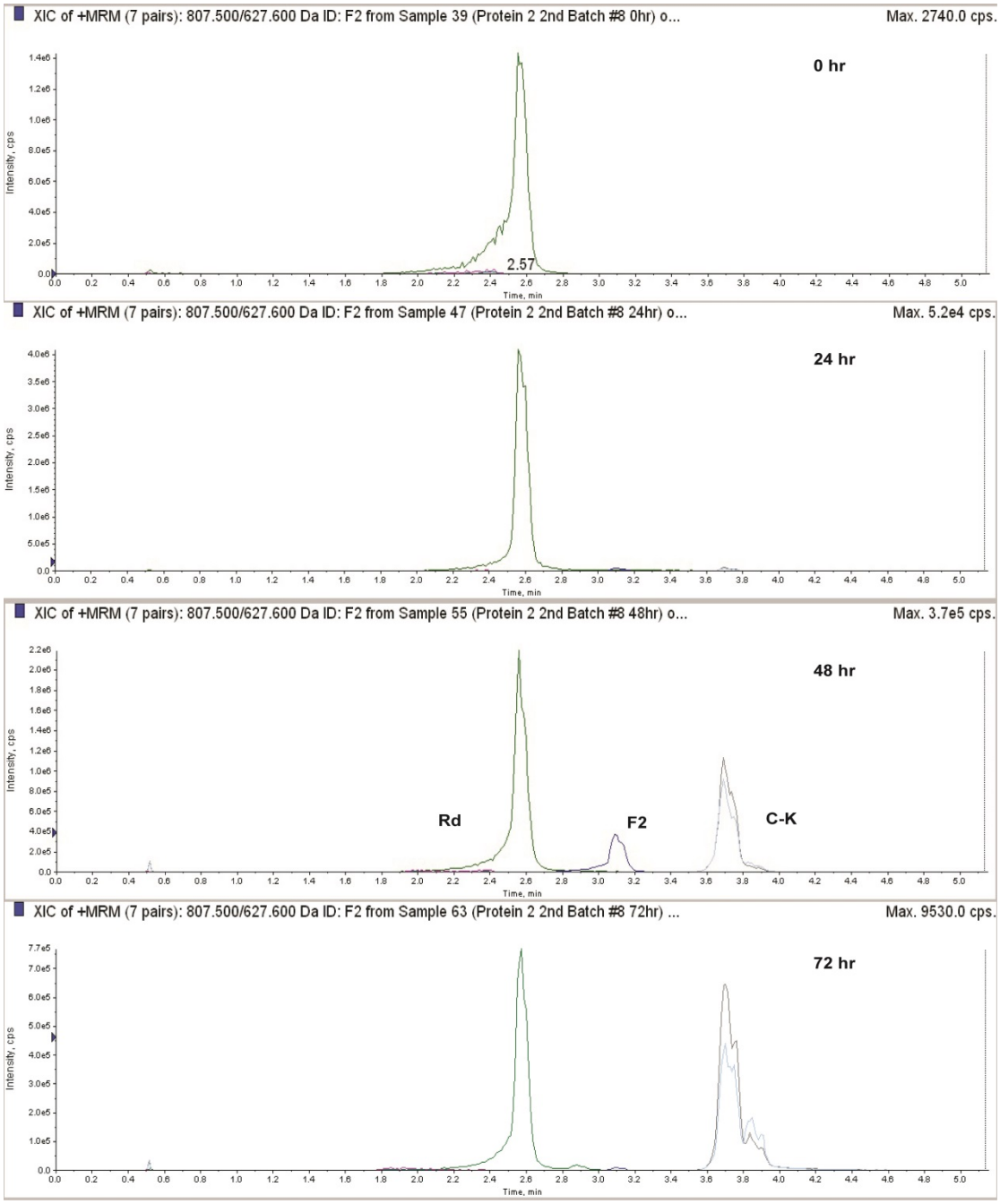


Figure 37. Chromatograms of metabolites of ginsenoside Rd produced by engineered *Escherichia coli* (glycosidase (gi: 501268188)

Incubation was carried out at 37℃ for up to 72 hr using ginsenside Rd (5 μM) as the probe substrate.

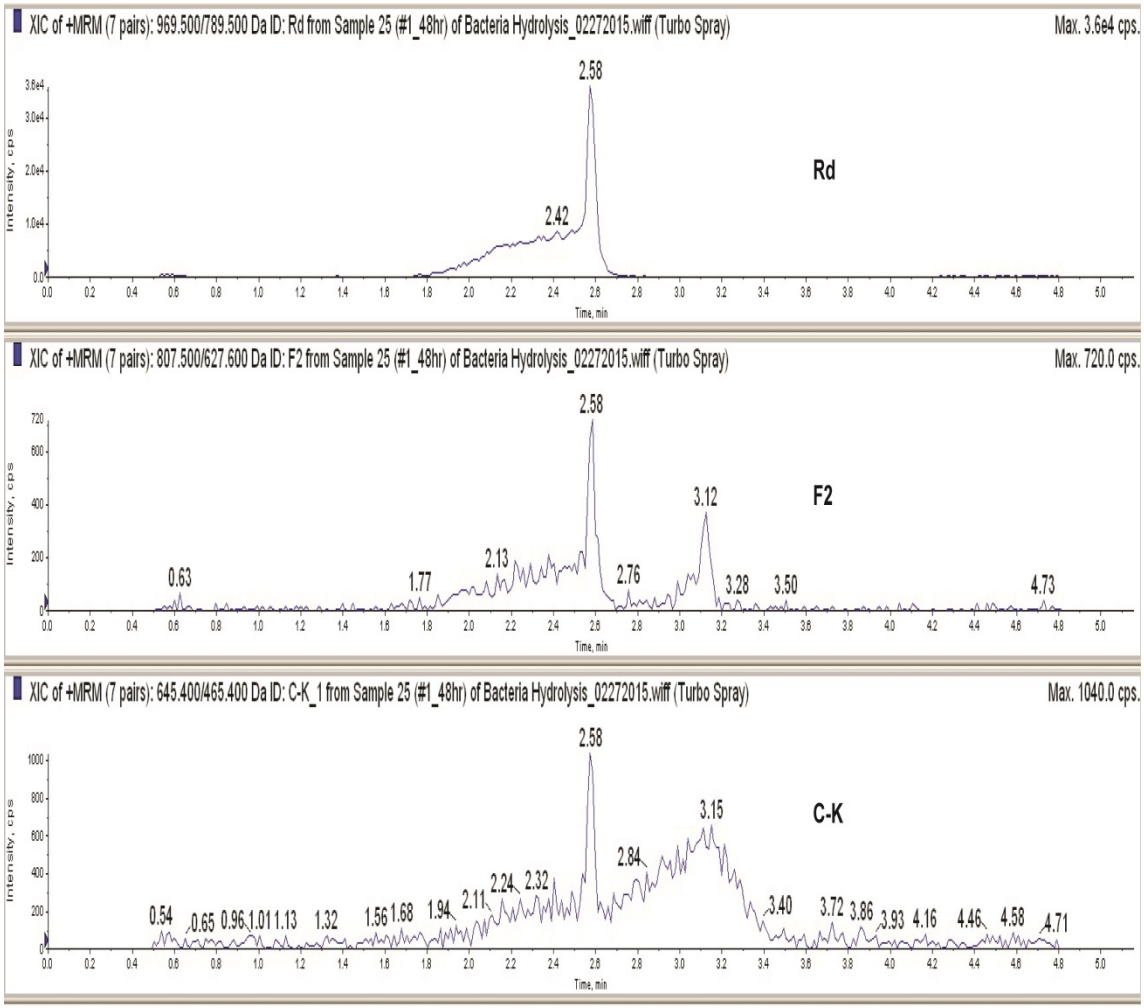


Figure 38. Chromatogram of metabolites of ginsenoside Rd produced by blank pET-3a vectora

Incubation was carried out at 37℃ for up to 48 hr using ginsenside Rd (1 μM) as the probe substrate. The 48 hr chromatogram was shown above.

5.5. Discussion

Researchers have been searching for bacterial glycosidases that would produce ginsenoside Compound K efficiently, mostly due to the difficulty in the synthesis of ginsenoside Compound K. Therefore, numerous efforts have been made to search specific bacteria species that express ginsenoside hydrolyzing glycosidases. Fungus, such as Paecilomyces Bainier sp. 229 (Yan, Zhou et al. 2008), Penicillium sclerotiorum (Wei, Zhao et al. 2011), and Aspergillus niger (Chang 2012), have been used as sources for purification of ginsenoside hydrolyzing enzymes. These fungi are mostly isolated from ginseng field soil. Cloning and expression of recombinant β-glucosidases from GRAS bacteria Bifidobacterium lactis AD011 (Kim, Wang et al. 2012) has also been reported. Additionally, a broad specificity β-glucosidase has been purified and characterized from sheep liver tissues (Chinchetru, Cabezas et al. 1989). To find bacteria present in the GI tract of A/J mouse and express ginsenoside hydrolyzing enzymes, A/J mouse feces were used to prepare the S9 fraction (crude enzyme) in this study. Using a classic protein chemistry approach, we identified two bacterial glycosidases responsible for RGE metabolism and the bacteria species expressing these enzymes. Our findings would help design specific probiotics and answer an important question, if chemoprevention efficacy of RGE could be altered through probiotic intervention using specific bacteria strains.

Dry feces are normally composed of undigested food remnants, solidified components of digestive juices, intestinal microbiota, fat, bile salts etc. There are large amount of interfering proteins, such as protease and bacterial membrane proteins, together with

bacterial glycosidases in the feces. These interfering proteins could not be eliminated to a large extent by the conventional ion exchange and gel filtration chromatography, suggested by the relatively low purification fold in our study (36 fold). Additionally, multiple protein bands, such as bacterial membrane proteins (i.e. phosphate ABC transporter substrate-binding protein, MW 34649; outer membrane porin OprF, MW 34534), were also found in gel electrophoresis. Ideally, immunopurificaiton would be the best approach for enrichment of targeted proteins from A/J mouse feces due to the specificity between bacterial glycosidases and the corresponding antibodies. Unfortunately, a brief literature search revealed that species reactivity of glycosidase antibodies on the market are for human and rodents but not for bacteria. In this case, antibody production might be a viable approach.

Generally, peptide identifications are accepted if they could be established at more than 95.0% probability, as specified by the Peptide Prophet algorithm (Keller 2002). Protein identifications are accepted if they could be established at more than 95.0% probability and contain at least two identified unique peptides. In our study, only one unique peptide (NGVLFPR) was identified at more than 95.0% probability, rendering the identified proteins questionable. Thus, gene synthesis, molecular cloning and transformation were performed as a confirmation test. The pET-3a vector (4640 bp) was selected because it carries NdeI (550) and BamHI (510) cloning sites. More importantly, there is no LacZα gene found in the map of pET-3a vector (Figure 39), making the blank pET-3a vector an approprioate negative control in the functional activity assay. LacZ gene has two fragments, LacZα and LacZΩ. The LacZα fragment encodes the N-terminus of β-

galactosidase, while the LacZΩ fragment, which is found in *E. coli* chromosome, encodes the C-terminus of the protein. Neither of these two fragments is active by itself. However, when both the fragments are present, a functional β-galactosidase is spontaneously reassembled.

