2008

CREATION OF A MOUSE WITH A HUMANIZED fpgs GENE COMPATIBLE WITH NORMAL DEVELOPMENT

Linying Xie
Virginia Commonwealth University

Virginia Commonwealth University
School of Medicine

This is to certify that the thesis prepared by Linying Xie entitled CREATION OF A MOUSE WITH A HUMANIZED *fpgs* GENE COMPATIBLE WITH NORMAL DEVELOPMENT has been approved by her committee as satisfactory completion of the thesis or dissertation requirement for the degree of Master of Science.

Moran, Richard G. Ph. D. Major Advisor, School of Medicine

Miles, Michael F. M.D., Ph. D. School of Medicine

Taylor, Shirley M. Ph. D. School of Medicine

Windle, Jolene J. Ph. D. School of Medicine

Christie, Gail E. Ph. D. Chair, Molecular Biology and Genetics

Strauss, Jerome F. M.D., Ph.D., Dean, School of Medine

Boudinot, F. Douglas, Ph.D., Dean, Graduate School

12/08/08
Date

CREATION OF A MOUSE WITH A HUMANIZED *fpgs* GENE COMPATIBLE WITH NORMAL DEVELOPMENT

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

Linying Xie

Advisor: Richard G Moran, Ph.D.
Professor, Department of Pharmacology and Toxicology

Virginia Commonwealth University
Richmond, Virginia
December, 2008

Acknowledgement

I would like to express my deepest gratitude to my mentor, Dr. Richard G. Moran, for his great support and advice throughout my research. I truly appreciated his encouragement in my graduate studies and everything he has taught me. I would like to thank my committee members Drs. Michael Miles, Shirley Taylor and Jolene Windle for their great support, suggestions, and for taking the time to read my thesis over the thanksgiving break. Many thanks are given to Helen Zhang for her mouse expertise and technical assistance in generation of *fpgs*-P1 knockout mice. I am glad to have such a good friend like her. Many thanks are extended to all the Lab members, Alexandra Racanelli, Dr. Shane Kasten, Guoyan Gao, Erica Peterson, Scott Rothbart, Scott Lawrence, and Courtney Heyer and Lisa Sale for their friendship, encouragement and help. I am fortunate to have the opportunity to work with them.

Finally, the most of all, I would like to thank my husband Jinxing Yu, my son Jiade, Jonathan, and my daughter Jacqueline for their love, patients, and support that I needed to complete this study. I could not have done this study without all my friends and family.

Table of Contents

Page

List of Tables

Page

List of Figures

Page

Abbreviations

5-CH_3-H_4PteGlu	5-methyl-tetrahydrofolatemonoglutamate
5'-RACE	Rapid amplification of 5' complementary DNA ends
10-CHO-H_4PteGlu	10-fomyl-tetrahydrofolatemonoglutamate
ABC	ATP-binding cassette
ALL	acute lymphoblastic leukemia
cDNA	complementary DNA
CHO	Chinese hamster ovary
cM	centimorgans
DDATHF	5,10-Dideazatetrahydrofolic acid
DHF	dihydrofolate
DHFR	dihydrofolate reductase
dTMP	deoxythymidine monophosphate
dUMP	deoxyuridine monophosphate
ER	estrogen receptor
ES	embryonic stem cells
FDA	Food and Drug Administration
FPGS	folylpoly-γ-glutamate synthetase
FR	folate receptor
GANC	gancyclovir
GARTF	glycinamide ribonucleotide transformylase
GGH	γ-glutamyl hydrolase

GH	ɤ-glutamylhydrolase
H_4PteGlu	tetrahydrofolatemonoglutamate
HSV-*tk*	herpes simplex virus thymidine kinase
kb	kilobase
MCS	Multiple cloning site
MRPs	multidrug resistance proteins
MTX	methotrexate
Neo	neomycin phosphotransferase gene
nM	nanomolar
NTD	neural tube defect
P1	*fpgs* upstream promoter
P2	*fpgs* downstream promoter
PGK	a phosphoglycerate kinase I promoter
PCFT	proton-coupled folate transporter
qPCR	quantitative real time polymerase chain reaction
RFC	reduced folate carrier
RT	reverse transcription
SAM	*S*-adenosylmethionine
SNP	single nucleotide polymorphisms
TE	Tris-EDTA beffer
tet	tetracycline
TS	thymidylate synthase

Abstract:

Folylpoly-γ-glutamate synthetase (FPGS) catalyzes the formation of polyglutamate forms of the reduced folates and antifolates such as methotrexate (MTX) and pemetrexed; this allows the retention of folates and antifolate cancer drugs inside the cell. The enzyme activity of FPGS is essential for cell proliferation and survival. The mouse *fpgs* gene contains two promoters spaced 10 kb apart which are activated in a tissue-specific manner. The upstream promoter (P1) and exons A1a and A1b are used in some differentiated tissues, mainly liver and kidney, whereas the downstream promoter (P2) and exon 1 are used in rapidly dividing cells. In contrast, the human *fpgs* gene expresses virtually all transcripts from the downstream promoter. In order to more faithfully mimic human folate metabolism in the mouse, we have deleted the upstream promoter and the associated two small exons of *fpgs* in the mouse genome by homologous recombination. Homozygous deletion mice survive embryonic development, grow to adulthood, and reproduce through several generations, they appear to be normal. The results of Q-RT-PCR analysis on RNA from adult mouse liver of three different genotypes (A1aA1b +/+, +/-, -/-) indicated that deletion of P1 results in the release of promoter interference of P2, and activation of the downstream P2 promoter is increased by 3-5 fold. Interestingly, the total FPGS mRNA expression in KO mouse liver is 20-100 fold lower than in liver from wild-type mice. However when the FPGS activity was measured using an FPGS enzyme assay, the liver of knockout mice appeared to have only 2 fold lower enzyme activity than liver from wild-type mice. In conclusion, we have successfully generated a mouse which reflects human folate

metabolism much closer than seen in wild type mice. The FPGS- humanized mouse liver model would be an appropriate *in vivo* tool for the study of the antifolate drug toxicity and inhibition.

Chapter 1: Introduction

Folate Metabolism

Folate is the generic term for all chemical derivatives with the biological activity of folic acid. It is water soluble B vitamin that naturally occurs in meat, dark leafy vegetables, and some fruits, and exists as various reduced forms of folate with one or more glutamate moieties. Mammals do not have the ability to synthesize folates *de novo* as do prokaryotes and plants, and thus all mammalian species require preformed folates in their diet. Most dietary folates are polyglutamated and reduced derivatives, and are hydrolyzed to monoglutamate forms by ɤ-glutamylhydrolase (GH) in the gut prior to absorption across the intestinal mucosa. After crossing the mucosal cell barrier the folates are released into the portal circulation. 5-methyl-tetrahydrofolatemonoglutamate (5-CH_3-H_4PteGlu) is the major form of folate in plasma and is loosely associated with plasma albumin in circulation. Most of the monoglutamate folate is then taken up by the liver, where it is metabolized to polyglutamated derivatives by folypolyglutamate synthase (FPGS) and retained or released into blood and bile *(Bailey, 1994, Selhub, 1983).* Folate monoglutamates are hydrophilic, bivalent anions present in serum at tens of nanomolar (nM) concentrations. There are three folate transport systems used for cellular transport of folate into the cells; 1.) The folate binding proteins or folate receptors (FRα,β,ɤ). These are membrane-bound receptors that mediated the unidirectional transport of folate into the cell. These receptors have high binding affinity for folic acid and 5- CH_3-H_4PteGlu. Previous studies have shown that internalization of the monoglutamate form of folate bound to the folate receptor is through receptor-mediated endocytosis (*Antony, 1992. 1996).* 2.) Reduced folates and

most antifolates can be transported by the reduced folate carrier (hRFC1, SLC19A1), which is a bidirectional anionic exchanger with high affinity for reduced folates *(Sirotnak, 1999)*. The RFC is involved in intestinal and renal folate absorption. The RFC-mediated uptake of reduced folates is coupled with efflux of cytoplasmic organic phosphate (adenine nucleotide and thiamine mono- or pyrophosphate) *(Zhao, 2001)*. 3.) The proton-coupled folate transporter (PCFT/SLC46A1) is another route of folate uptake, which has high affinity for both folic acid and reduced folates and functions optimally at acidic pH. This is a unidirectional transport route of folates into the cell and is mediated by a protein mainly expressed in the apical membrane of the intestine and other epithelial cells *(Subramanian et al, 2008)*. At the cellular level, the uptake of folates mainly depends on FR, RFC and PCFT transporters (Fig. 1). Folate efflux from the cell is mediated by the RFC, perhaps by the PCFT, and also by a series of membrane transport proteins with ATP-binding cassette domains. ATP-binding cassette transporters couple the energy of ATP hydrolysis to the translocation of various anion and cation compounds, including folates and antifolates. Once inside the cell, folylpolyglutamate synthetase (FPGS) adds glutamate residues to the monoglutamate intrinsic to the folate structure. This increases its size and prevents its loss from the cell via the folate export pumps.

In eukaryotes, cellular folates are located in two compartments, the mitochondria and the cytoplasm. About 15-35% of folate resides in the mitochondria and the remainder is in the cytoplasm. Folate is also present in the nucleus at very low concentrations. In the cytoplasm, folate derivatives (5,10-CH_2-H_4PteGlu or 10-CHO-

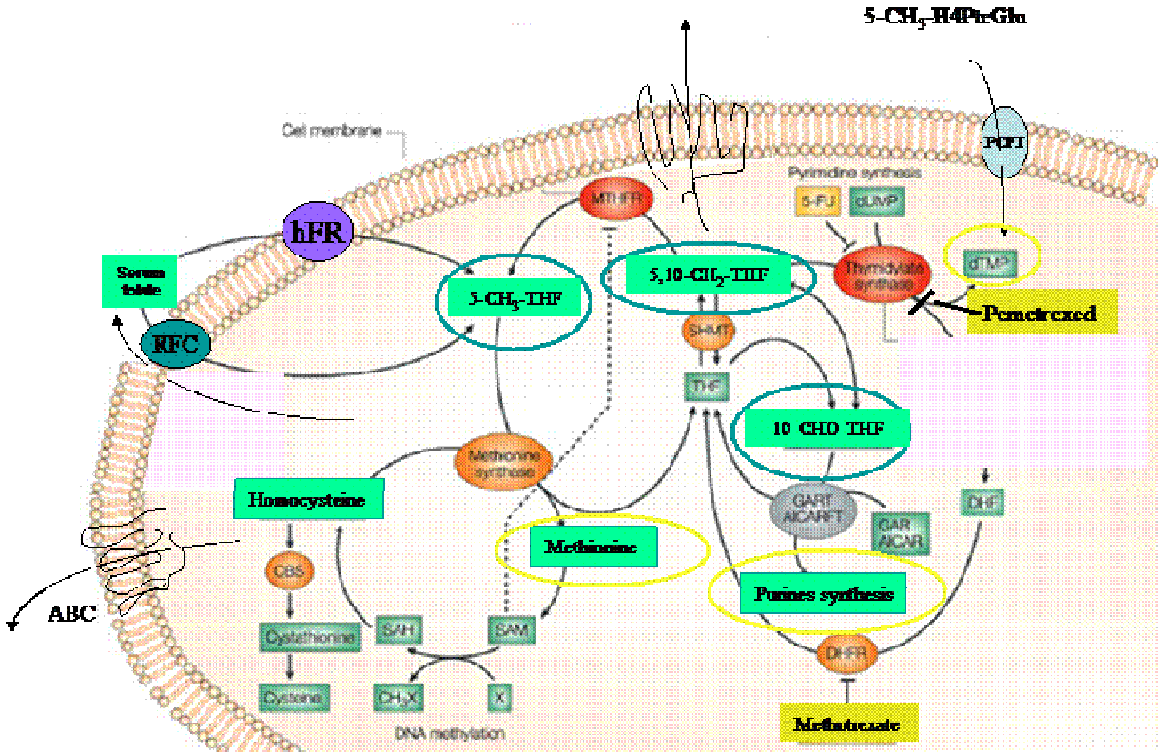


Fig 1. Folate metabolism and related pathways. Folate uptake through three folate transporters: RFC, FR, PCFT. Enzymes that require folates as a cofactor have been targets for chemotherapeutic agents. (*This figure was modified from Nat Rev Cancer 3(12):912-920, 2003*)