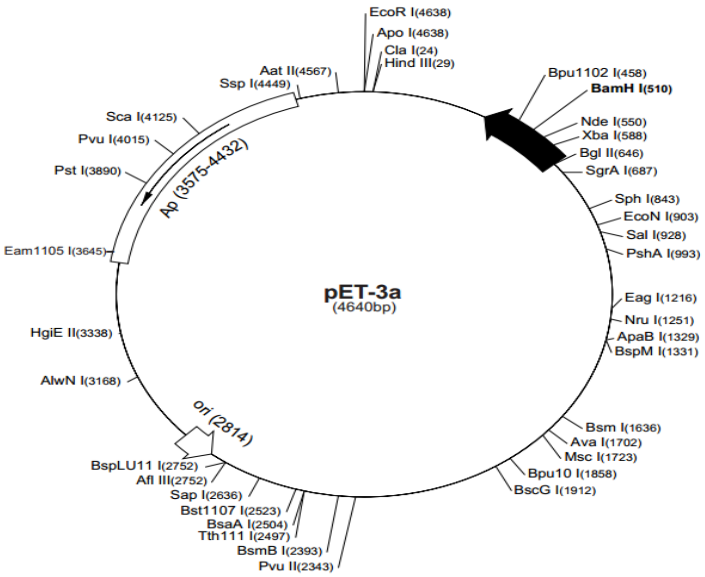


Figure 39. Map of bacterial expression pET-3a vector

5' Nde I(550) and 3' BamH I(510) were chosen as the cloning sites, (Source: http://www.helmholtz-muenchen.de/fileadmin/PEPF/pET_vectors/pET-3a-d_Map.pdf)

The functional activity study of the the engineered *Escherichia coli* showed that ginsenoside Rd was hydrolyzed to F2 and then Compound K in a stepwise manner, in accordance with our previous finding in Aim I. These bacterial glycosidases did not hydrolyze ginsenoside Rd in an efficient fashion, suggested by the percentage (10%) of formed metabolites (ginsenoside F2 and Compound K) after 72 hr incubation. By comparing the amino acid sequences with the NCBI Non-redundant protein sequences (NR) database using BLAST, the identified bacterial glycosidases belong to GH43 (glycoside hydrolase 43) family. GH43 family includes β-xylosidase (EC 3.2.1.37), α-L-arabinofuranosidase (EC 3.2.1.55), arabinanase (EC 3.2.1.99), xylanase (EC 3.2.1.8), galactan 1,3-β-galactosidase (EC 3.2.1.145), α-1,2-L-arabinofuranosidase (EC 3.2.1.-), exo-α-1,5-L-arabinofuranosidase (EC 3.2.1.-), exo-α-1,5-L-arabinanase (EC 3.2.1.-), β-1,3-xylosidase (EC 3.2.1.-) (http://www.cazy.org/GH43.html). Of the 7 enzyme classes listed above, β-xylosidase (EC 3.2.1.37), whose function is to remove successive D-xylose residues from the non-reducing termini, has been reported to possess exo-glucosidase activities (Chinchetru, Cabezas et al. 1989). Most of the purified ginsenoside hydrolyzing enzymes (Yan, Zhou et al. 2008; An, Cui et al. 2010; Kim, Wang et al. 2012; Quan, Wang et al. 2013) belong to β-glucosidase (EC. 3.2.1.21), a class of enymzes that hydrolyze terminal, non-reducing β-D-glucosyl residues with release of β-D-glucose. A common trait of β-glucosidase is its broad specificity. For example, a novel β-glucosidase in sheep liver has been reported to show β-D-galactosides, α-L-arabinosides, β-D-xylosides, and β-D-fucosides activities (Chinchetru, Cabezas et al. 1989). Therefore, we speculate that the bacterial glycosidases identified in this study belong to β-xylosidase (EC 3.2.1.37).

Determination of the DNA genome sequence of bacteria species, *Thermotoga maritima* Msb8, Thermotoga sp. RQ2, *Thermotoga petrophila* RKU-1, and *Bacteroides stercoris* ATCC 43183 have been done thanks to the Whole Genome Shortgun Project. The corresponding DNA genome sequences have been fed to computational analysis programs, such as GeneMark, Glimmer, tRNAscan-SE, RNAmmer for prediction of mRNA and amino acid sequences. Of the five identified bacterial glycosidases, three hypothetical proteins (chain A, crystal structure of a predicted glycosidase (tm1225), hypothetical protein TM1225, and hypothetical protein BACSTE_03540) were not cloned since there is no experimental evidence that they are expressed *in vivo*. Chain A, crystal structure of a predicted glycosidase (tm1225) (http://www.ncbi.nlm.nih.gov/protein/49259571) possesses a N-terminal His-Tag (MGSDKIHHHHHH) in addition to the amino acid sequence of hypothetical protein TM1225 (http://www.ncbi.nlm.nih.gov/protein/NP_229030.1). Additionally, chain A, crystal structure of a predicted glycosidase (tm1225) shares the same amino acid sequence as hypothetical protein TM1225 , other than having 5 wobble amino acids (X). A closer look of the data revealed that the recombinant *Escherichia coli* strain was cultured in a medium supplemented with seleno-methionine. Therefore, the wobble amino acid should be methionine and these two proteins are supposed to possess the same functions.

The identified bacterial glycosidases are from *Thermotoga sp.* RQ2 and *Thermotoga petrophila* RKU-1. Unfortunately, species in genus *Thermotoga* are not major components of A/J mouse intestinal microbiota since they generally live deep underground in oil reservoirs as well as in other high temperature environments.

Therefore, searching for GRAS bacteria in A/J mouse intestinal microbiota that express bacterial glycosidases remains necessary in an attempt to identify probiotic intervention agents for chemoprevention of lung cancer.

In summary, we have successfully enriched and identified glycosidase PH1107-related [*Thermotoga petrophila* RKU-1] and glycosidase [*Thermotoga sp.* RQ2] from A/J mouse feces. These two lead bacterial glycosidases were found, for the first time, to hydrolyze ginsenoside Rd to F2 and Compound K.

Chapter 6 Summary

The key role of gut microbiota in chemoprevention efficacy of RGE has been recognized. Understanding of interactions between the intestinal microbiome and RGE assumes great importance since the interactions impact the production rates and extent of active ginsenosides and thereby the chemoprevetion efficacy of RGE. In this thesis work, we propose an intestinal microbiota based approach to elucidate the mechanism of bioactivation of RGE in A/J mouse due to the *in vitro* and *in vivo* discrepancy of bioactivities of RGE. Specifically, we aim to elucidate the contribution of gut microbiota to the chemoprevention efficacy of RGE at both the enzyme functional and genomic level. In the absence of whole genome shotgun pyrosequencing, 16s rRNA pyrosequencing together with ginsenoside hydrolysis assay have been used for this purpose.

We firstly elucidated the *in vitro* and *in vivo* discrepancy of bioactivities of RGE by investigating the anti-proliferative activities of various ginsenosides *in vitro*, the metabolic pathways of ginsenosides in A/J mouse fecal lysate and *in vivo* pharmacokinetics of ginsenosides in A/J mice (**Aim I**). Secondary ginsenoside Compound K, produced by bacterial glycosidases via Rb1-Rd-F2, exhibited higher anti-proliferative activities than primary ginsenosides. Significant amounts of secondary ginsenosides (F2 and Compound K) were found in the blood of A/J mice following oral administration of the primary ginsenoside Rb1. The results show that the efficacy of RGE demonstrated previously in A/J mice (Yan, Wang et al. 2006) was largely explained by enzymatic activity of bacterial glycosidases in the intestinal microbiota. Furthermore, we

characterized the kinetics of microbiota mediated biotransformation of ginesnosides by determining the corresponding kinetic parameters (metabolite formation rates) (**Aim I**). The formation of F2 from Rd was found, for the first time, to be the rate-limiting step in the biotransformation of Rb1 to Compound K. The findings indicate that there were multiple bacterial glycosidases with different catalytic capacities involved in the metabolism of ginsenoside Rb1. Therefore, bacterial glycosidases responsible for formation of ginsenoside F2 from Rd were studied in the following aims. Additionally, the intestinal microbiome of the A/J mouse capable of producing ginsenoside Compound K was characterized preliminary, which formed the basis for futher studies of how changes in intestinal microbiome will impact bioactivities of RGE *in vivo* (**Aim II**).

The central hypothesis, lung cancer chemoprevention of ginsenosides is mediated by the glycosidases activities of the A/J mouse intestinal microbiome defined by pyrosequencing, is futher validated in **Aim II**, where ginsenoside hydrolysis assay and 16s rRNA pyrosequencing were employed to examine the changes of glycosidases activities and intestinal microbiome in response to RGE treatment. Two sets of experiments (Dose Response study and RGE Interaction study) were performed to provide strong evidence: (1) 50 mg/kg RGE daily by oral gavage was the optimal dose to A/J mouse based on the activity of bacterial glycosidase; (2) A/J mouse could be classified into RGE responders (significant enhancement of bacterial glycosidases activities) and non-responders in response to RGE treatment (50 mg/kg daily by oral gavage); (3) the diversity and overall composition of intestinal microbiome, together with

relative abundance of certain genera (*Lactobacillus*) capable of RGE metabolism changed significantly in response to RGE treatment (50 mg/kg daily by oral gavage).

Lastly, a classic chromatographic “bottom up” approach was applied to enrich and identify bacterial glycosidases from A/J mouse feces, aiming at identifying bacteria species as agents for probiotic intervention (**Aim III**). The enriched and identified bacterial glycosidases (gi: 147736211, 501268188), were most likely β-xylosidases (EC 3.2.1.37) and were found, for the first time, to hydrolyze ginsenoside Rd to F2 and Compound K. The bacteria species expressing the enzymes were *Thermotoga petrophila* RKU-1 and *Thermotoga sp.* RQ2, respectively.

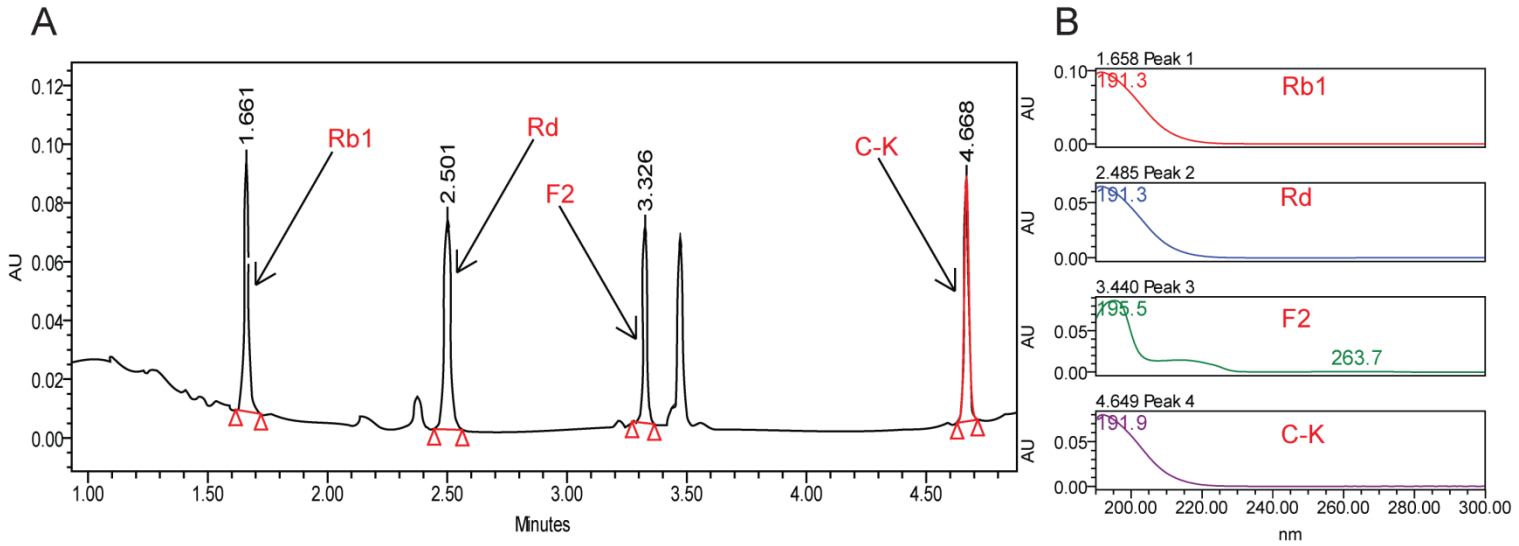
RGE as a chemoprevention agent against lung cancer remains to be largely tested in clinical trials as the data are extremely limited (Suh, Kroh et al. 2002; Sun, Lin et al. 2006; Yun, Zheng et al. 2010). This thesis demonstrats that the efficacy of RGE demonstrated previously in A/J mice was dependent on activation of bacterial glycosidases in the intestinal microbiota and proposes a novel approach for the chemoprevention of lung cancer by probiotic intervention. The work therefore opens the door to the prebiotic/probiotic intervention studies for cancer chemoprevention. It is noteworthy that our current studies are not without limitations. First, this work focuses only on ginsenoside Rb1, the most abundant primary ginsenosides in RGE. We believe that a comprehensive study incorporating other major primary ginsenosides in RGE, including Rg1, Re, Rc, Rb2 etc., will be of greater value. Second, our microbiome results in **Aim II** are based on 16s rRNA pyrosequencing data, revealing relative abundance of bacteria community in the gut of A/J mice at the genus level. A deeper sequencing

technique, such as whole genome shotgun (WGS) pyrosequencing is more straightforward in an attempt to elucidate relative abundance of bacteria community at the species level. Third, the classic “bottom up” approach identified two lead bacterial glycosidases that belong to genus *Thermotoga*. Species in genus *Thermotoga* are not major components of A/J mouse intestinal microbiota since they generally live deep underground in oil reservoirs as well as in other high temperature environments. Searching for GRAS bacteria appears to be necessary in an attempt to identify probiotic intervention agents for chemoprevention of lung cancer.

Taken together, this thesis represents great efforts towards a better understanding of contribution of gut microbiota to the chemoprevention efficacy of RGE. We anticipate that our novel approach (chemoprevention of lung cancer by probiotic intervention) may be applied to other drugs and dietary supplements with a similar *in vivo* activation mechanism.

Appendix A: UPLC chromatograms and UV spectrum of ginsenosides in fecal lysate in the UPLC assay (Section 3.3.6)

In the UPLC assay, chromatogram (panel A) and UV spectrum (panel B) of ginsenosides Rb1, Rd, F2 and Compound K in fecal lysate were presented below. The concentration of each ginesnoside was Rb1 (25 μg/mL), Rd (25 μg/mL), F2 (19.6 μg/mL) and Compound K (15.55 μg/mL).



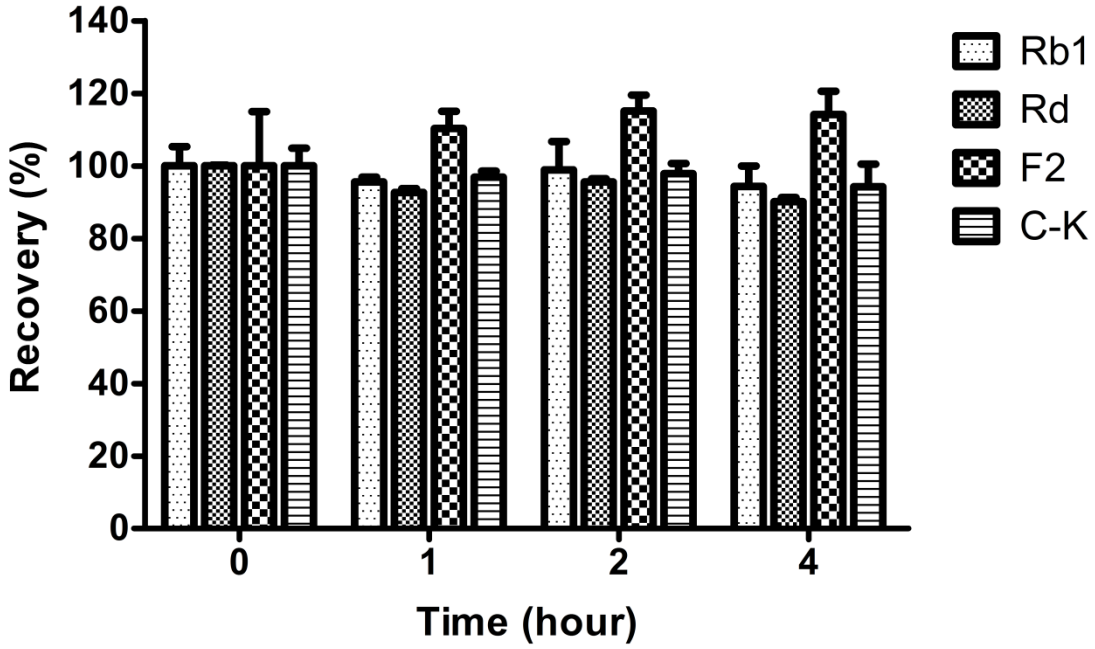
Appendix B: Extraction recovery, intra-day and inter-day precision and accuracy for ginsenosides Rb1, Rd, F2 and C-K in fecal lysate of the UPLC assay (Section 3.3.6)

Analyte	Concentration (μg/mL)	Extraction Recovery (n=3)	Intra-day (n=6)		Inter-day (n=6)	
		Average±SD (%)	Precision (CV, %)	Accuracy (Bias, %)	Precision (CV, %)	Accuracy (Bias, %)
Rb1	25	95.08 ± 3.46	1.88	111.46	2.63	109.99
	6.25	95.91 ± 2.73	2.01	111.73	2.77	111.98
	0.78	83.37 ± 4.47	5.87	102.61	4.86	102.57
Rd	25	88.27 ± 3.85	3.32	104.58	4.28	105.15
	6.25	86.50 ± 2.17	3.84	105.85	4.69	105.06
	0.78	81.86 ± 4.61	5.27	106.60	4.97	105.18
F2	19.6	90.82 ± 2.41	3.67	106.76	4.30	106.23
	4.9	80.61 ± 1.90	2.35	114.91	2.78	113.32
	0.306	71.76 ± 3.93	5.15	109.03	5.99	106.39
C-K	15.55	72.77 ± 2.94	4.47	104.86	4.64	105.02
	3.89	66.53 ± 0.26	2.05	114.67	2.27	111.27
	0.243	51.37 ± 0.93	14.18	109.33	15.73	110.78

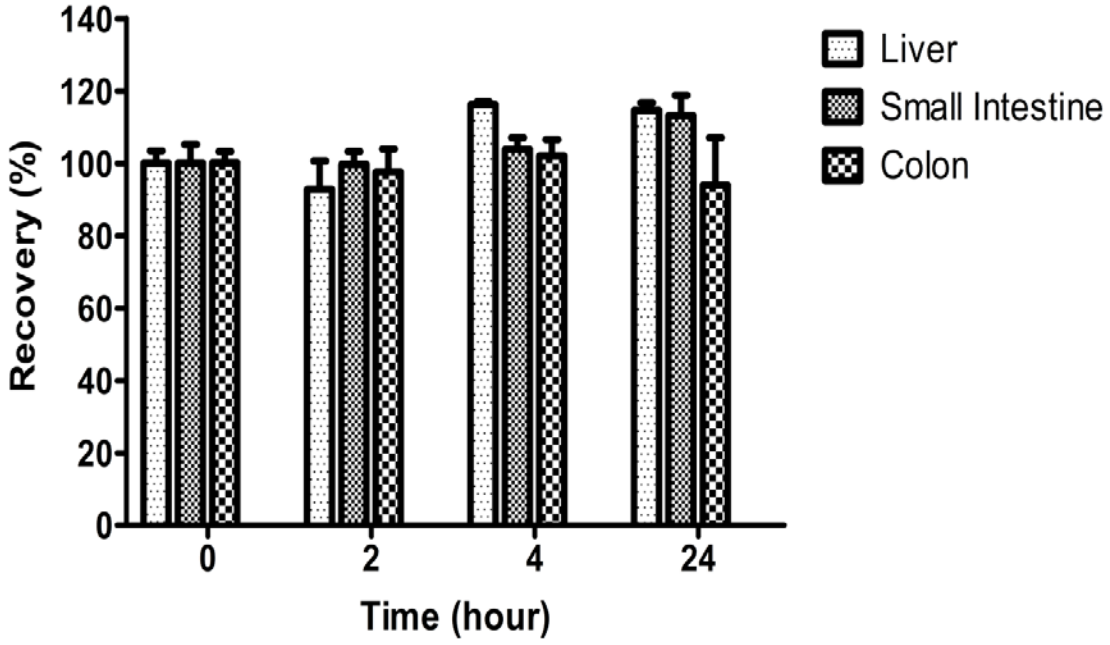
Appendix C: Stability of ginsenosides and glycosidases in various sample matrices

In this thesis work, the stability of ginsensoides in various sample matrices was tested. The stability of glycosidases in the fecal lysate was also evaluated.

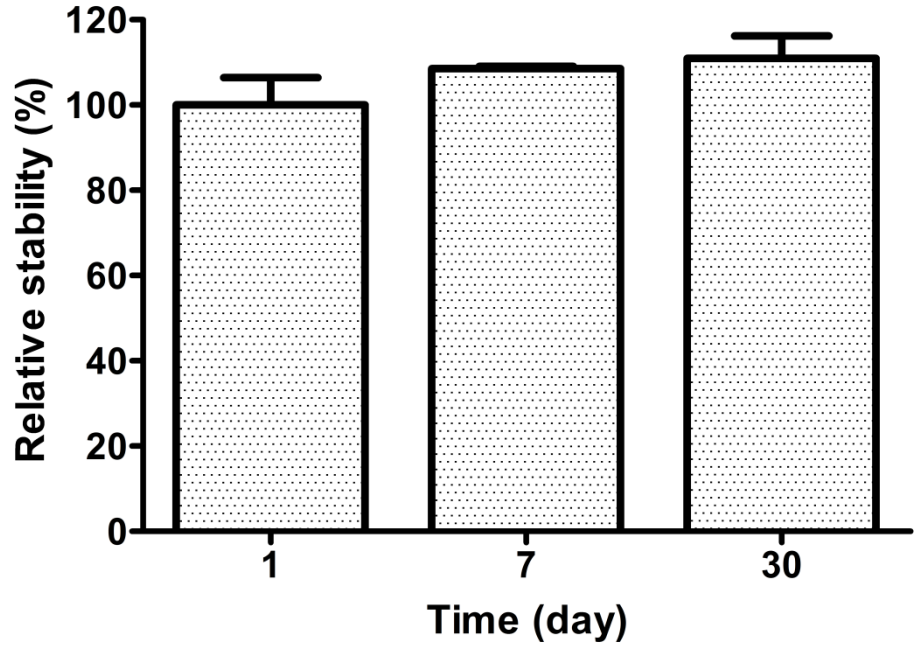
1. Stability of ginsenosides Rb1, Rd, F2 and Compound K in PBS buffer[a],
2. Stability of ginsenoside Rb1 in A/J mice liver, small intestine and colon S9[b],
3. Stability of glycosidases in fecal lysate up to 30 days[c],
4. Stability of glycosidases in fecal lysate up to 10 freeze-thaw cycles[d].