H_4PteGlu) function as donors of one-carbon units in a series of metabolic pathway involving *de novo* synthesis of purines and thymidylate, amino acid synthesis of methionine, serine and glycine, as well as methyl group metabolism (Fig. 1). In the mitochondria, folate is required for glycine metabolism and mitochondrial DNA-encoded protein synthesis (*Appling, 1991).* Therefore, folates play a key role in DNA synthesis and are essential for the survival of dividing cells and for growth. Folate deficiency may impair the *de novo* biosynthesis of purines and thymidylate and thereby disrupt DNA and RNA metabolism. Folate deficiency results in megaloblastic anemia, which is characterized by many large immature and dysfunctional red blood cells (megaloblasts) in the bone marrow, and also by multisegmented neutrophils. Folate deficiency also impairs the homocysteine remethylation reaction, methionine synthase. 5-CH_3-H_4PteGlu is a co-substrate required by methionine synthase when it converts homocysteine to methionine. Methionine is subsequently metabolized to *S*-adenosyl-L-methionine (SAM), which is the principal methyl donor in the methylation of macromolecules. Folate deficiency may lead to homocysteine accumulation, and the resulting homocysteinemia is involved in cardiovascular disease and, embryonic defects, particularly, neural tube defects (NTDs) and congenital heart defects. Folate deficiency has been shown to contribute to genome-wide DNA hypomethylation and has been claimed to result in activation of oncogenes and malignant transformation *(Wu, 1994 Kim, 2004).*

Folate plays an essential role in the synthesis of purines, thymidylate, and remethylation of homocysteine to methionine *(Anguera, 2006).* Enzymes that require

folates as a cofactor have been targets for chemotherapeutic agents, and several therapeutic drugs have been developed that are aimed at these enzymes. For example, dihydrofolate reductase (DHFR) plays a key role in the activation and regeneration of folates. DHFR is the target for several early antifolate drugs such as methotrexate (MTX) (Fig. 1). MTX is a folic acid analog that binds tightly to and inhibits DHFR, prevents the formation of tetrahydrofolate, and ultimately disrupts DNA synthesis and cell proliferation *(Takimoto, 1996)*. 5,10-CH_2-H_4PteGlu is an important coenzyme for the enzyme thymidylate synthase (TS), which transfers a one carbon unit from 5,10-CH_2- H_4PteGlu to deoxyuridine monophosphate (dUMP) to generate thymidine monophosphate (dTMP). Two antifolate drugs that are potent inhibitors of thymidylate synthase have been developed, ralitrexid and pemetrexed. Inhibition of thymidylate synthese by pemetrexed results in decreased thymidine triphosphate which is necessary for DNA synthesis. Pemetrexed is now FDA-approved for first-line therapy for non-small cell lung cancer. It was originally discovered in this laboratory in collaboration with Eli Lilly company and Professor Ted Taylor of Princeton University. It has been shown that polyglutamylation of these antifolates greatly enhances their cytotoxicity and that the enzyme folylpolyglutamate synthetase (FPGS) is responsible for the conversion of the drug to the more active form *(Synold, 1996)*.

Folate and antifolate polyglutamation

Folypoly-ɤ-glutamate synthase (FPGS) is essential for the survival of the mammalian cell and is a major determinant of cytotoxicity and selectivity for antifolates.

It is an ATP-dependent enzyme that catalyzes the addition of multiple glutamate molecules to the naturally occurring folates, and the classical anti-folate cancer drugs, such as methotrexate and pemetrexed. The highly anionic polyglutamate chain of cellular folates (four glutamate residues or more) severely reduces export through the cell membrane, which allows the retention and accumulation of folates and folate analogues pools inside tumor cells *(Moran, 1976.)*. Polyglutamyl derivatives of endogenous reduced folates have a much higher affinity than the corresponding unpolyglutamated forms for some, but not all, tetrahydrofolate cofactor-requiring enzymes *(Schirch, 1989.)*. For example, thymidylate synthase shows a 60% increase in substrate affinity with the diglutamate derivative of 5,10-CH_2-H_4PteGlu compared to the monoglutamate. This is also true for most antifolates; polyglutamated methotrexate and pemetrexed have much higher affinity for thymidylate synthase and glycinamide ribonucleotide formyltransferase than the parent drugs *(Shih,et al. 1998)*.

The FPGS Protein

A crystal structure of mammalian FPGS has not been solved. However, the structure of the MgATP complex of the enzyme from *Lactobacillus casei* has been extensively studied. The structural analysis of *Lactobacillus casei* FPGS (Fig. 2) reveals that the enzyme consists of two domains, joined by a six-residue linker. The N-terminal domain (a.a. 1-294) consists of a large central β-sheet composed of seven strands flanked by six α-helices and forms an ATP binding domain. The C-terminal domain (a.a. 295-300) comprises a central six-stranded β-sheet flanked by four α-

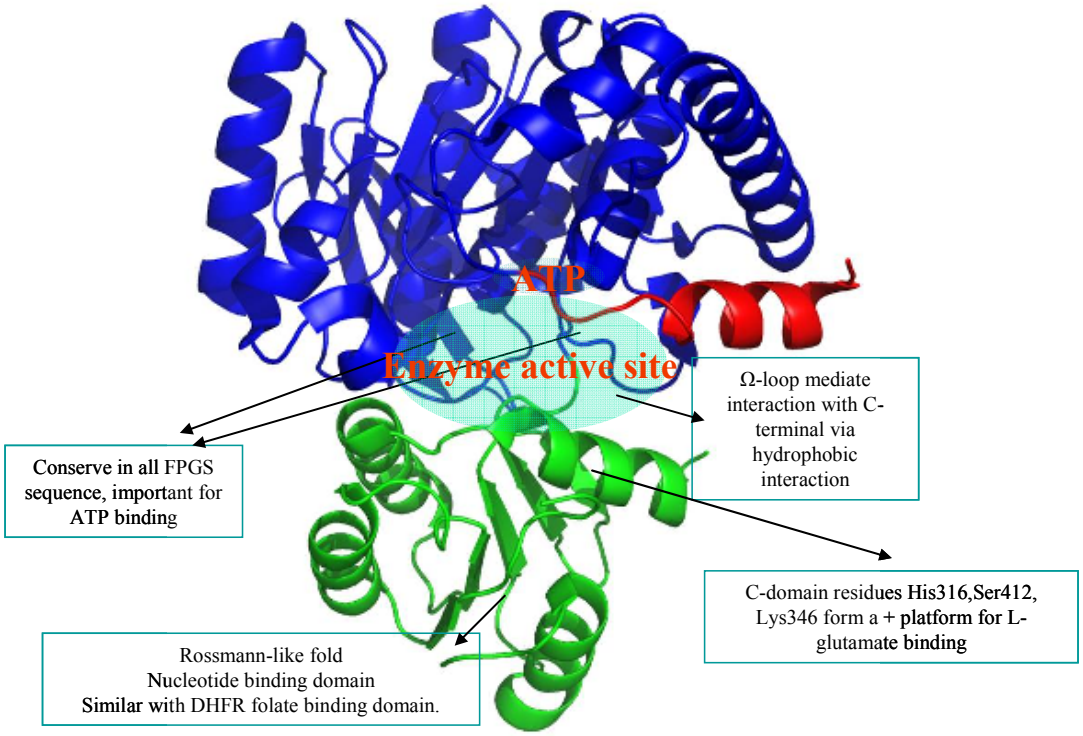


Figure 2. Ribbon diagram of L. casei FPGS

The N-domain is shown in blue, the C-domain in green and N-terminal in red. The binding domains were indicated in the boxes. *(This figure was modified from Sun, et al 2001. J. Mol. Biol.).*

helices with a Rossmann-like fold. This structure is composed of three or more parallel beta strands linked by two alpha helices in the topological order beta-alpha-beta-alpha-beta, which has structural similarly to dihydrofolate reductase, and contain a folate binding site. The active site of the enzyme is located in a large inter-domain cleft formed between the N- and C-domains, adjacent to two highly-conserved structural motifs, an ATP-binding P-loop (a.a. 46-50 GTN(K)GK GS) and a Ω-loop (a.a. 72-81) (Fig 2) *(Sun, 1998, 2001).* Binding of ATP and folate substrate triggers a large-scale rigid-body rotation of the C-terminal domain. This conformational change brings residues His 316, Ser412 and Lys346 of *L.casei* FPGS closer to the active site and provides a positively charged platform for the binding of L-glutamate. Mutation of these amino acid residues (H316A, S412A, Y414) have been shown to dramatically decrease the enzyme affinity for L-glutamate and an almost complete loss of FPGS activity. Substrate binding activate the FPGS and allow the enzyme to add a polyglutamate tail to folate *(Sun et al. 2001).*

Amino acid substitutions in the C-domain could influence the three dimensional structure of FPGS and disrupt folate and glutamic acid binding. The molecular docking studies of *L.casei* FPGS by Tan et al., indicated that there are two possible sites for binding folate substrates. One is for the folate tail and the other for the attacking glutamate. Both binding pockets were capable of binding the longer folate and involve the association of N-terminus domain residues R15-R85 and C-domain residue H316-Y414. It appears that N-terminus residues R15, R82 and S73 are important in complementing the growing glutamate chain *(Tan et al, 2005).* Mutagenesis

experiments suggest that the N-terminal domain plays a critical role at the enzyme active site. Deletion of the first eight amino acids of human FPGS and modification of Tyr3 had significant effect on FPGS activity *(Qi et al, 1999)*.

The amino acid sequences of two species of FPGS expressed in the mouse differ only by 18 amino acids at their N-terminus (Fig 2). The results of protein steady state kinetic studies have shown that both isoforms have the same catalytic activity for the production of polyglutamated folates (*Andreassi and Moran, 2002*). The isoform made by P2 transcription was more sensitive to feedback inhibition by long chain folypolyglutamates than the isoform made from P1. The enzyme produced in mouse liver (the P1 isoform) was not affected by 5,10-CHO-H_4PteGlu5, and H_4PteGlu$_5$ at the concentrations that inhibited the activity of P2 isoform. The steady-state kinetics analysis also showed some differences in substrate preference between the two mouse isoforms. For example, the average turnover number (Kcat) for P1 isoform using 5-CH_3-H_4PteGlu, 10-CHO-H_4-PteGlu, and H_4PteGlu as substrates was about double that of the P2 isoform (Kcat =6.5/min). But the P1 isoform also has higher Km (7.6 uM) for DDATHF, about 6 fold higher than the Km of the Ex1 isoform *(Andreassi and Moran. 2002)*. The difference in substrate preference and feedback inhibition between the two mouse isoforms may result from the differences of N-terminal amino acid sequence which influence the three-dimensional structure of the folate binding site in the active site. The decreased sensitivity of the FPGS P1 isoform to feedback inhibition by folypolyglutamate cofactors may be an explanation of a larger folate cofactor

accumulation in mouse liver and kidney than in rapidly proliferating tissues (*Andreassi and Moran. 2002)*

The *fpgs* gene - Control of the two *fpgs* promoters

FPGS enzyme activity is found in both cytosolic and mitochondrial compartments of mammalian cells; there are two consensus translational initiation codons (AUG start codons) on *fpgs* mRNA species, and these are used for the production of mitochondrial and cytosolic forms of FPGS (Fig 3). The translation products differ by the presence of an amino-terminal mitochondrial leader peptide. The first transcriptional initiation site generates a longer transcript encoding the mitochondrial isoform of the enzyme; the second start site generates a shorter transcript encoding only the cytosolic protein *(Freemantle and Moran, 1995).* The FPGS species made from the upstream ATG was capable of trafficking to the mitochondria and is responsible for the production of glycine (*Turner et al, 1999).*

The mouse *fpgs* gene is 21.5 kb long, contains 17 exons and 16 introns, and is located on chromosome region 2qB (Fig. 3). Two species of cDNA have been isolated from mouse liver and mouse leukemia cell L1210 cDNA libraries and by 5'-RACE; they share identical sequences from exon 2-15, but utilize different initial exons *(Roy, 1997; Turner et al, 1999).* In dividing normal cells and all tumor cells, *fpgs* transcription initiates from the downstream promoter (P2) and uses exon1. The alternative transcript initiating at P1 and including exons A1a and A1b was used in mouse liver and kidney. P1 is located about 10 kb upstream of P2, and both species of

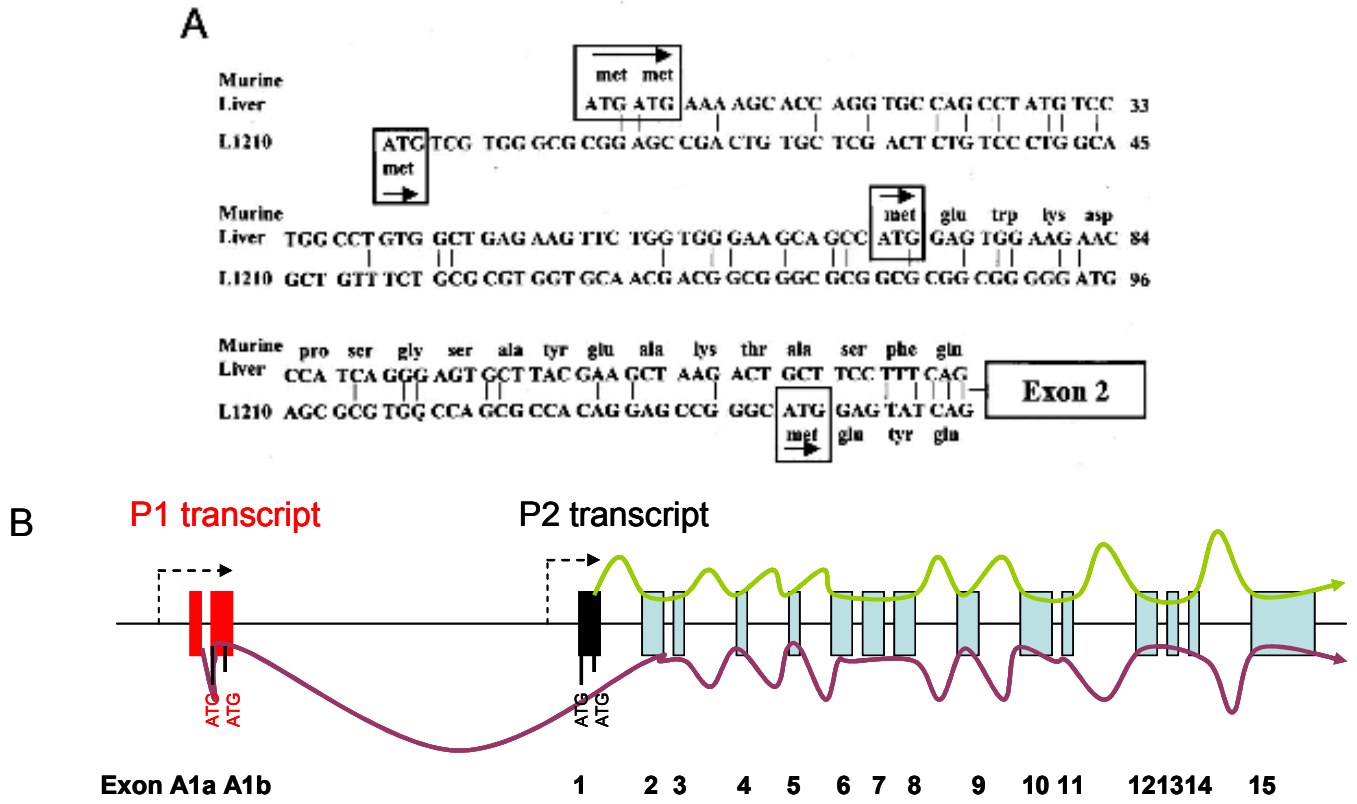


Fig. 3. Comparison of N-terminal sequences of the fpgs P1 transcript in mouse liver and P2 transcript in leukemic (L1210). *A*, sequence alignment of cDNA derived from P1 and P2. Both transcripts P1 and P1 contain identical exons sequences after exons 2. The NH2-terminal domain of the cytosolic forms of the P1 FPGS and the P2 FPGS differ in sequence of 18 amino acids, as indicated. *B*, a schematic of the alternative splicing patterns of P1 and P2 in the genome. Both transcripts contain two ATG start codons the sequence between the 2 ATG is in frame and free of stop codons. The FPGS species made from upstream ATG for either promoter was capable of trafficking to mitochondria. (*This figure was modified from Turner et al 1999 Cancer Research).*

mRNA encode functional protein (61kd - 62 kd) *(Andreassi and Moran,2002).* The human *fpgs* is located in chromosome region 9q34.1, and mRNA transcription initiates from downstream promoter P2 and uses exon 1 in both dividing and differentiated tissues. A recent study in this laboratory found that the tissue-specific expression of mouse *fpgs* gene is controlled by at least three mechanisms. One was linked to DNA methylation to control tissue-specific expression from the CpG-sparse upstream P1 promoter. The other is promoter occlusion preventing transcription from the downstream P2 promoter when P1 is active *(Racanelli et al, 2008).* The third mechanism is a poorly understood process involving tissue specific elongation of RNA polymerase complexes assembled at P2. Binding of regulatory tissue specific proteins to alternative promoters are often responsible for tissue-specific gene expression or developmental stage-specific expression *(Torik, 1996).* However, the role of tissue specific regulatory protein in control of transcriptional mutation from P1 or P2 is not yet known.

Homologous recombination as a tool to study gene function

The mouse provides a powerful system to study mammalian genetics and to produce animal models for human diseases. Gene targeting is the introduction of defined sequence changes at a specific locus in the genome by homologous recombination. This powerful technique has produced many advances in understanding the function of specific gene products in mammalian development and is one of the

most valuable and definitive tools in current biology. The major benefit of gene knockout technology is that it enables the analysis of the function of a protein produced from a specific gene *in vivo* in every cell type during development and into adulthood.

Gene targeting is achieved by transfecting embryonic stem (ES) cells isolated from mouse blastocysts with specific gene-targeting vectors followed by introduction of these recombinant ES cells into the mouse germline *in vivo*. Two type of targeting have been used: conventional and conditional.

Conventional gene targeting vectors

Typically, this type of vector contains two regions of homology to the targeted genomic locus placed on either side of a positive selection marker, such as the neomycin (Neo), hygromycin or puromycin-resistance genes. Outside of the homologous region, the herpes simplex virus thymidine kinase gene (HSV-tk) is most often included in the construct to allow negative selection. When the targeting vector is transfected into ES cells, a correct homologous recombination event will replace the target gene sequences irreversibly with the replacement construct sequences. However, genomic integration can occur not only by homologous recombination, but also by random insertion. To produce ES cell clones that contain a correct homologous recombination event, a combined positive and negative selection strategy is used. Positive selection for resistance to neomycin, for instance, will eliminate most cells that have not stably incorporated the construct into the genome. The negative selection for

exclusion of *HSV-tk* will enrich for homologous recombinant clones, since the sequences of *HSV-tk* gene in the targeting vector, which are located outside of the homology regions to the target gene, are lost during proper homologous recombination. In contrast, during random integration, all sequences in the vector tend to be retained because recombination usually occurs at the ends of the construct. In the presence of the *HSV-tk* gene, the cells are sensitive to acyclovir and its analog gancyclovir. *HSV-tk* allows the incorporation of gancyclovir into DNA, resulting in DNA chain termination and cell death. Thus, cells that contain randomly integrated targeting vector which carry the *HSV-tk* gene will be resistant to positive selection but will be gancyclovir sensitive, whereas cells that contain homologous recombinants will be resistant to both positive and negative drug selection *(McCarrick, et al, 1993).* A correct homologous recombination event has to be confirmed either by southern blot hybridization with a DNA probe that is designed to be able to distinguish the correct homologous recombination from the wild-type allele, usually by unique restriction enzyme digestion pattern, or by PCR analysis with a gene-specific primer located outside the region of homologous recombination and a targeting vector specific primer. The genomically modified pluripotent ES cells will then be introduced into mouse germ line by injecting the recombinant ES cells into recipient blastocysts *(Waldman, 1988. Deng, 1992).*