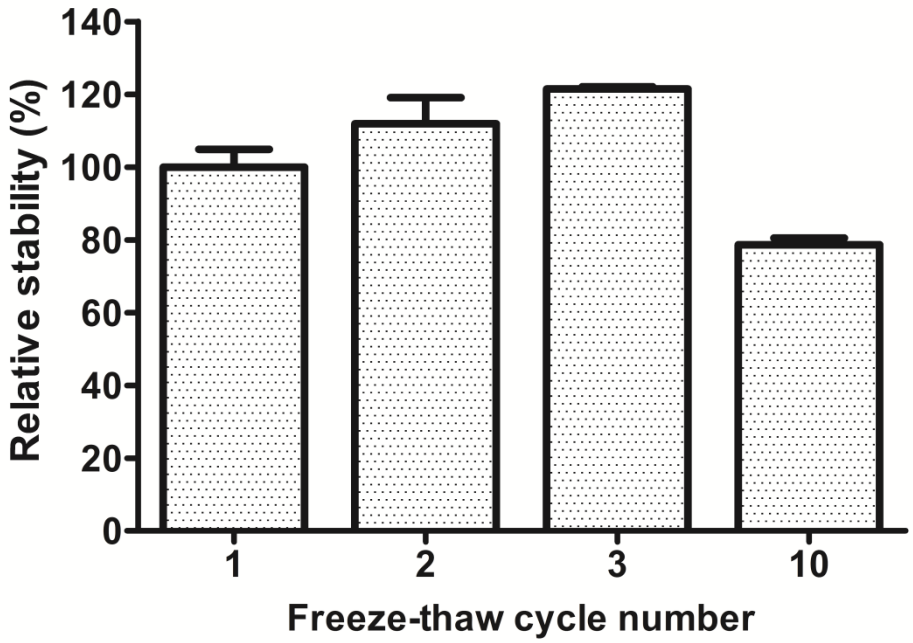
[a] Incubation was carried out at 37 ^{o}C up to 4 hours. The concentration of tested ginsenosides was 20 μM.



[b] Incubation was carried out at 37 ℃ up to 24 hours, with 20 μM Rb1 as the substrate.



[c] Fecal lysate was stored in -80° C freezer until use, with 20 μM Rb1 as the substrate.



[d] Fecal lysate was stored in -80℃ freezer until use, with 20 μM Rb1 as the substrate.

Appendix D: Codon optimization results of four potential ginsenoside hydrolyzing enzymes

In this thesis work, genes for glycosidase [*Thermotoga sp.* RQ2] (gi: 501268188), glycosidase PH1107-related [*Thermotoga petrophila* RKU-1] (gi: 147736211) were codon optimizatied for expression in *Escherichia coli*.

i, glycosidase [*Thermotoga sp.* RQ2] (gi: 501268188)

1, Codon usage bias adjustment

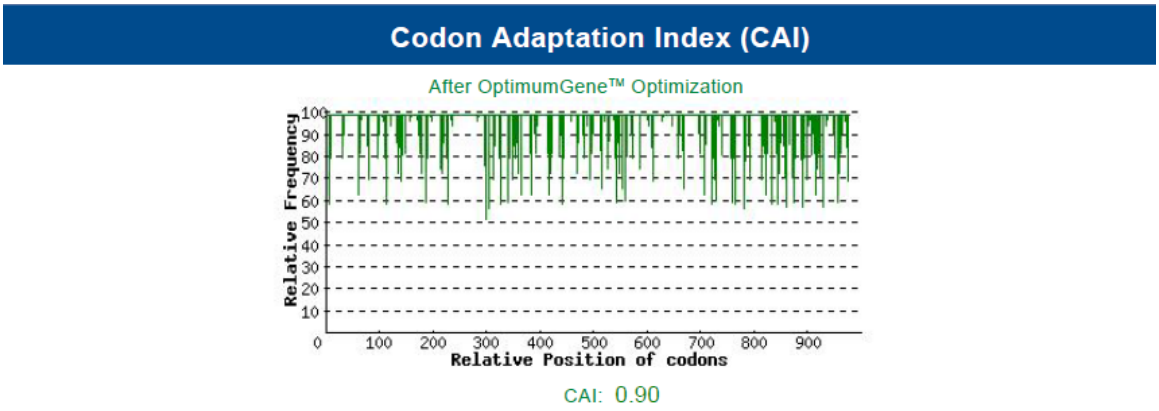


Figure 1a. The distribution of codon usage frequency along the length of the gene sequence. A CAI of 1.0 is considered to be perfect in the desired expression organism, and a CAI of > 0.8 is regarded as good, in terms of high gene expression level.

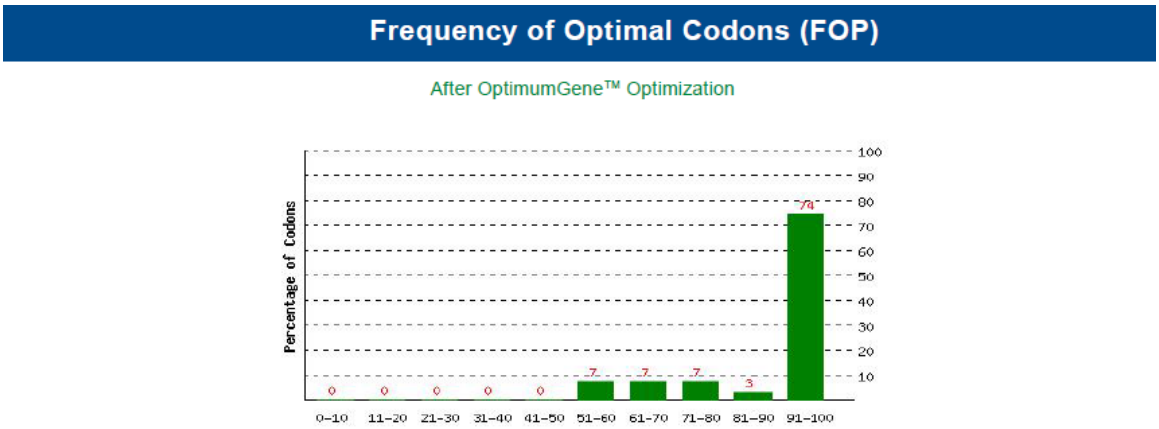
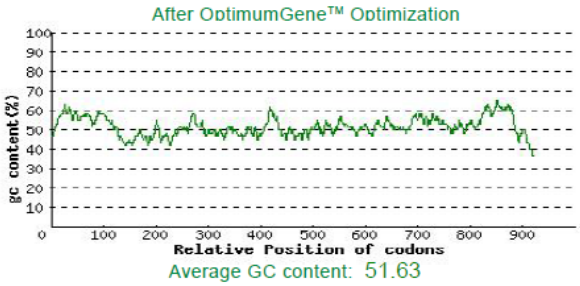


Figure 1b. The percentage distribution of codons in computed codon quality groups. The value of 100 is set for the codon with the highest usage frequency for a given amino acid in the desired expression organism.

2, GC Content Adjustment

GC Content Adjustment



Average GC content: 51.63

Figure 2. The ideal percentage range of GC content is between 30-70 %. Peaks of %GC content in a 60 bp window have been removed.

3, Restriction Enzymes and CIS-Acting Elements

Restriction Enzymes	Optimized

** Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.*

NdeI(CATATG)	1(1)
BamHI(GGATCC)	1(985)
BglII(AGATCT)	0
EcoRI(GAATTC)	0
HindIII(AAGCTT)	0
KpnI(GGTACC)	0
NcoI(CCATGG)	0
NotI(GCGGCCGC)	0
SalI(GTCGAC)	0
XbaI(TCTAGA)	0
XhoI(CTCGAG)	0
PflMI(CCANNNNNTGG)	0

CIS-Acting Elements	Optimized
E.coli_RBS(AGGAGG)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0
Chi_sites(GCTGGTGG)	0
T7Cis(ATCTGTT)	0
SD_like(GGRGGT)	0

4, Remove Repeat Sequences

Max Direct Repeat: Size: 9 Distance:620 Frequency:2

Max Inverted Repeat: Size: 9 Tm: 24.0 Start Positions: 296, 823

Max Dyad Repeat: Size: 10 Tm: 28.3 Start Positions: 331, 771

ii, glycosidase PH1107-related [*Thermotoga petrophila* RKU-1] (gi: 147736211)

1, Codon usage bias adjustment

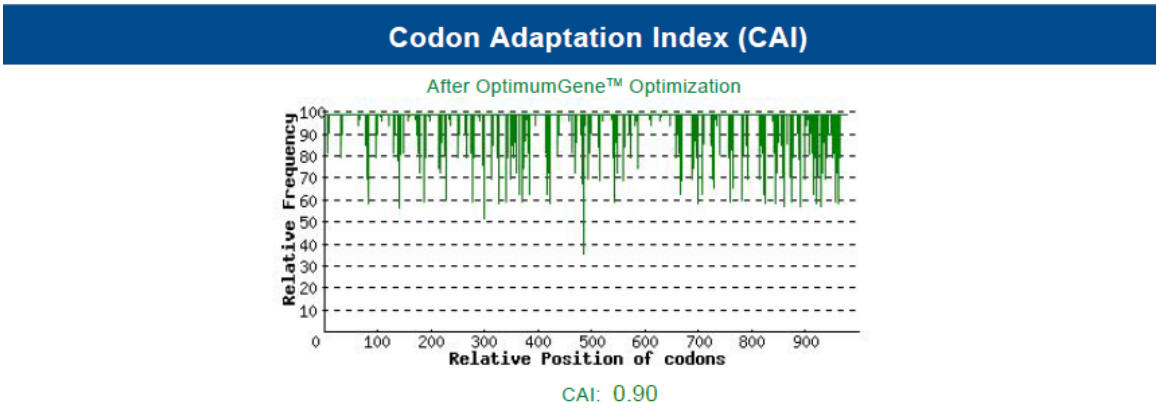


Figure 1a. The distribution of codon usage frequency along the length of the gene sequence. A CAI of 1.0 is considered to be perfect in the desired expression organism, and a CAI of > 0.8 is regarded as good, in terms of high gene expression level.