Disruption of a gene by conventional gene knockout in the mouse has advanced our understanding of the function of numerous gene products in mammalian development. The consequences of gene disruption in the germ line may have some

limitations. For example, the lack of a protein that serves one or more essential functions during embryogenesis can result in early embryonic lethality and hence blocks the experimental access to the study of its roles in later development or adulthood. For some genes, the phenotype affects multiple lineages, making it difficult to dissect individual functions. Furthermore, the effect of a germ-line mutation may also be compensated for during development by redundant gene products, thus preventing the appearance of an abnormal phenotype in the adult animal *(Marth 1996. Rajewsky 1996. Lobe 1998. Metzger 2001).*

Another limitation of conventional gene targeting also comes from the presence of a selection marker gene in the targeted locus. PGK–Neo is a hybrid gene, consisting of the phosphoglycerate kinase I (PGK) promoter driving the neomycin phosphotransferase gene (Neo). It is a widely used selection marker for homologous recombination in ES cells. Previous studies have suggested that targeted mutations that retain the PGK–Neo cassette may yield unexpected phenotypes in ‘‘knockout’’ mice due to the altered expression of neighboring genes within a locus *(Olson, 1996).* For instance, a study by Christine et al. (1996) has shown that the insertion of PGK–Neo into the granzyme B gene, the most 5′ gene in the granzyme B gene cluster, severely reduced the normal expression of multiple genes within the locus, even at distances greater than 100 kb from the mutation. The insertion of a PGK–Neo cassette into the β-globin locus control region also abrogates the expression of multiple globin genes downstream from the cassette *(Christine, 1996).* Several other examples are known for

which the positive selection gene disrupts normal function (Olson, 1996). For all these reasons, the selection marker cassettes should optimally be removed from the targeted locus after homologous recombination. This can be down by applying the *Cre/loxP* system to generate conditional gene knockouts *(Sauer, 1988)* with the correct placement of two *loxP* sites surrounding the drug resistance cassette (Fig. 6).

***Cre/loxP* conditional gene targeting vector**

Conventional gene targeting leads to inactivation or modification of a gene in all tissues of the body starting from the onset of development and continuing throughout the whole life span. The conditional gene targeting approach, in contrast, is to specifically delete a gene in a particular organ, cell type, or stage of development *(Smith, 2000)*. The control of gene targeting in a time-dependent manner allows the differentiation between effects of chronic versus acute depletion of a protein and also the analysis of multiple protein functions at different time points in development. The tissue-specific gene inactivation may define physiological roles of the gene product in a certain tissue without compromising other functions in the organism *(Muller, 1999)*. Conditional knockout mice have other benefits over the conventional type. 1) They often survive longer than conventional knockout mice. 2) Conditional knockout animal allow more precise study of tissue or cell type specificity of the function of a gene product. There are many conditional knockout models; however the most widely used is

that based on the specific excision of a genetic element using the *Cre-loxP* recombinase system.

Cre recombinase is a 35-kd enzyme isolated from bacteriophage P1 that acts as a site-specific DNA recombinase *(Kilby, 1993)*. The sequence recombined by *Cre* is a 34-bp sequence known as a *LoxP* site. A *LoxP* site consists of two 13 base pair (bp) inverted repeats (a *Cre* recognition site) separated by an 8bp spacer (Fig. 4). The asymmetrical 8bp spacer is responsible for the directionality of the *loxP* site. Binding of *Cre* recombinase to two *loxP* sites that are in the same orientation will result in excision of the intervening DNA, leaving one intact *loxP* site within the genome and the other on the excised circularized fragment. If the *loxP* sites are arranged in the opposite orientation relative to each other, *Cre* recombination will result in an inversion of the intervening sequence (Fig. 4) *(Martin J. 2000)*

To generate a conditional gene knockout using the *Cre/loxP* system, two lines of transgenic mice are required;

1) A mouse strain that carries the target gene or gene fragment flanked by two *loxP* sites in the same orientation and positioned in a way that it does not prevent normal gene activity. Thus, the conditional targeting vector is designed in the manner that *loxP* sites are inserted outside of coding regions, usually in an intron in a position that does not interfere with regulatory regions or exon splicing. The *Cre* recombination of these *loxP* sites will result in excision of a DNA fragment that inactivates the gene or results in an inactivate gene product. The neomycin or similar drug resistance gene is also

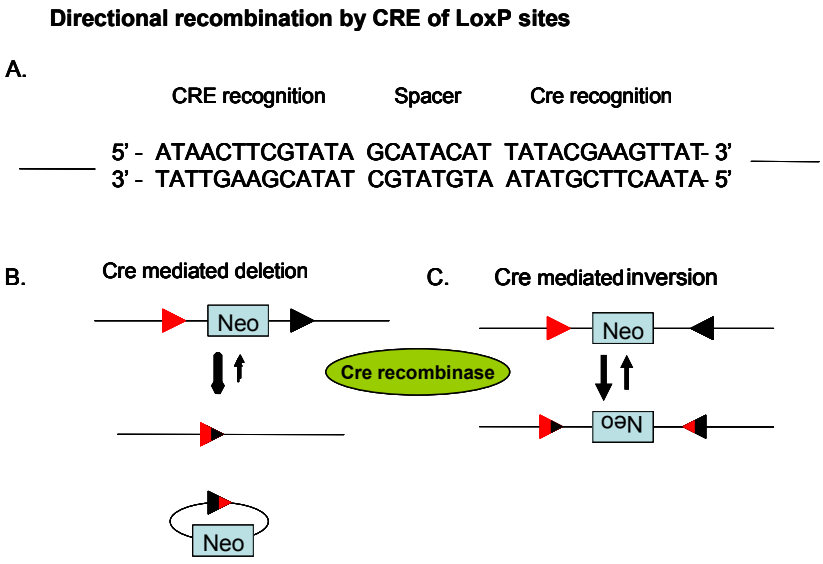


Fig 4. Recombination of *loxP* by Cre recombinase (A). 34bp sequences recognized by Cre recombinase. Two 13bp inverted *Cre* recognition sequences separated by an 8-bp spacer construct a *loxP* site. (B) and (C) Directional of l*oxP* sites and *Cre* recombination, ▶ indicates *loxP* site and orientation. Where sites are in the opposite orientation the DNA is inverted during the recombination event.

included in the vector for the selection of correctly targeted ES clones. However, the drug resistance gene included for the selection usually includes more than a kilobase (kb) of foreign DNA, which has the potential to influence the expression of the targeted gene prior to excision. To minimize this influence and remove the resistance gene, an additional *loxP* site is often inserted to flank the resistance gene. This will allow excision of the resistance gene from correctly targeted ES cell clones *in vitro* by transient expression of *Cre* recombinase prior to blastocysts microinjection. The final targeting vector would thus include three *loxP* sites: two positioned flanking the resistance gene, the third located to cause the desired deletion of the target gene when timed recombination occurs. For efficient and site-specific integration of the targeting vector into the ES cell genome, the targeting vector also contains two homologous regions identical to DNA sequences that flank the targeting region up- and down-stream of the negative selection marker. *HSV-TK* gene is also included in the construct at either end of the homologous recombination region to decrease the frequency of survival of random recombinants and, hence, enrich for homologous recombination ES cell clones.

2) To achieve spatially or temporally controlled somatic mutations, the second *Cre* transgenic mouse line is required that expresses *Cre* recombinase under the control of either a tissue-specific or a time-specific inducible promoter. This mouse is generated by standard oocyte injection techniques, or by knock-in gene targeting behind a suitable promoter. Up to now, many different transgenic mouse lines expressing tissue-specific *Cre* recombinase have been generated and are commercial available from the Jackson

Laboratory (Bar Harbor, ME). For instance, Postic and Magnuson (2000) have generated an albumin-*Cre* transgenic mice line in which the promoter and upstream enhancer of the rat albumin gene were used to direct the liver-specific expression of *Cre*. This modified version of *Cre* contains a rat albumin enhancer/promoter fragment, a nuclear localization signal sequence, and a translation start site optimized for eukaryotic cells *(Postic, 2000).* Metzger and Chambon have generated a mouse line expressing the tamoxifen-activatable *Cre* recombinase (*Cre-ER)* specifically in the proximal tubule cells of the kidney. In this system the *Cre* recombinase is encoded as part of a hybrid gene consisting of the *Cre* gene fused to a mutant form of the ligand binding domain of the estrogen receptor (ER), which is insensitive to the natural ligand (17β-estradiol) but, following binding of the synthetic estrogen antagonist tamoxifen, permits Cre–ER translocation from the cytoplasm to the nucleus, and allows the temporally-controlled inactivation of floxed target genes by excision (*Metzger, 2001*). Also, the tetracycline-dependent system developed by Gossen and Bujard activates transcription of the recombinase in a spatio-temporal manner. In this system, *Cre* expression is under the control of the *tet* system in which a transactivator tTA binds in the absence of tetracycline (*tet)* to an engineered *tet*-operator and activates the transcription of the *Cre* gene *(Gossen, 1992).* Mating of these mouse lines with a target mouse carrying a modified allele flanked with two *loxP* sites will generate a generation of animals with a tissue-specific gene deletion.

Genetic background of homologous recombinant mice

Genetically modified mice provide a powerful system for understanding the function of specific gene products in mammalian development and as models of human genetic diseases. However, transgenic and knockout mice are generated in a variety of inbred strain backgrounds. These background strains have been selected for ease and convenience of generating the transgenic or knockout mouse lines. For example, many targeted mouse strain have been generated by using ES cells derived from 129 sub-strain of mice, because the ES cells derived from this strain have the ability to contribute to the mouse germ line even after extensive manipulation in culture. C57BL/6-derived blastocysts are the most common hosts of ES cells lines for generation of germ-line transmitting chimeras, especially with the 129-derived ES cells. It is most often the case that 129-derived chimeras are bred with C57BL/6 females to allow monitoring germ-line transmission by coat-color markers and also to speed up breeding, since 129 strains are poor breeders *(Festing, 1996)*. This will generate a mouse with a mixed 129 x C57BL/6 genetic background. However, 129 mice are not derived from a single strain of mice; they are actually a diverse and complex family of sub-strains. The genetic background of 129 strains could interfere with phenotypic analysis of a targeted mouse. For example: 129 and C57BL/6 are very different genetically and behaviorally. Genetically differences were demonstrated by their polymorphic genetic markers. Behaviorally differences involve learning, memory,

anxiety, and pain response. C57BL/6 mice are not anxious, are more active and good learners, and perform well in a water maze. In contrast, 129 mice are more anxious, less active, poor learners and perform poorly in a water maze. Another example of background dependent phenotype variation is the postnatal phenotype of TGF-β1 knockout mice. On the mixed background of 129/sv X C57BL/6 approximately 50% of TGF-β1 KO mice survive to birth and develop an autoimmune-like inflammatory disorder that affects most organs (Marion. 1995). In constrast, only 1% of TGFβ1 knockout embryos in C57BL/6 genetic background live to birth and develop inflammatory disorder, the rest die before birth. This indicated that the phenotype of the embryonic lethality of TGFβ1 knockout is quite strain dependent. Donehower et al. monitored a large number of p53-deficient mice in two different genetic backgrounds 129/sv and C57BL/6 X 129/sv. They found that both *p53*+/- and *p53*-/- with 129/Sv genetic background mice had an accelerated tumorigenesis rate compared with their *p53*-deficient counterparts of mixed C57BL/6 × 129/Sv genetic background (*Donehower. 1995*). These studies also clearly demonstrated that the genetic differences between inbred strains can profoundly influence the knockout mouse phenotype. To reduce this influence, successfully targeted mice initially in the 129/sv background can be bred to a desired background strain for eight generations to obtain more than 95% pure genetic background. C57BL/6 is the most widely used mouse strain in a variety of research areas including cardiovascular biology, developmental biology, diabetes and obesity, genetics, immunology, and neurobiology research.

Many gene-targeted mouse strains are commonly backcrossed to C57BL/6 for more than ten generations to generated a congenic mouse strain. However, crossing two inbred stains (for example, 129sv X C57BL/6) results in a mixed genetic background. The F1 generation will be isogenic, because they share all wildtype genes and are heterozygous for the alleles inherited from the parental strains (one allele from 129 sv and the other from C57BL/6). Further intercrossing of F1, results in the generation of F2 and F3, where the alleles will segregate according to the chromosomal locations and cross-over frequencies during gametogenesis. Gene alleles on the same chromosome can be separated and go to different daughter cells. The further apart the alleles are on the chromosome, the greater the probability of allelic segregation. In general, the contribution of genetic material from parental strains to daughter strain is usually 50/50. Subsequent C57BL/6 backcrossing reduces the number of 129 genes that flank the mutated locus. However, even after 12 generations of backcrosses, at least 6.6 centimorgans (cM*) of the genetic material surrounding the targeted locus will be derived from the ES cell strain, usually 129/sv, rather than from the breeding strain (C57BL/6). The 6.6 cM of DNA corresponds to 13 million base pairs (Mbp) of DNA sequence and could have as many as 110 genes in it *(Schalkwyk, 2007).* Therefore, it would be dificult to determine whether phenotypic differences observed between homozygous deletion mice and wild-type control littermates are the result of the gene deletion or due to the effect of other genes closely linked to the mutated gene locus. Recently, the “speed congenics” method have been developed to aid in the generation of

congenic strains (Wong, 2002). This method uses polymorphic genetic markers to screen the progeny of each backcross generation for the "best male" for subsequent mating. Speed congenic breeding greatly accelerates the production of congenic mouse with maximal recipient genome contribution.

Statement of the problem addressed in this thesis

As explained in more detail above, there are two isoforms of the FPGS protein expressed in mouse tissues (Fig. 2). In mouse liver and kidney, the *fpgs* transcript is initiated from an upstream promoter (P1) and includes A1a and A1b exons in the mature mRNA. In all dividing tissues, the *fpgs* transcript is initiated from the downstream promoter (P2) and includes exon1 in the mature mRNA. To better understand the potential pathophysiological and pharmacological importance of the FPGS isoforms expressed in mouse liver and kidney, we decided to generate a knockout mouse specifically lacking P1 and exons A1a and A1b, but to still contain an otherwise normal *fpgs* locus. It was not known whether the deletion of the upstream promoter and exons (A1a and A1b) of *fpgs* would lead to early embryonic lethality. To overcome the possibility of an early lethality problem, a *Cre/loxP* conditional gene-knockout strategy was chosen for our targeting vector construction.

In this study we try to understand why the mouse has evolved to use the upstream P1 promoter instead of downstream P2 in liver and kidney, whether the two promoter system used in the mouse is essential for development, whether the deletion of

the upstream promoter and exons activate the down stream promoter, and whether the promoter interference observed in mouse FPGS collapses to the simpler system used in humans upon genetic deletion of the promoter used in mouse liver and kidney.

* cM, is a unit of recombination frequency for measuring genetic linkage. It is often used to indecate the distance along a chromosome. 1cM = 1% chance that a marker at one genetic locus on a chromosome will be separated from marker at a second locus due to crossing over in a single generation.

CHAPTER 2: Materials and Methods

***LoxP* cassette generation**

The *LoxP* cassette (Fig 7) was generated by annealing two equal amounts of complementary oligonucleotides (5ug) in a total volume of 90 ul in a 1.5 ml mirocentrifuge tube. The tube was placed in a glass beaker containing 300 ml of boiling water for 5 min; the sample was then allow to cool to 70^0C, and 10 ul of 10X SuRE/Cut buffer B (100 mM Tris HCl (pH8), 100 mM NaCl, 50 mM $MgCl_2$, and 10 mM 2-mercaptoethanol) were added and the reation mix was allowed to continue to cool to room temperature in the beaker containing 300 ml H_2O. The sample was phenol/chloroform extracted, EtOH precipitated, resuspended in H_2O and used for ligation. This *LoxP* cassette was designed with *Sal*I and *Xho*I ends for convenient insertion of the loxP site upstream of the P1 promoter in the targeting vector.

Gene targeting on ES cell

The PKO2lx FPGS-A1aA1b construct (Fig. 6) was linearized by digesting with *Not* I and an aliquot was run on an 0.8% agarose gel to confirm complete linearization. The linearized targeting vector was then phenol-chloroform extracted, ethanol precipitated and re-suspended in sterile TE (10 mM Tris, pH 7.5, 0.1 mM EDTA) at a concentration of 1 ug/ul. A total of 30 ug DNA was electroporated (250 v volts, 500 uF capacity) into 5 X 10^7 129/Sv ES cells by Mrs. Herju Zhang at the Transgenic core Facility. Forty-eight hours after electroporation, 150 ug/ml of the positive selection drug G418 and 2 mM of negative selection drug ganciclovir (GANC) was added. The culture

medium was changed every 2 days and selection continued for 12-14 days. Surviving clones (192) were picked into individual wells of two 96 well plates that contained primary mouse embryo fibroblasts (MEFs) feeder cells and expanded into three plates. One plate was cryopreserved in 10% DMSO and stored at -80 ^{o}C to provide early passage stocks of the clones. Genomic DNAs were extracted from the remaining two plates for genotyping to identify ES clones in which homologous recombination had occurred.

Genomic DNA isolation from ES cells in 96 well plates

Once the selected clones reach confluence, the plates were washed twice with 1x PBS (100 ul/well/96well plate) and lysed in 50 ul/well of Lysis Buffer (10mM Tris pH 7.5, 10mM EDTA pH 8.0, 10mM NaCl, 0.5% Sarcosyl, and 1 mg/ml Proteinase *K*) overnight at 60^{o}C in a humidified chamber. The next day, the genomic DNA was precipitated with 100 ul of cold absolute ethanol that contained 75mM NaCl. The plates were incubated at the room temperature for 15 min – 30 min or until the precipitated DNA was clearly visible. DNA was washed with 70% ETOH in the well and allow to air dry. Genomic DNA from each clone was re-suspended in 50 ul ddH2O, and 3 ul of genomic DNA from each clone was used for PCR analysis.

Identification of homologous recombinant ES clones

To identify the ES clones that contained the *fpgs* homologous recombinant allele, a long range PCR assay was designed using two sets of primers. 1) Primer 4309 and –Neo1B: Primer FPGS+4309 (CAG AAG TTT CTG CCA CCT CCC AAT) sequences were located at the 5' intron outside of the homologous recombination region contained in the targeting vector. The anti-sense primer Neo 1B (CAT TCG ACC ACC AAG CGA AAC ATC) contained *Neo* gene specific sequence. Together, these primers would amplify a 6.4kb DNA fragment if the genomic DNA of a cell contained the properly inserted homologous recombinant *fpgs* sequence. 2.) Primer-15459 and +ECB (Fig. 9a). The primer "+ECB" (GAT CGA TCC TAG AGA ATT CCT AGA CCG) was targeting vector-specific sequence, located over the unique restriction sites (*EcoR*I *Cla*I *BamH*I) downstream of the *loxP*3. The anti-sense primer -15459 (ATT GCA TAT CTG TGG AGG TCA GGG) was located in the third intron of *fpgs* gene sequence, outside of the sequence used as a 3' homologous region in the targeting vector. This set of primers would amplify a 5.2 kb DNA fragment from the targeted *fpgs* genomic locus.

The genomic DNA of ES cells were assessed by PCR using 20 ul Platinum HIFI Taq DNA polymerase High Fidelity Master Mix (Invitrogen) combined with 0.31 U *Pfu* Turbo DNA polymerase (Stratagene) and 2ul of PCR primers (75 ng of each primer was added in each reaction). A touch-down PCR program was used: initial denaturation at 94 ^{o}C for 4 min, followed by 5 cycles of denaturation at 94 ^{o}C for 30 second, annealing at a melting temperature of 60 ^{o}C for 30 sec, extension at 72 ^{o}C for 12

min, then followed by 30 cycles of denaturation at 94 °C for 30 sec., annealing at a 58°C for 30 sec, extension at 72°C for 6 min. followed by one final extension at 72°C for 15min. PCR results were analyzed by electrophoresis on 1% agarose gels in 1X TAE buffer at 30 v overnight (at room temperature).

Southern blot analysis

The positive clones identified by PCR analysis, were recovered from a well of cryopreserved cells on a 96 well plates and expanded into three 100 mm tissue culture dishes with MEF feeder cell. After the cells reached confluence, genomic DNA was extracted from these ES cells for DNA analysis in accordance with the manufacturer's (Gentra) protocol. After purification, 20 ug of genomic DNA were digested with *Sac*II (Fig. 11), fractionated by electrophoresis in 1% agarose gels with "1kb plus" size markers (Invitrogen), and transferred to a Hybond-N nylon membrane (Amersham). After blotting, membranes were UV cross-linked by using a Stratalinker (Stratagene). The 275 bp DNA probe was obtained by PCR from plasmid pGEM1Z-A1a with FPGS 9047 and FPGS 9322 (Table 1), primers which are located on *fpgs* gene upstream exon A1aA1b. DNA probe was randomly primed and labeled with [α-32P]-dCTP using a random primer DNA labeling system from Invitrogen. Hybridization was performed overnight at 65°C in Denhardts hybridization solution. Membranes were washed at 50°C (15 min each) in the solution that contained 0.1%SDS and 2xSSC. The washed

membrane was exposed to film (X-OMAT AR, Kodak) for 2 days at -80 ^{o}C using double intensifying screens (Fig8).

Cre Excision in Targeted ES cells

To eliminate the interference of *PGK-Neo*, the pCMV-cre plasmid (10ug) was introduced into the correct targeted ES clone 2D-6 ($5x10^6$) by electroporation. After two days cultured, the cell were seeded at 500-2000 cells per 100 mm dish to form colonies. About twenty-four clones were picked after two weeks of culture and each clone was seeded in duplicated plates. One set was grown in selection media with 180 ug /ml G418. The duplicate plate was used to expand the G418-sensitive (This was done by Mrs. Herju Zhang in the Transgenic Mouse Facility at VCU).