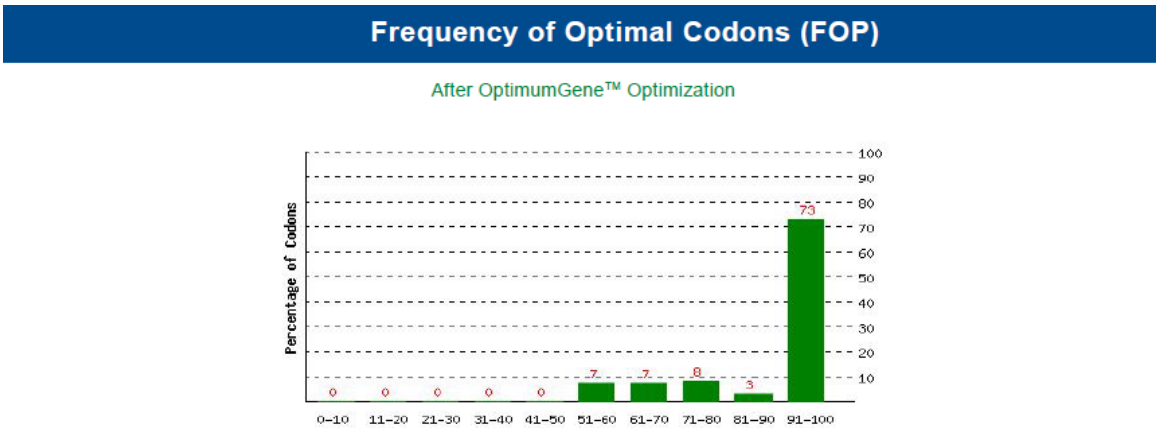
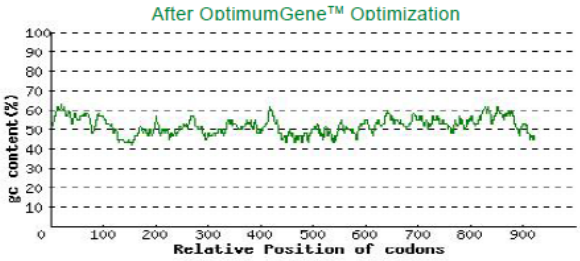


Figure 1b. The percentage distribution of codons in computed codon quality groups. The value of 100 is set for the codon with the highest usage frequency for a given amino acid in the desired expression organism.

2, GC Content Adjustment

GC Content Adjustment



Average GC content: 51.91

Figure 2. The ideal percentage range of GC content is between 30-70 %. Peaks of %GC content in a 60 bp window have been removed.

3, Restriction Enzymes and CIS-Acting Elements

Restriction Enzymes	Optimized

** Green: filtered sites; Blue: checked sites (not filtered); Red: kept sites.*

NdeI(CATATG)	1(1)
BamHI(GGATCC)	1(985)
BglII(AGATCT)	0
EcoRI(GAATTC)	0
HindIII(AAGCTT)	0
KpnI(GGTACC)	0
NcoI(CCATGG)	0
NotI(GCGGCCGC)	0
SalI(GTCGAC)	0
XbaI(TCTAGA)	0
XhoI(CTCGAG)	0
PflMI(CCANNNNNTGG)	0

CIS-Acting Elements	Optimized
E.coli_RBS(AGGAGG)	0
PolyT(TTTTTT)	0
PolyA(AAAAAAA)	0
Chi_sites(GCTGGTGG)	0
T7Cis(ATCTGTT)	0
SD_like(GGRGGT)	0

4, Remove Repeat Sequences

Max Direct Repeat: Size: 9 Distance: 546 Frequency:2

Max Inverted Repeat: Size: 9 Tm: 32.0 Start Positions: 911, 102

Max Dyad Repeat: Size: 10 Tm: 29.8 Start Positions: 772, 330

Appendix E: 20 Amino acids, their single-letter data-base codes (SLC), and their corresponding DNA codons

Amino Acid	SLC	DNA codons
Isoleucine	I	ATT, ATC, ATA
Leucine	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	V	GTT, GTC, GTA, GTG
Phenylalanine	F	TTT, TTC
Methionine	M	ATG
Cysteine	C	TGT, TGC
Alanine	A	GCT, GCC, GCA, GCG
Glycine	G	GGT, GGC, GGA, GGG
Proline	P	CCT, CCC, CCA, CCG
Threonine	T	ACT, ACC, ACA, ACG
Serine	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Y	TAT, TAC
Tryptophan	W	TGG
Glutamine	Q	CAA, CAG
Asparagine	N	AAT, AAC
Histidine	H	CAT, CAC
Glutamic acid	E	GAA, GAG
Aspartic acid	D	GAT, GAC
Lysine	K	AAA, AAG
Arginine	R	CGT, CGC, CGA, CGG, AGA, AGG
Stop codons	Stop	TAA, TAG, TGA

In this table, the twenty amino acids found in proteins are listed, along with the single-letter code used to represent these amino acids in protein data bases. The DNA codons representing each amino acid are also listed. All 64 possible 3-letter combinations of the DNA coding units T, C, A and G are used either to encode one of these amino acids or as one of the three stop codons that signals the end of a sequence. While DNA can be decoded unambiguously, it is not possible to predict a DNA sequence from its protein sequence. Because most amino acids have multiple codons, a number of possible DNA sequences might represent the same protein sequence.

References

Abu-Shanab, A. and E. M. Quigley (2010). "The role of the gut microbiota in nonalcoholic fatty liver disease." Nat Rev Gastroenterol Hepatol **7**(12): 691-701.

Akao T, K. H., Kanaoka M, Hattori M, Kobashi K (1998). "Intestinal bacterial hydrolysis is required for the appearance of compound K in rat plasma after oral administration of ginsenoside Rb1 from Panax ginseng." J Pharm Pharmacol. **50**(10): 1155-1160.

Akao T, K. H., Kanaoka M, Hattori M, Kobashi K (1998). "Intestinal bacterial hydrolysis is required for the appearance of compound K in rat plasma after oral administration of ginsenoside Rb1 from Panax ginseng." J Pharm Pharmacol. **50**(10): 1155-1160.

Akaza, H. (2012). "Prostate cancer chemoprevention by soy isoflavones: Role of intestinal bacteria as the "second human genome"." Cancer Science **103**(6): 969-975.

An, D. S., C. H. Cui, et al. (2010). "Identification and characterization of a novel Terrabacter ginsenosidimutans sp. nov. beta-glucosidase that transforms ginsenoside Rb1 into the rare gypenosides XVII and LXXV." Appl Environ Microbiol **76**(17): 5827-5836.

Angiuoli, S. V., M. Matalka, et al. (2011). "CloVR: A virtual machine for automated and portable sequence analysis from the desktop using cloud computing." Bmc Bioinformatics **12**.

Artursson, P. and J. Karlsson (1991). "Correlation between Oral-Drug Absorption in Humans and Apparent Drug Permeability Coefficients in Human Intestinal Epithelial (Caco-2) Cells." Biochemical and Biophysical Research Communications **175**(3): 880-885.

Arumugam, M., J. Raes, et al. (2011). "Enterotypes of the human gut microbiome." Nature **473**(7346): 174-180.

Atkinson, C., C. L. Frankenfeld, et al. (2005). "Gut bacterial metabolism of the soy isoflavone daidzein: Exploring the relevance to human health." Experimental Biology and Medicine **230**(3): 155-170.

Azcarate-Peril, M. A., M. Sikes, et al. (2011). "The intestinal microbiota, gastrointestinal environment and colorectal cancer: a putative role for probiotics in prevention of colorectal cancer?" American Journal of Physiology-Gastrointestinal and Liver Physiology **301**(3): G401-G424.

Backhed, F., J. K. Manchester, et al. (2007). "Mechanisms underlying the resistance to diet-induced obesity in germ-free mice." Proc Natl Acad Sci U S A **104**(3): 979-984.

Bae, E. A., M. K. Choo, et al. (2002). "Metabolism of ginsenoside R(c) by human intestinal bacteria and its related antiallergic activity." Biol Pharm Bull **25**(6): 743-747.

Bae, E. A., M. J. Han, et al. (2001). "In vitro anti-Helicobacter pylori activity of panaxytriol isolated from ginseng." Arch Pharm Res **24**(4): 297-299.

Bae, E. A., M. J. Han, et al. (2002). "Metabolism of 20(S)- and 20(R)-ginsenoside Rg3 by human intestinal bacteria and its relation to in vitro biological activities." Biol Pharm Bull **25**(1): 58-63.

Baumeister, P., M. Reiter, et al. (2012). "Curcumin and other polyphenolic compounds in head and neck cancer chemoprevention." Oxid Med Cell Longev **2012**: 902716.

Belinsky, S. A., K. J. Nikula, et al. (1998). "Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis." Proc Natl Acad Sci U S A **95**(20): 11891-11896.

Benamouzig, R., J. Deyra, et al. (2003). "Daily soluble aspirin and prevention of colorectal adenoma recurrence: one-year results of the APACC trial." Gastroenterology **125**(2): 328-336.

Bennett, W. P., T. V. Colby, et al. (1993). "p53 protein accumulates frequently in early bronchial neoplasia." Cancer Res **53**(20): 4817-4822.

Bettuzzi, S., M. Brausi, et al. (2006). "Chemoprevention of human prostate cancer by oral administration of green tea catechins in volunteers with high-grade prostate intraepithelial neoplasia: A preliminary report from a one-year proof-of-principle study." Cancer Research **66**(2): 1234-1240.

Bhatia, Y., S. Mishra, et al. (2002). "Microbial beta-glucosidases: cloning, properties, and applications." Crit Rev Biotechnol **22**(4): 375-407.

Bohannan, B. J. and J. Hughes (2003). "New approaches to analyzing microbial biodiversity data." Curr Opin Microbiol **6**(3): 282-287.

Bolam, D. N., E. D. Sonnenburg, et al. (2010). "Specificity of Polysaccharide Use in Intestinal Bacteroides Species Determines Diet-Induced Microbiota Alterations." Cell **141**(7): 1241-U1256.

Boleij, A. and H. Tjalsma (2012). "Gut bacteria in health and disease: a survey on the interface between intestinal microbiology and colorectal cancer." Biological Reviews **87**(3): 701-730.

Burcelin, R., V. Crivelli, et al. (2002). "Heterogeneous metabolic adaptation of C57BL/6J mice to high-fat diet." Am J Physiol Endocrinol Metab **282**(4): E834-842.

Castagnini, C., C. Luceri, et al. (2009). "Reduction of colonic inflammation in HLA-B27 transgenic rats by feeding Marie Menard apples, rich in polyphenols." Br J Nutr **102**(11): 1620-1628.

Chang, K. H., Jo, M. N., Kim, K.-T., Paik, H.-D. (2012). "Purification and Characterization of a Ginsenoside Rb1-Hydrolyzing β-Glucosidase from Aspergillus niger KCCM 11239." International Journal of Molecular Sciences. **13**(9): 12140-12152.

Chen, G., M. Yang, et al. (2007). "Microbial transformation of 20(S)-protopanaxatriol-type saponins by Absidia coerulea." J Nat Prod **70**(7): 1203-1206.

Chen, G., M. Yang, et al. (2008). "Comparative analysis on microbial and rat metabolism of ginsenoside Rb1 by high-performance liquid chromatography coupled with tandem mass spectrometry." Biomed Chromatogr **22**(7): 779-785.

Chen, G. T., M. Yang, et al. (2008). "Microbial transformation of ginsenoside Rb(1) by Acremonium strictum." Appl Microbiol Biotechnol **77**(6): 1345-1350.

Cheng, C. C., S. M. Yang, et al. (2005). "Molecular mechanisms of ginsenoside Rh2-mediated G1 growth arrest and apoptosis in human lung adenocarcinoma A549 cells." Cancer Chemother Pharmacol **55**(6): 531-540.

Cheng, L. Q., M. K. Kim, et al. (2006). "Conversion of major ginsenoside Rb1 to ginsenoside F2 by Caulobacter leidyia." Biotechnol Lett **28**(14): 1121-1127.