Mice

a. Mice housing

All experimental mice were born and raised in the animal facilities at the Virginia Commonwealth University. The A1aA1b knockout mice used in this study were C57BL/6 and 129Sv mixed genetic background or backcrossed for more than six generations in C57BL/6. C57BL/6 mice were obtained from Charles River Laboratories. Unless used for the folate deficiency dietary study, all mice were maintained on standard rodent chow mix. Mice were kept under a 12:12-h light-dark cycle in a temperature-controlled facility (25^{o}C) with free access to food and water. The mouse

reaches sexual maturity by six weeks of age; all the mice matings were set up after six weeks of age, and at the ratio of 1-2 female / male. Gestation normally lasts 19-21 days. The progeny of these matings were genotyped at weaning age (3-4 weeks of age) and the genotypes were determined by PCR of tail DNA (as described below). After weaning age, 4-6 mice were housed per cage. Heterozygous F1 females were intercrossed with heterozygous F1 males to generated F2 homozygous KO mice.

Growth curve analysis

Twenty four weaning age mice (4 weeks of age) obtained from each genotype group with C57BL/6 background (F6), (homozygous (-/-), heterozygous A1aA1b knockout (+/-) and wild type (+/+)), were used for a growth rate study. Each group contained twelve males and twelve females mice, the male or female mice were then divided into two sub-groups; one sub-group (six male or female mice) received normal amino acid-based diet, which contained 1300 ug folate per kg diet and 1% succinylsufathiozole (Purina, Test Diet, 5S4Q, irradiated ½" pellets). The other sub-group (six male or female mice) received a low folate diet, which contained 63 ug folate per kg diet and 1% succinylsufathiozole (Purina, Test Diet; 5T2L, irradiated ½" pellet). The body weight of mice were measured twice a week, start at weaning age. The mouse weights were recorded for ten weeks after initiation of feeding with each test diet.

Genotyping of offspring

Tail genomic DNA was prepared using an in-house protocol: Three week old weaning mice were marked by ear punching and a 0.5 cm tail tip was cut and digested in 500 ul of Tail blot solution (50mm Tris-HCl pH 8, 100 mM EDTA, 100 mM NaCl and 1% SDS) at 55^{o}C overnight. The mixture was extracted with phenol/chloroform (500 ul) once and the DNA was precipitated in 99-100% isopropanol and spool with a capillary pipette, washed in 75% EtOH and 100% EtOH, air dried for 5 minutes, and then re-hydrolyzed in 200 ul TE buffer (10 mM Tris-Cl, pH 7.5 1 mM EDTA).

Genotyping were performed by PCR analyses. Primers FPGS+9047 5'-TGC GAA CCT CTC AGG TAA ACA CTC-3' and FPGS-9322 5'-CCT GAA AGG AGG CAG TCT TAG CTT-3' were used to amplified a 275 bp DNA band from the wild type allele, Primer FP1 5'-TGA GTC AGT AGG CTC AGT GTG AGA-3' and -ECB 5'-GAT CGA TCC TAG AGA ATT CCT AGA CCG-3' were used to amplified a 175 bp DNA band from any knockout allele (Fig 15). The conditioned knockout allele was analyzed by *fpgs* intronic primer FPGS-9860 5'-TCG CTT GTC TAC CAC GCA CTG TAA-3'; and FPGS-10355 5'-TGT TCA GGC ATG TCA CAG AGT CCA-3'; This PCR will produce a 495bp fragment from the wild-type gene and a 576 bp fragment from the conditional knockout allele (Fig 10). PCR amplification was carried out in 20 ul Platinum HIFI Taq DNA polymerase High Fidelity Master Mix (Invitrogen) containing 75 ng of each primer and 80 ng of genomic DNA as template. After a 3 minute denaturation at 94^{o}C, the PCR was performed on the following program for 35

cycle: 35 sec 94°C; 35 sec 58 °C; 45 sec 72 °C, with a final extension step of 5 minute at 72 °C. The PCR products were analyzed on 1.2% agarose gels containing ethidium bromide (Fig 15).

RT-qPCR

Total RNA was isolated from liver, kidney, spleen, bone mallow and intestine of congenic litter mates using the TRIZOL Reagent (Invitrogen) according to the manufacturer's instructions; Mouse liver was first perfused through the portal vein with 5-10 ml cold 1x PBS, and about 150-200 mg of liver tissue was isolated and homogenized in 3ml TRIZOL Reagent with a Douncer homogenizer. SuperScript III First-Strand Synthesis System (Invitrogen) was used for first-strand cDNA synthesis; 5ug total RNA and oligo-dT was used in this reaction. About 25ug of cDNA was synthesized and ~ 1 ug or 1 ul cDNA were used for real time PCR.

Primers for mouse *fpgs* and *β-actin*, were designed using IDT (Integrated DNA Technologies) SciTools' Primer Quest and these primers were synthesized by IDT (Table 1a). All primers were then tested to ensure amplification of single bands with no primer-dimers. Primer sequences were as follow: for *fpgs* gene expression analysis: For β-actin gene expression; forward primer + actin sn 5'-GAC CCA GAT CAT GTT TGA GAC C-3', reverse primer: - actin asn, 5' - ATC AGA ATG CCT GTG GTA CGA C - 3'. Together these primers would produce a 150 bp PCR product. Primer A1aA1b-A was located in the sequences of exon A1a (5'-TGT GCA CAG GAT GGA AAT GCG

AAC C-3'), and its anti-sense primer –Ex3a (5'- TGT TTA GCC GGT TCA AGT CCT CCA – 3') was located in the *fpgs* cDNA sequences of exon 3. Together, these primers would amplify a 340bp cDNA fragment from P1 *fpgs* mRNA species. Primer Ex1 (5'- CTA GCT TGC GGT GGC ATT AAG ACT- 3') was located in the sequences of exon 1(Fig. 21), which was used in the transcript made from P2. Primer Ex1 and anti-sense primer -Ex3a can amplify a 343 bp cDNA fragment from P2 *fpgs* mRNA species. The real time PCR results were normalized to the mouse *actin* gene to control for the cDNA loading in the reaction.. The FPGS gene expression levels were quantified either using absolute external plasmid standards quantification or calculated as $2^{\Delta Ct}$ method using a reference gene (Bio-Rad Real-time PCR applications Guide).

PCR products using the primer pairs fpgs-A1aA1b-Ex3a, fpgs-Ex1-Ex3a, and mouse β-actin sn and asn were ligated into the TOPO-TA cloning vector (Invitrogen), and transformed into TOP10 competent cells (Invitrogen). Mini preps of isolated plasmid DNA were then prepared (Promega). Before use, the plasmid concentration was determined by measuring the absorbance at 260 nM (A_{260}) in a spectrophotometer (NanoDrop). To generated standard curves, the plasmids were serially (1:100) diluted in HPLC grade water (Fisher) over the appropriate concentration range (usually from 1ng/ul, 10^{-2}ng/ul, 10^{-4}ng/ul, 10^{-6}ng/ul, and 10^{-8}ng/ul).

Bio-Rad Mini Opticon Real-Time PCR System was used for DNA amplification and data collection. Reactions were carried out in triplicate in a 25ul volume containing 2 ul reverse-transcribed cDNAs, 12.5ul 2xSYBR Green Master Mix, and 75ng of each

forward and reverse primer. The reactions were cycled with a pre-incubation step of 15 minutes at 95°C followed by 40 cycles of 30 seconds at 94 °C, 30 seconds at 58 °C, and 30 seconds at 72 °C. Fluorescence data was collected at the end of the extension. The raw data were analyzed and normalized with respect to β-Actin expression. Ct values were exported into a Microsoft Excel worksheet for calculation of gene expression in accordance with the power (2- delta Ct) method (*Bio-Rad, Real-time PCR applications Guide*). The levels of gene expression in knockout mice were represented with respect to those in wild-type animals as knockout/wild-type ratios.

$$\text{Ratio} = \frac{\text{Average } 2^{\Delta Ct \text{ target (target-}Actin)}}{\text{Average } 2^{\Delta Ct \text{ ref (control-}Actin)}} \quad X\ 100$$

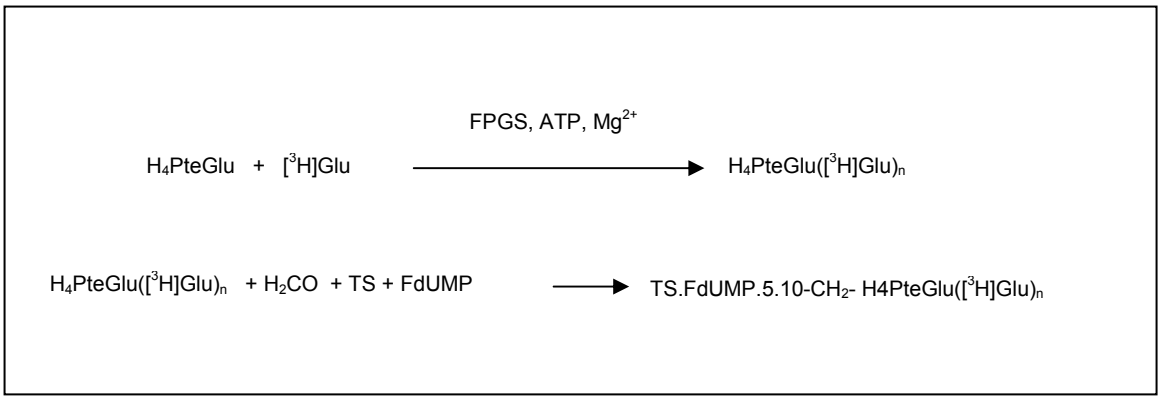
Table 1a – Primers for qPCR analysis of fpgs gene and housekeeping (GAPDH and β-Actin) gene

Target gene	primer name		Primer sequence	Product size
Fpgs -A1aA1b isoform:	A1aA1b-a	Sense	5'-TGT GCA CAG GAT GGA AAT GCG AAC C-3'	300 bp
	-Ex3a	Anti-sense	5'-TGT TTA GCC GGT TCA AGT CCT CCA-3'	
Fpgs – Ex1 isoform:	+Ex1	Sense	5'-CTA GCT TGC GGT GGC ATT AAG ACT-3'	300 bp
	-Ex3	Anti-sense	5'-TGT TTA GCC GGT TCA AGT CCT CCA-3'	
Fpgs	Ex3F	Sense	5'-TCATGTCACTGGGACCAAAGGGAA-3'	218 bp
	Ex6R	Anti-sense	5'-AGACATGGCTGTCGTCCTTGAACT-3'	
GAPDH	F177	Sense	5'-ATG ACA TCA AGA AGG TGG TG-3'	177 bp
	R177	Anti-sense	5'-CAT ACC AGG AAA TGA GCT TG-3'	
β -actin	sn	Sense	5'-GAC CCA GAT CAT GTT TGA GAC C-3',	155 bp
	asn,	Anti-sense	5'-ATC AGA ATG CCT GTG GTA CGA C - 3',	

FPGS microassay.