Cheng, L. Q., J. R. Na, et al. (2008). "Conversion of major ginsenoside Rb1 to 20(S)-ginsenoside Rg3 by Microbacterium sp. GS514." Phytochemistry **69**(1): 218-224.

Cheng, L. Q., J. R. Na, et al. (2007). "Microbial conversion of ginsenoside Rb1 to minor ginsenoside F2 and gypenoside XVII by Intrasporangium sp. GS603 isolated from soil." J Microbiol Biotechnol **17**(12): 1937-1943.

Chi, H. and G. E. Ji (2005). "Transformation of ginsenosides Rb1 and Re from Panax ginseng by food microorganisms." Biotechnol Lett **27**(11): 765-771.

Chi, H., D. H. Kim, et al. (2005). "Transformation of ginsenosides Rb2 and Rc from Panax ginseng by food microorganisms." Biol Pharm Bull **28**(11): 2102-2105.

Chinchetru, M. A., J. A. Cabezas, et al. (1989). "Purification and characterization of a broad specificity beta-glucosidase from sheep liver." Int J Biochem **21**(5): 469-476.

Chow, H. H. and I. A. Hakim (2011). "Pharmacokinetic and chemoprevention studies on tea in humans." Pharmacol Res **64**(2): 105-112.

Cimino, S., G. Sortino, et al. (2012). "Polyphenols: key issues involved in chemoprevention of prostate cancer." Oxid Med Cell Longev **2012**: 632959.

Clark, J. and M. You (2006). "Chemoprevention of lung cancer by tea." Mol Nutr Food Res **50**(2): 144-151.

Clemente, J. C., L. K. Ursell, et al. (2012). "The impact of the gut microbiota on human health: an integrative view." Cell **148**(6): 1258-1270.

Compare, D. and G. Nardone (2011). "Contribution of Gut Microbiota to Colonic and Extracolonic Cancer Development." Digestive Diseases **29**(6): 554-561.

Costello, E. K., C. L. Lauber, et al. (2009). "Bacterial Community Variation in Human Body Habitats Across Space and Time." Science **326**(5960): 1694-1697.

Cotter, P. D., C. Stanton, et al. (2012). "The Impact of Antibiotics on the Gut Microbiota as Revealed by High Throughput DNA Sequencing." Discovery Medicine **70**: 193-199.

Craciun, S. and E. P. Balskus (2012). "Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme." Proc Natl Acad Sci U S A **109**(52): 21307-21312.

Daly, A. K. (2012). "Genetic Polymorphisms Affecting Drug Metabolism: Recent Advances and Clinical Aspects." Current Concepts in Drug Metabolism and Toxicology **63**: 137-167.

Damman, C. J., S. I. Miller, et al. (2012). "The Microbiome and Inflammatory Bowel Disease: Is There a Therapeutic Role for Fecal Microbiota Transplantation?" American Journal of Gastroenterology **107**(10): 1452-1459.

Davis, C. D. and J. A. Milner (2009). "Gastrointestinal microflora, food components and colon cancer prevention." Journal of Nutritional Biochemistry **20**(10): 743-752.

de Vos, W. M. and E. A. de Vos (2012). "Role of the intestinal microbiome in health and disease: from correlation to causation." Nutr Rev **70 Suppl 1**: S45-56.

Decroos, K., S. Vanhemmens, et al. (2005). "Isolation and characterisation of an equol-producing mixed microbial culture from a human faecal sample and its activity under gastrointestinal conditions." Archives of Microbiology **183**(1): 45-55.

Dong, D., R. Ako, et al. (2012). "Understanding substrate selectivity of human UDP-glucuronosyltransferases through QSAR modeling and analysis of homologous enzymes." Xenobiotica **42**(8): 808-820.

Dong, D., R. Ako, et al. (2012). "Crystal structures of human sulfotransferases: insights into the mechanisms of action and substrate selectivity." Expert Opinion on Drug Metabolism & Toxicology **8**(6): 635-646.

Duperron, C. and A. Castonguay (1997). "Chemopreventive efficacies of aspirin and sulindac against lung tumorigenesis in A/J mice." Carcinogenesis **18**(5): 1001-1006.

Eckburg, P. B., E. M. Bik, et al. (2005). "Diversity of the human intestinal microbial flora." Science **308**(5728): 1635-1638.

Elmqvist, T., C. Folke, et al. (2003). "Response diversity, ecosystem change, and resilience." Frontiers in Ecology and the Environment **1**(9): 488-494.

Eloe-Fadrosh, E. A. and D. A. Rasko (2013). "The human microbiome: from symbiosis to pathogenesis." Annu Rev Med **64**: 145-163.

Euhus, D. M. and J. Diaz (2015). "Breast cancer prevention." Breast J **21**(1): 76-81.

Fielding, J. E. (1985). "Smoking: health effects and control (1)." N Engl J Med **313**(8): 491-498.

Frank, D. N., A. L. St Amand, et al. (2007). "Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases." Proc Natl Acad Sci U S A **104**(34): 13780-13785.

Gao, S., S. Basu, et al. (2012). "Bioavailability challenges associated with development of saponins as therapeutic and chemopreventive agents." Curr Drug Targets **13**(14): 1885-1899.

Goossens, D., D. Jonkers, et al. (2003). "The effect of Lactobacillus plantarum 299v on the bacterial composition and metabolic activity in faeces of healthy volunteers: a placebo-controlled study on the onset and duration of effects." Aliment Pharmacol Ther **18**(5): 495-505.

Govan, J. R. and V. Deretic (1996). "Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia." Microbiol Rev **60**(3): 539-574.

Guarner, F. and J. R. Malagelada (2003). "Gut flora in health and disease." Lancet **361**(9356): 512-519.

Guengerich, F. P. (2008). "Cytochrome P450 and chemical toxicology." Chemical Research in Toxicology **21**(1): 70-83.

Haiser, H. J. and P. J. Turnbaugh (2012). "Is It Time for a Metagenomic Basis of Therapeutics?" Science **336**(6086): 1253-1255.

Han, M., L. Ma, et al. (2015). "A nanoparticulate drug-delivery system for 20(S)-protopanaxadiol: formulation, characterization, increased oral bioavailability and anti-tumor efficacy." Drug Deliv: 1-9.

Hasegawa, H., J. H. Sung, et al. (1997). "Role of human intestinal Prevotella oris in hydrolyzing ginseng saponins." Planta Med **63**(5): 436-440.

Hecht, S. S. (2000). "Inhibition of carcinogenesis by isothiocyanates." Drug Metab Rev **32**(3-4): 395-411.

Hecht, S. S., F. Kassie, et al. (2009). "Chemoprevention of lung carcinogenesis in addicted smokers and ex-smokers." Nat Rev Cancer **9**(7): 476-488.

Henrissat, B. (1991). "A classification of glycosyl hydrolases based on amino acid sequence similarities." Biochem J **280 (Pt 2)**: 309-316.

Henrissat, B. and A. Bairoch (1993). "New families in the classification of glycosyl hydrolases based on amino acid sequence similarities." Biochem J **293 (Pt 3)**: 781-788.

Henrissat, B. and A. Bairoch (1996). "Updating the sequence-based classification of glycosyl hydrolases." Biochem J **316 (Pt 2)**: 695-696.

Herzog, C. R., R. A. Lubet, et al. (1997). "Genetic alterations in mouse lung tumors: implications for cancer chemoprevention." J Cell Biochem Suppl **28-29**: 49-63.

Holmes, E., J. V. Li, et al. (2012). "Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk." Cell Metab **16**(5): 559-564.

Hou, J. G., J. J. Xue, et al. (2012). "Highly selective microbial transformation of major ginsenoside Rb1 to gypenoside LXXV by Esteya vermicola CNU120806." Journal of Applied Microbiology **113**(4): 807-814.

Hou, J. G., J. J. Xue, et al. "Highly selective microbial transformation of major ginsenoside Rb(1) to gypenoside LXXV by Esteya vermicola CNU120806." J Appl Microbiol.

Howitt, M. R. and W. S. Garrett (2012). "A complex microworld in the gut: gut microbiota and cardiovascular disease connectivity." Nat Med **18**(8): 1188-1189.

Hu, M., J. Chen, et al. (1994). "Mechanism and kinetics of transcellular transport of a new beta-lactam antibiotic loracarbef across an intestinal epithelial membrane model system (Caco-2)." Pharm Res **11**(10): 1405-1413.

Hull, M. A., A. C. Sandell, et al. (2013). "A randomized controlled trial of eicosapentaenoic acid and/or aspirin for colorectal adenoma prevention during colonoscopic surveillance in the NHS Bowel Cancer Screening Programme (The seAFOod Polyp Prevention Trial): study protocol for a randomized controlled trial." Trials **14**: 237.

Huse, S. M., Y. Z. Ye, et al. (2012). "A Core Human Microbiome as Viewed through 16S rRNA Sequence Clusters." Plos One **7**(6).

Hylemon, P. B., H. P. Zhou, et al. (2009). "Bile acids as regulatory molecules." Journal of Lipid Research **50**(8): 1509-1520.

Jalbert, G. and A. Castonguay (1992). "Effects of NSAIDs on NNK-induced pulmonary and gastric tumorigenesis in A/J mice." Cancer Lett **66**(1): 21-28.

Johnson, T. E., D. Hermanson, et al. (2011). "Lung tumorigenesis suppressing effects of a commercial kava extract and its selected compounds in A/J mice." Am J Chin Med **39**(4): 727-742.

Johnston, B. C., J. Z. Goldenberg, et al. (2011). "Probiotics for the prevention of pediatric antibiotic-associated diarrhea." Cochrane Database Syst Rev(11): CD004827.

Johnston, B. C., A. L. Supina, et al. (2007). "Probiotics for the prevention of pediatric antibiotic-associated diarrhea." Cochrane Database Syst Rev(2): CD004827.

Ju, J., G. Lu, et al. (2007). "Inhibition of carcinogenesis by tea constituents." Seminars in Cancer Biology **17**(5): 395-402.

Kallus, S. J. and L. J. Brandt (2012). "The Intestinal Microbiota and Obesity." Journal of Clinical Gastroenterology **46**(1): 16-24.

Kamangar, F., G. M. Dores, et al. (2006). "Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world." J Clin Oncol **24**(14): 2137-2150.

Keith, R. L., V. Karoor, et al. (2010). "Chemoprevention of murine lung cancer by gefitinib in combination with prostacyclin synthase overexpression." Lung Cancer **70**(1): 37-42.

Keller, A., A. Nesvizhskii, E. Kolker, and R. Aebersold (2002). "Empirical statistical model to estimate the accuracy of peptide identifications made by ms/ms and database search." **74**: 5383-5392.

Kelly, K., J. Kittelson, et al. (2009). "A Randomized Phase II Chemoprevention Trial of 13-CIS Retinoic Acid with Or without alpha Tocopherol or Observation in Subjects at High Risk for Lung Cancer." Cancer Prevention Research **2**(5): 440-449.

Khan, N., F. Afaq, et al. (2008). "Cancer chemoprevention through dietary antioxidants: Progress and promise." Antioxidants & Redox Signaling **10**(3): 475-510.

Kim, E. M., J. Kim, et al. (2012). "Identification and characterization of the Rhizobium sp. strain GIN611 glycoside oxidoreductase resulting in the deglycosylation of ginsenosides." Appl Environ Microbiol **78**(1): 242-249.

Kim, J., Y. Wang, et al. (2012). "Cloning and expression of β-glucosidases from Bifidobacterium lactis AD011." Food Science and Biotechnology **21**(3): 731-738.