Animals were anesthetized with Isoflurane (Baxter) and sacrificed by cervical dislocation. Whole mouse liver were quickly perfused with PBS, then removed and whole liver was homogenized in 1mL lysis buffer [100mM Tris, pH 7.9 (at 25 °C), 2 mM dithiothreitol (DTT), 2 mM ATP, 5 mM MgCl2, 0.05% n-octylglucoside, 0.01mg/ml leupeptin, 0.01mg/ml Aprotinin, 5mM EDTA and 2 mM EGTA]; Tissues were disrupted using a dounce homogenizer for 20 strokes. Homogenate was centrifuged at 160000g (36000 rpm at Beckman optima-L-80 XP autra-centrifuge in SW 55 Ti Rotor) for 1 hr at 2^{o}C. The high-speed supernatant fraction 110 ul was passed through a Sephadex G-50 spin column (1 ml) to remove any glutamic acid present in the sample that would interfere with the assay and 2.5 ul of supernatant were used directly for FPGS assay. The first reaction was carried out in a total volume of 10 ul for crude cytosols: FPGS substrate solution 5 ul (800 mM Tris-HCl, pH 8.5 (25^{o}C) 40 mM $MgCl_2$, 120 mM KCl, 20 mM ATP 4 mM L-[^{3}H]glutamic acid, the specific activity: 1000cpm/pmole of glutamic acid, 20 x 10^{-6}M (6-S)-tetrahydrofolate and 0.5% 2-mercaptoethanol), 2.5 ul of crude sample and 2.5 ul buffer A (50 mM Tris-HCl pH 8.5, 10 mM $MgCl_2$, 2 mM DTT, 1 mM phenylmenthylsulfonyl fluoride) were mixed in 0.6 ml Eppendorf tubes on ice, the reaction was started by transferring the tubes to a 37^{o}C water bath, and the stander reaction was last for 15 min and ended by transfer to ice for 2 min., Any FPGS present in a sample mixture catalyzes the addition of [^{3}H]-glutamate to tetrahydrofolate to form -tetrahydropteroyl-diglutamate. The second reaction was

started by addition of 100 ul of fresh made thymidylate synthase solution (30 mM Na_2HPO_4 pH 7.2, 8 mM 2-mercaptoethanol, 10 mM formaldehyde, 2 uM FdUMP, 0.1 mg/ml bovine serum albumin, 0.3 uM thymidylate synthase) and 30 min incubation in 37^{o}C. The reaction was ended by transfer the tube to ice. During the second reaction, tetrahydrofolate was nonenzymatically converted to 5,10-methylenetetrahydrofolate and a covalent complex was formed among 5-10-methylenetetrahydrofolate, FdUMP , and thymidylate synthase. To ensure all the [^{3}H]-tetrahydropteroyldiglutamate in the reaction wereight complex containg the e bound in a covalently bound macromolecular complex, the excess FdUMP and thymidylate synthase were added. The high molecular [^{3}H]-tetrahydropteroyl-diglutamate was separated from free L-[^{3}H]glutamic acid by passing the reaction mixture through a Sephadex G-50 spun-column (1.1 ml). The columns were centrifuged at 400 rpm for 5 min and followed by 2 min. at 2400 rpm in Beckman J-6 centrifuge and the radioactive samples were collected in the mini-scintillation vials. 5 ml Scintillation liquid was added to each vial, and the tritium in the samples was determined by liquid scintillation counting. Enzyme activity was calculated by count per/min./ microgram of total protein *(Antonsson and Moran , 1990)*. Protein concentrations were determined by the Bradford dye-binding procedure (*Bradford,: 1976)*



CHAPTER 3: Results

In mouse tissues, two isoforms of the FPGS protein are expressed. One isoform is expressed in all rapidly dividing cells, and transcription that produces this isoform is initiated from P2 and the mature transcript includes exon1. The other isoform is only expressed in a few differentiated tissues (liver and kidney); this transcript uses the upstream promoter (P1) and the initial sequence is from exons A1a and A1b. In the adult human, almost all *fpgs* transcripts in both dividing and differentiated tissues are initiated from Exon1 and use only the P1 promoter. The role of the *fpgs* transcript expressed in mouse liver has not yet been studied during development. Is the two promoter system used in the mouse essential for development? Will the deletion of the upstream promoter activate or up-regulate the FPGS P2 transcription in mouse liver and kidney? Will the promoter interference used in mouse *fpgs* collapse to the simpler system used in humans upon genetic deletion of the promoter used in mouse liver and kidney? What are the differences in folate metabolism in these mutant mice? To better understand the evolution of two mouse FPGS species, we decided to knock out the P1 promoter and Exons A1a and A1b in the mouse genome.

Gene targeting vector construction

There are two types of gene targeting: conventional and conditional. Conventional gene targeting leads to the deletion or mutation of a gene in all tissues of the mouse from embryonic development throughout adulthood. Conditional gene targeting is aimed at controlling gene targeting in a time- and tissue-dependent manner.

When the global removal of the gene of interest results in embryonic lethality, the conditional knockout will be the method of choice for studying the gene function during later stages of development (*Muller, 1999)*. It was not known whether the deletion of the upstream promoter and exons A1a and A1b of *fpgs* would lead to early embryonic lethality. What is the function of the P1-derived FPGS isoform in late development and in the adult? To overcome the possible early lethality problem, a *loxP/Cre* conditional gene-knockout strategy was chosen for our targeting vector construction.

Selection of homologous recombination regions for targeting vector construction

Two contiguous *Hind*III genomic DNA fragments of *fpgs* from mouse 129/sv BAC library were previously cloned into a modified p GEM4 (Promega) vector (pGEM-LZ) by a former graduate student in this laboratory, Fiona Turner *(Turner et al. 1999).* The plasmid pGEM-LZ-A1a contained an 8.5 kb *Hind*III genomic fragment, which includes the sequences of the P1 promoter and exons A1a and A1b (Fig. 5). The other plasmid, pGEMLZ-Ex1, contains a 7.1 kb *Hind*III genomic fragment, which included the P2 promoter and Exon 1. The sequences of these two clones were analyzed by extensive restriction mapping during Dr. Turner's dissertation, and more recently was compared to the NCBI mouse genome resource sequence of the *fpgs* gene. The location of the experimentally determined restriction endonuclease sites relative to the positions of intron and exon sequences on the clones were identical to the NCBI mouse *fpgs* sequence (Fig. 5). The results show that two *Hind*III sites were located

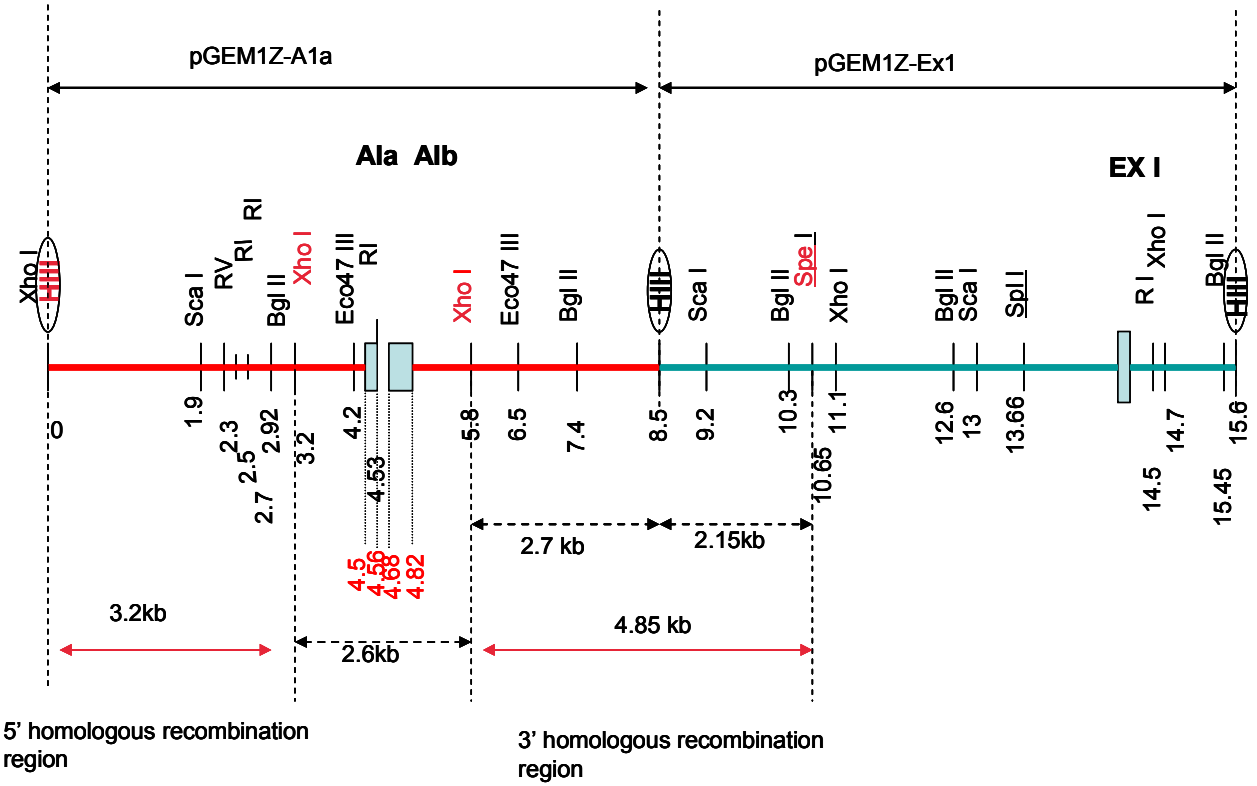


Fig 5. The Mapping of restriction enzyme sites on mouse *fpgs* gDNA sequences surrounding exons A1a and A1b and exon 1. A *Hind* III-*Xho*I fragment was selected as the 5' homologous recombination region, and the ajacent *Xho*I – *Spe*I fragment was selected as the 3' homologous recombination region.

sequentially in the mouse genome; these two clones (pGEM1Z-A1a and pGEM1Z-Ex1) were used as the backbone of the targeting construct.

The total length of the homologous DNA regions on the targeting vector is one of the most important parameters that affect the targeting frequency at a particular genetic locus. It has been shown that increasing the length of total homology dramatically increased the absolute targeting frequence *(Deng,C and Capecchi, M.R. 1992).* Therefore, 3.2 kb of *Hind*III -*Xho*1 region present in the 5' *fpgs* gene intron upstream of the P1 was selected as a 5' homologous recombination region. The 4.85 kb of *Xho*1-*Spe*1 region, present in the intron 3' to exon A1b, was selected as a 3' homologous recombination region (Fig. 5).