Kim, M. K., J. W. Lee, et al. (2005). "Microbial conversion of major ginsenoside rb(1) to pharmaceutically active minor ginsenoside rd." J Microbiol **43**(5): 456-462.

Kim, S. N., Y. W. Ha, et al. (2007). "Simultaneous quantification of 14 ginsenosides in Panax ginseng C.A. Meyer (Korean red ginseng) by HPLC-ELSD and its application to quality control." J Pharm Biomed Anal **45**(1): 164-170.

Kim, Y., D. Lee, et al. (2008). "Inhibition of proliferation in colon cancer cell lines and harmful enzyme activity of colon bacteria by Bifidobacterium adolescentis SPM0212." Arch Pharm Res **31**(4): 468-473.

Kim, Y. S., J. J. Kim, et al. (2008). "Biotransformation of ginsenoside Rb1, crocin, amygdalin, geniposide, puerarin, ginsenoside Re, hesperidin, poncirin, glycyrrhizin, and baicalin by human fecal microflora and its relation to cytotoxicity against tumor cells." J Microbiol Biotechnol **18**(6): 1109-1114.

Kishimoto, Y., K. Sugio, et al. (1995). "Allele-specific loss in chromosome 9p loci in preneoplastic lesions accompanying non-small-cell lung cancers." J Natl Cancer Inst **87**(16): 1224-1229.

Kiyotani, K., T. Mushiroda, et al. (2012). "Pharmacogenomics of Tamoxifen: Roles of Drug Metabolizing Enzymes and Transporters." Drug Metabolism and Pharmacokinetics **27**(1): 122-131.

Klepac-Ceraj, V., K. P. Lemon, et al. (2010). "Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and Pseudomonas aeruginosa." Environmental Microbiology **12**(5): 1293-1303.

Kong, H., M. Wang, et al. (2009). "Bioconversion of red ginseng saponins in the gastro-intestinal tract in vitro model studied by high-performance liquid chromatography-high resolution Fourier transform ion cyclotron resonance mass spectrometry." J Chromatogr A **1216**(11): 2195-2203.

Lee, B. H., S. J. Lee, et al. (1998). "In vitro antigenotoxic activity of novel ginseng saponin metabolites formed by intestinal bacteria." Planta Med **64**(6): 500-503.

Lee, J. H., J. S. Shim, et al. (2009). "Inhibition of pathogen adhesion to host cells by polysaccharides from Panax ginseng." Biosci Biotechnol Biochem **73**(1): 209-212.

Lee, J. H., J. S. Shim, et al. (2006). "Pectin-like acidic polysaccharide from Panax ginseng with selective antiadhesive activity against pathogenic bacteria." Carbohydr Res **341**(9): 1154-1163.

Lee, Y. S., J. S. Seo, et al. (1991). "Inhibitory effects of biochanin A on mouse lung tumor induced by benzo(a)pyrene." J Korean Med Sci **6**(4): 325-328.

Lemon, K. P., G. C. Armitage, et al. (2012). "Microbiota-Targeted Therapies: An Ecological Perspective." Science Translational Medicine **4**(137).

Ley, R. E., F. Backhed, et al. (2005). "Obesity alters gut microbial ecology." Proc Natl Acad Sci U S A **102**(31): 11070-11075.

Ley, R. E., D. A. Peterson, et al. (2006). "Ecological and evolutionary forces shaping microbial diversity in the human intestine." Cell **124**(4): 837-848.

Li, H., J. H. Lee, et al. (2008). "Effective purification of ginsenosides from cultured wild ginseng roots,red ginseng, and white ginseng with macroporous resins." J Microbiol Biotechnol **18**(11): 1789-1791.

Liao, J., G. Y. Yang, et al. (2004). "Inhibition of lung carcinogenesis and effects on angiogenesis and apoptosis in A/J mice by oral administration of green tea." Nutrition and Cancer-an International Journal **48**(1): 44-53.

Limburg, P. J., M. R. Mahoney, et al. (2011). "Randomized Phase II Trial of Sulindac, Atorvastatin, and Prebiotic Dietary Fiber for Colorectal Cancer Chemoprevention." Cancer Prevention Research **4**(2): 259-269.

Limburg, P. J., S. J. Mandrekar, et al. (2013). "Randomized phase II trial of sulindac for lung cancer chemoprevention." Lung Cancer **79**(3): 254-261.

Liu, Y. and M. Hu (2002). "Absorption and metabolism of flavonoids in the Caco-2 cell culture model and a perfused rat intestinal model." Drug Metabolism and Disposition **30**(4): 370-377.

Luciano, F. B. and R. A. Holley (2009). "Enzymatic inhibition by allyl isothiocyanate and factors affecting its antimicrobial action against Escherichia coli O157:H7." International Journal of Food Microbiology **131**(2-3): 240-245.

Malkinson, A. M. (1992). "Primary lung tumors in mice: an experimentally manipulable model of human adenocarcinoma." Cancer Res **52**(9 Suppl): 2670s-2676s.

Mandard, S., F. Zandbergen, et al. (2006). "The fasting-induced adipose factor/angiopoietin-like protein 4 is physically associated with lipoproteins and governs plasma lipid levels and adiposity." J Biol Chem **281**(2): 934-944.

Marteau, P., P. Pochart, et al. (1990). "Effect of chronic ingestion of a fermented dairy product containing Lactobacillus acidophilus and Bifidobacterium bifidum on metabolic activities of the colonic flora in humans." Am J Clin Nutr **52**(4): 685-688.

Maslowski, K. M., A. T. Vieira, et al. (2009). "Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43." Nature **461**(7268): 1282-1286.

Maynard, C. L., C. O. Elson, et al. (2012). "Reciprocal interactions of the intestinal microbiota and immune system." Nature **489**(7415): 231-241.

McBain, A. J. and G. T. Macfarlane (1998). "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 **47**(5): 407-416.

McDoniels-Silvers, A. L., C. R. Herzog, et al. (2001). "Inactivation of both Rb and p53 pathways in mouse lung epithelial cell lines." Experimental Lung Research **27**(3): 297-318.

Merlo, A., J. G. Herman, et al. (1995). "5' CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers." Nat Med **1**(7): 686-692.

Mimoto, J., K. Kiura, et al. (2000). "(-)-epigallocatechin gallate can prevent cisplatin-induced lung tumorigenesis in A/J mice." Carcinogenesis **21**(5): 915-919.

Miyanaga, N., H. Akaza, et al. (2012). "Prostate Cancer Chemoprevention Study: An investigative randomized control study using purified isoflavones in men with rising prostate-specific antigen." Cancer Science **103**(1): 125-130.

Morse, M. A., S. G. Amin, et al. (1989). "Effects of aromatic isothiocyanates on tumorigenicity, O6-methylguanine formation, and metabolism of the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in A/J mouse lung." Cancer Res **49**(11): 2894-2897.

Mouzaki, M. and J. P. Allard (2012). "The role of nutrients in the development, progression, and treatment of nonalcoholic fatty liver disease." J Clin Gastroenterol **46**(6): 457-467.

Muthyala, R. S., Y. H. Ju, et al. (2004). "Equol, a natural estrogenic metabolite from soy isoflavones: convenient preparation and resolution of R- and S-equols and their differing binding and biological activity through estrogen receptors alpha and beta." Bioorganic & Medicinal Chemistry **12**(6): 1559-1567.

Nag, S. A., J. J. Qin, et al. (2012). "Ginsenosides as anticancer agents: in vitro and in vivo activities, structure-activity relationships, and molecular mechanisms of action." Frontiers in Pharmacology **3**.

Nagar, S. and R. L. Blanchard (2006). "Pharmacogenetics of uridine diphosphoglucuronosyltransferase (UGT) 1A family members and its role in patient response to irinotecan." Drug Metabolism Reviews **38**(3): 393-409.

Nicholson, J. K., E. Holmes, et al. (2012). "Host-gut microbiota metabolic interactions." Science **336**(6086): 1262-1267.

Niu, T., D. L. Smith, et al. (2013). "Bioactivity and bioavailability of ginsenosides are dependent on the glycosidase activities of the A/J mouse intestinal microbiome defined by pyrosequencing." Pharm Res **30**(3): 836-846.

Oakley, A. (2011). "Glutathione transferases: a structural perspective." Drug Metabolism Reviews **43**(2): 138-151.

Ohashi, K., N. Takigawa, et al. (2009). "Chemopreventive effects of gefitinib on nonsmoking-related lung tumorigenesis in activating epidermal growth factor receptor transgenic mice." Cancer Res **69**(17): 7088-7095.

Panwar, M., R. Samarth, et al. (2005). "Inhibition of benzo(a)pyrene induced lung adenoma by panax ginseng extract, EFLA400, in Swiss albino mice." Biol Pharm Bull **28**(11): 2063-2067.

Park, C. S., M. H. Yoo, et al. (2010). "Biotransformation of ginsenosides by hydrolyzing the sugar moieties of ginsenosides using microbial glycosidases." Applied Microbiology and Biotechnology **87**(1): 9-19.

Phillips, M. L. (2009). "Gut Reaction Environmental Effects on the Human Microbiota." Environmental Health Perspectives **117**(5): A198-A205.

Qi, L. W., C. Z. Wang, et al. (2010). "American ginseng: Potential structure-function relationship in cancer chemoprevention." Biochemical Pharmacology **80**(7): 947-954.

Qin, J., R. Li, et al. (2010). "A human gut microbial gene catalogue established by metagenomic sequencing." Nature **464**(7285): 59-65.

Quan, L. H., C. Wang, et al. (2013). "Isolation and characterization of novel ginsenoside-hydrolyzing glycosidase from Microbacterium esteraromaticum that transforms ginsenoside Rb2 to rare ginsenoside 20(S)-Rg3." Antonie Van Leeuwenhoek **104**(1): 129-137.

Rafter, J., M. Bennett, et al. (2007). "Dietary synbiotics reduce cancer risk factors in polypectomized and colon cancer patients." American Journal of Clinical Nutrition **85**(2): 488-496.

Relman, D. A., L. Dethlefsen, et al. (2008). "The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing." Plos Biology **6**(11): 2383-2400.

Ridlon, J. M., D. J. Kang, et al. (2006). "Bile salt biotransformations by human intestinal bacteria." Journal of Lipid Research **47**(2): 241-259.

Riesenfeld, C. S., P. D. Schloss, et al. (2004). "Metagenomics: genomic analysis of microbial communities." Annu Rev Genet **38**: 525-552.

Saavedra, M. J., A. Borges, et al. (2010). "Antimicrobial Activity of Phenolics and Glucosinolate Hydrolysis Products and their Synergy with Streptomycin against Pathogenic Bacteria." Medicinal Chemistry **6**(3): 174-183.

Samuel, B. S. and J. I. Gordon (2006). "A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism." Proceedings of the National Academy of Sciences of the United States of America **103**(26): 10011-10016.

Samuel, B. S., A. Shaito, et al. (2008). "Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41." Proceedings of the National Academy of Sciences of the United States of America **105**(43): 16767-16772.

Schloss, P. D., S. L. Westcott, et al. (2009). "Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities." Appl Environ Microbiol **75**(23): 7537-7541.

Scott, E. N., A. J. Gescher, et al. (2009). "Development of Dietary Phytochemical Chemopreventive Agents: Biomarkers and Choice of Dose for Early Clinical Trials." Cancer Prevention Research **2**(6): 525-530.

Scott, K. P., S. W. Gratz, et al. (2013). "The influence of diet on the gut microbiota." Pharmacol Res **69**(1): 52-60.

Sears, C. L. and D. M. Pardoll (2011). "Perspective: Alpha-Bugs, Their Microbial Partners, and the Link to Colon Cancer." Journal of Infectious Diseases **203**(3): 306-311.

SH, Y. T. a. K. (1988). "Inhibition of development of benzo(a)-pyrene-induced mouse pulmonary adenoma by natural products in medium-term bioassay system." J Korean Cancer Assoc **20**: 133 - 142.

Shibata, S. (2001). "Chemistry and cancer preventing activities of ginseng saponins and some related triterpenoid compounds." J Korean Med Sci **16 Suppl**: S28-37.

Shin, K. C., H. J. Lee, et al. (2015). "Substrate specificity of beta-glucosidase from Gordonia terrae for ginsenosides and its application in the production of ginsenosides Rg3, Rg2, and Rh1 from ginseng root extract." J Biosci Bioeng **119**(5): 497-504.

Siegel, R., J. Ma, et al. (2014). "Cancer statistics, 2014." CA Cancer J Clin **64**(1): 9-29.

Sobhani, I., J. Tap, et al. (2011). "Microbial dysbiosis in colorectal cancer (CRC) patients." PLoS One **6**(1): e16393.

Son, J. W., H. J. Kim, et al. (2008). "Ginsenoside Rd production from the major ginsenoside Rb(1) by beta-glucosidase from Thermus caldophilus." Biotechnol Lett **30**(4): 713-716.

Sozzi, G., M. Miozzo, et al. (1992). "Deletions of 17p and p53 mutations in preneoplastic lesions of the lung." Cancer Res **52**(21): 6079-6082.

Stagos, D., G. D. Amoutzias, et al. (2012). "Chemoprevention of liver cancer by plant polyphenols." Food Chem Toxicol **50**(6): 2155-2170.

Steinbach, G., P. M. Lynch, et al. (2000). "The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis." N Engl J Med **342**(26): 1946-1952.

Streit, W. R. and R. A. Schmitz (2004). "Metagenomics--the key to the uncultured microbes." Curr Opin Microbiol **7**(5): 492-498.

Strober, W., I. J. Fuss, et al. (2002). "The immunology of mucosal models of inflammation." Annu Rev Immunol **20**: 495-549.

Suh, S. O., M. Kroh, et al. (2002). "Effects of red ginseng upon postoperative immunity and survival in patients with stage III gastric cancer." American Journal of Chinese Medicine **30**(4): 483-494.

Sun, S., L. W. Qi, et al. (2011). "Red notoginseng Higher ginsenoside content and stronger anticancer potential than Asian and American ginseng." Food Chemistry **125**(4): 1299-1305.

Sun, Y., H. Lin, et al. (2006). "[A randomized, prospective, multi-centre clinical trial of NP regimen (vinorelbine+cisplatin) plus Gensing Rg3 in the treatment of advanced non-small cell lung cancer patients]." Zhongguo Fei Ai Za Zhi **9**(3): 254-258.

Swidsinski, A., V. Loening-Baucke, et al. (2008). "Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora." Inflammatory Bowel Diseases **14**(2): 147-161.

Tan, X. L. and S. D. Spivack (2009). "Dietary chemoprevention strategies for induction of phase II xenobiotic-metabolizing enzymes in lung carcinogenesis: A review." Lung Cancer **65**(2): 129-137.

Thun, M. J. and B. Blackard (2009). "Pharmacologic effects of NSAIDs and implications for the risks and benefits of long-term prophylactic use of aspirin to prevent cancer." Recent Results Cancer Res **181**: 215-221.

Tremaroli, V. and F. Backhed (2012). "Functional interactions between the gut microbiota and host metabolism." Nature **489**(7415): 242-249.

Turnbaugh, P. J., R. E. Ley, et al. (2007). "The human microbiome project." Nature **449**(7164): 804-810.

Turnbaugh, P. J., R. E. Ley, et al. (2006). "An obesity-associated gut microbiome with increased capacity for energy harvest." Nature **444**(7122): 1027-1031.

Turnbaugh, P. J., C. Quince, et al. (2010). "Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins." Proceedings of the National Academy of Sciences of the United States of America **107**(16): 7503-7508.

Vipperla, K. and S. J. O'Keefe (2012). "The microbiota and its metabolites in colonic mucosal health and cancer risk." Nutr Clin Pract **27**(5): 624-635.

Wallace, B. D., H. W. Wang, et al. (2010). "Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme." Science **330**(6005): 831-835.

Wang, S. W., J. Chen, et al. (2006). "Disposition of flavonoids via enteric recycling: structural effects and lack of correlations between in vitro and in situ metabolic properties." Drug Metab Dispos **34**(11): 1837-1848.

Wang, Y. W., J. D. Hoenig, et al. (2009). "16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis." Isme Journal **3**(8): 944-954.

Wang, Z. N., E. Klipfell, et al. (2011). "Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease." Nature **472**(7341): 57-U82.

Wattenberg, L. W. and R. D. Estensen (1996). "Chemopreventive effects of myo-inositol and dexamethasone on benzo[a]pyrene and 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone-induced pulmonary carcinogenesis in female A/J mice." Cancer Res **56**(22): 5132-5135.

Wei, Y., W. Zhao, et al. (2011). "Purification and characterization of a novel and unique ginsenoside Rg1-hydrolyzing beta-D-glucosidase from Penicillium sclerotiorum." Acta Biochim Biophys Sin (Shanghai) **43**(3): 226-231.

Whitman, W. B., D. C. Coleman, et al. (1998). "Prokaryotes: the unseen majority." Proc Natl Acad Sci U S A **95**(12): 6578-6583.

Whittaker, R. H. (1972). "Evolution and Measurement of Species Diversity." Taxon **21**.

Witschi, H., I. Espiritu, et al. (1997). "The carcinogenic potential of the gas phase of environmental tobacco smoke." Carcinogenesis **18**(11): 2035-2042.

Witschi, H., I. Espiritu, et al. (1997). "The carcinogenicity of environmental tobacco smoke." Carcinogenesis **18**(3): 575-586.

Wong, J. M. W., R. de Souza, et al. (2006). "Colonic health: Fermentation and short chain fatty acids." Journal of Clinical Gastroenterology **40**(3): 235-243.

Wu, G. D., J. Chen, et al. (2011). "Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes." Science **334**(6052): 105-108.

Yan, Q., W. Zhou, et al. (2008). "Purification method improvement and characterization of a novel ginsenoside-hydrolyzing beta-glucosidase from Paecilomyces Bainier sp. 229." Biosci Biotechnol Biochem **72**(2): 352-359.

Yan, Y., Y. Wang, et al. (2006). "Efficacy of polyphenon E, red ginseng, and rapamycin on benzo(a)pyrene-induced lung tumorigenesis in A/J mice." Neoplasia **8**(1): 52-58.

Yang, C. S., S. Sang, et al. (2008). "Bioavailability issues in studying the health effects of plant polyphenolic compounds." Mol Nutr Food Res **52 Suppl 1**: S139-151.

Yang, C. S., X. Wang, et al. (2009). "Cancer prevention by tea: animal studies, molecular mechanisms and human relevance." Nat Rev Cancer **9**(6): 429-439.

Yang, Z., S. Gao, et al. (2011). "Enhancement of Oral Bioavailability of 20(S)-Ginsenoside Rh2 through Improved Understanding of Its Absorption and Efflux Mechanisms." Drug Metabolism and Disposition **39**(10): 1866-1872.

Yang, Z., S. Gao, et al. (2010). "Biopharmaceutical and pharmacokinetic characterization of matrine as determined by a sensitive and robust UPLC-MS/MS method." Journal of Pharmaceutical and Biomedical Analysis **51**(5): 1120-1127.

Yang Z, G. S., Wang J, Yin T, Teng Y, Wu B, You M, Jiang Z, Hu M. (2011). "Enhancement of Oral Bioavailability of Ginsenoside 20(s)-Rh2 through Improved Understanding of its Absorption and Efflux Mechanisms." Drug Metab Dispos.

Yang, Z., K. Kulkarni, et al. (2012). "Bioavailability and pharmacokinetics of genistein: mechanistic studies on its ADME." Anticancer Agents Med Chem **12**(10): 1264-1280.

Yang, Z., J. R. Wang, et al. (2012). "Inhibition of P-glycoprotein leads to improved oral bioavailability of compound K, an anticancer metabolite of red ginseng extract produced by gut microflora." Drug Metab Dispos **40**(8): 1538-1544.

Yang, Z., W. Zhu, et al. (2012). "Breast Cancer Resistance Protein (ABCG2) Determines Distribution of Genistein Phase II Metabolites: Reevaluation of the Roles of ABCG2 in the Disposition of Genistein." Drug Metabolism and Disposition **40**(10): 1883-1893.

Yue, P. Y., D. Y. Wong, et al. (2006). "The angiosuppressive effects of 20(R)- ginsenoside Rg3." Biochem Pharmacol **72**(4): 437-445.

Yun, T. K. (1991). "Usefulness of medium-term bioassay determining formations of pulmonary adenoma in NIH(GP) mice for finding anticarcinogenic agents from natural products." J Toxicol Sci **16 Suppl 1**: 53-62.

Yun, T. K. (1999). "Update from Asia. Asian studies on cancer chemoprevention." Ann N Y Acad Sci **889**: 157-192.

Yun, T. K. (2001). "Panax ginseng--a non-organ-specific cancer preventive?" Lancet Oncol **2**(1): 49-55.

Yun, T. K., S. H. Kim, et al. (1995). "Trial of a New Medium-Term Model Using Benzo(a)Pyrene Induced Lung-Tumor in Newborn Mice." Anticancer Research **15**(3): 839-845.

Yun TK, K. S., and Oh YR (1987). "Medium-term (9 weeks) method for assay of preventive agents against tumor." J Korean Cancer Assoc **19**: 1-7.

Yun TK, K. S., Lee YS (1995). "Trial of a new medium-term model using benzo(a)pyrene induced lung tumor in newborn mice." Anticancer Res **15**(3): 839-845.

Yun, T. K., Y. S. Lee, et al. (2001). "Anticarcinogenic effect of Panax ginseng C.A. Meyer and identification of active compounds." J Korean Med Sci **16 Suppl**: S6-18.

Yun TK, Y. Y., Han IW (1983). "Anticarcinogenic effect of long-term oral administration of red ginseng on newborn mice exposed to various chemical carcinogens." Cancer Detect Prev **6**(6): 515-525.

Yun, T. K., S. Zheng, et al. (2010). "Non-organ-specific preventive effect of long-term administration of Korean red ginseng extract on incidence of human cancers." J Med Food **13**(3): 489-494.

Zhou, P., W. Zhou, et al. (2008). "Biotransformation of Panax notoginseng saponins into ginsenoside compound K production by Paecilomyces bainier sp 229." Journal of Applied Microbiology **104**(3): 699-706.

Zhu, M., Y. Chen, et al. (2000). "Oral absorption and bioavailability of tea catechins." Planta Med **66**(5): 444-447.

Zimmerman, M. A., N. Singh, et al. (2012). "Butyrate suppresses colonic inflammation through HDAC1-dependent Fas upregulation and Fas-mediated apoptosis of T cells." American Journal of Physiology-Gastrointestinal and Liver Physiology **302**(12): G1405-G1415.