Construction of the targeting vector

The targeting vector pKO Scrambler NTKV (Stratagene) (Fig. 6) contains a drug resistance gene and the neomycin phosphotransferase gene (*Neo*) driven by the phosphoglycerate kinase gene (*PGK*) promoter. This marker would allow the positive selection of cells that contain the targeting construct integrated in the cellular DNA. On both sides of a *Neo* selection marker reside two multiple enzyme cloning sites (MCS) for the cloning of the homologous recombination regions. This targeting vector also contains the herpes simplex virus thymidine kinase gene (HSV-*tk*) - a negative selection marker; retention of the *HSV-tk* gene in a mammalian cell renders that cell sensitive to the guanosine analog, gangcyclovir (GANC). Some studies have shown that retention of

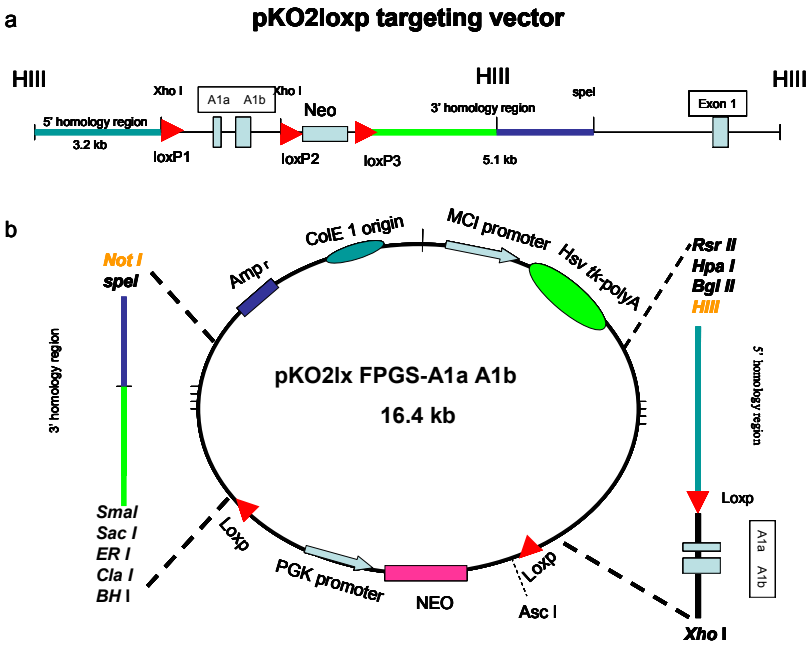


Fig. 6 The construction of the pKO2lxFPGS-A1aA1b conditional knockout targeting vector. (a) Schematic diagram showing the DNA enzyme restriction sites and relative location of *loxP* sites and homologous recombination regions after planned insertion into the mouse genome. The targeting vector was designed by placing a *PGK-Neo* selectable marker, flanked by *loxP* sites ~1kb downstream of coding exons A1a and A1b. The upstream promoter and coding exons A1a and A1b of fpgs were also flanked with *loxP* sites. (b) Plasmid map of the targeting vector following ligation of two homologous recombination regions and floxed A1aA1b region into two MCSs.

the 1.6kb PGK-Neo cassette in targeted loci can cause unexpected phenotypes in knockout mice due to disruption of expression of neighboring genes within a locus *(Christine, 1996)*. To permit the removal of the resistance gene at any point in this experiment, additional loxP sites were inserted, flanking the *Neo* resistance gene. Thus, the *fpgs* targeting vector was designed for the deletion of the P1 promoter and exons A1a and A1b in the *fpgs* gene in the mouse genome. The P1 promoter region, exons A1a, A1b and the 5' homologous recombination region recombination regions were cloned into the MCS downstream of the floxed *Neo* gene, and the 3' homologous recombination region was cloned into the other MCS. In order to permit the deletion of the upstream promoter and exons A1a and A1b in a time- and tissue-dependent manner, another *loxP* site was inserted upstream, between the promoter and the 5' homologous recombination region (Fig. 6). These procedures are described briefly below and the schematic diagram was shown in appendix A.

Since there was a paucity of useful restriction enzyme sites in the pKO Scrambler vector, a smaller vector, the plasmid pBluescript II SK$^-$, was used as an intermediate vector to introduce the restriction enzyme sites needed for the insertion of a *loxP* cassette at the desired site in the homologous region. The 5' homology arm was generated as follows:

(1) The 3.2kb *HindIII* - *XhoI* DNA fragment to be used as an upstream recombination region was excised from genomic clone pGEMLZ-A1a plasmid, gel-purified and

ligated into the *HindIII* and *XhoI* sites of the pBluscript II SK$^-$ vector (Stratagene), resulting in an intermediate plasmid pSK-A1aHIII-Xho.

(2) Two complementary DNA oligonucleotides (SalI-loxP-Xho-T 5'-TCG ACA TAA CTT CGT ATA GCA TAC ATT ATA CGA ATT GAT C-3' and SalI-loxP-Xho-B 5'-TCG AGA TAA CTT CGT ATA ATG TAT GCT ATA CGA AGT TAT G-3') Fig. 7) were designed to generate a loxP site cassette with *SalI* and *XhoI* ends, the oligonucleotides were annealed and slowly cooled to room temperature to allow the formation of double strand DNA fragment. These procedures are diagrammed in Fig. 7. The double-stranded cassette was then inserted downstream of the 5' recombination region of plasmid pSK-A1a HIII-XhoI from step (1). Ligation of *SalI* (at the end of *loxP* cassette) to the *Xho*I site (at the end of 5' recombination region) abolished the internal *Xho*I site at the end of the 5' recombination region, and left a unique *XhoI* site at the end of the *loxP* cassette.

(3) Plasmid pGEMLZ-A1a was cut with *Xho*I and the 2.6 kb fragment containing the exon A1a and A1b sequences was gel-purified (Fig. 5). The 3.3kb fragment representing the 5' recombination region with a *loxP* site was then cut out of pSK-A1aHIII-Xho-Loxp with *Hind*III and *Xho* I and gel-purified. The 2.6 kb pGEMLZ-A1aXHoI fragment and the 3.3 kb fragment representing the 5'-recombination region and loxP site were than moved into the *Hind*III – *Xho*I sites of the pKO Scrambler NTKV 1901 targeting vector by three piece ligation to form a intermediate targeting construct pKO2LX -A1a.

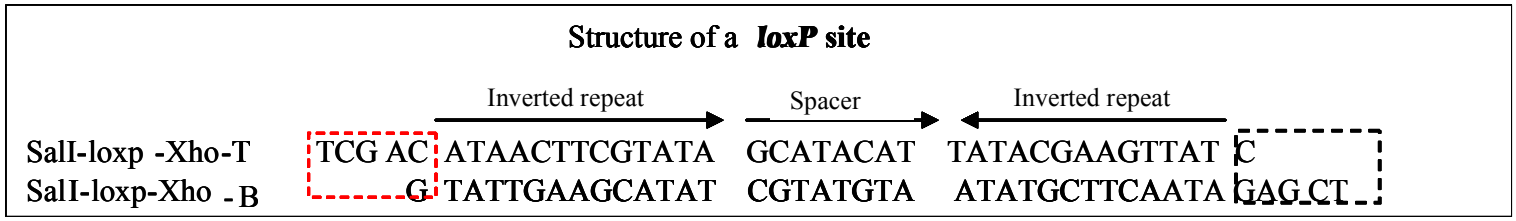


Fig 7. Sequence of a *loxP* site that was generated by two primers annealing with *Sal*I and *Xho*I sticky ends. The *lox*P sequence consists of two 13-bp inverted repeats flanking an 8-bp spacer region that determines the orientation of the site. The sequence in a red box is a *Sal*I recognition site, that in the black box is a *Xho*I recognition site.

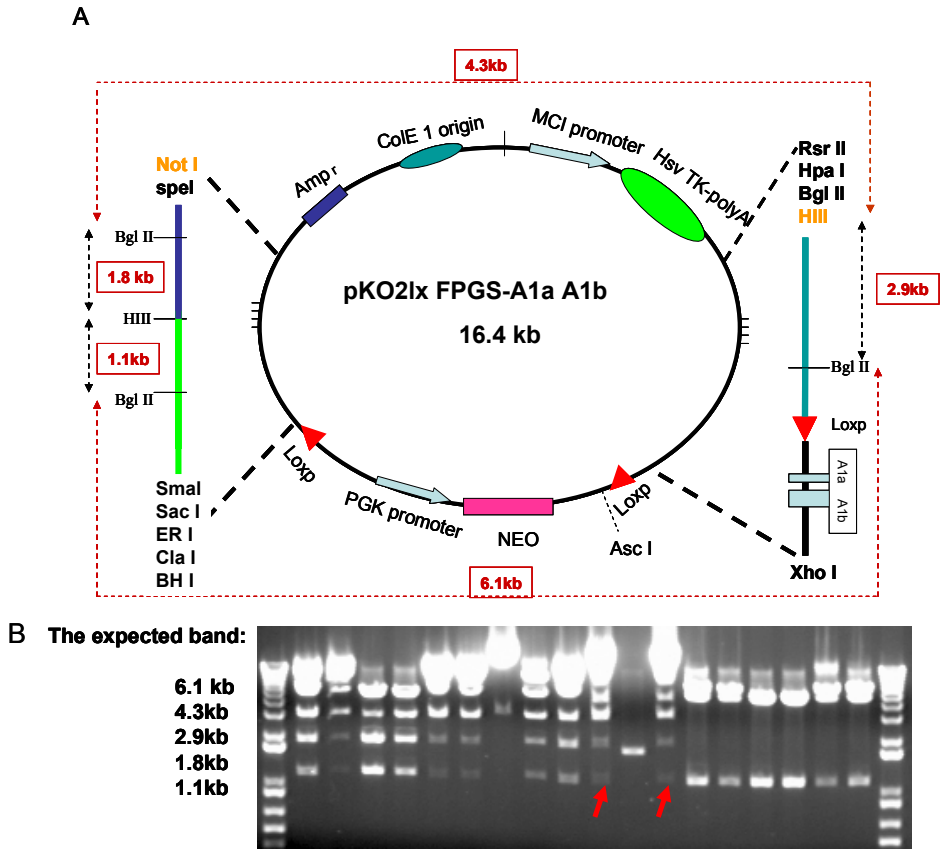


Fig 8. A. Schematic diagram showing the DNA enzyme restriction sites *Hind*III and *Bgl*II and the relative location of the *loxP* sites and homologous recombination regions in the targeting vector. B. Constructed Targeting vector was transformed into Stbl2 competent cells (Invitrogen). Plasmid DNA from overnight bacterial cultures of recombinant clones was purified, and analyzed by double digestion with *Hind* III and *Bgl* II and loaded on a 1% agarose gel. Red arrows indicate the clones that were chosen for large DNA amplification, sequence analysis and ES cell transfection.

A 4.85kb *Xho*I – *Spe*I fragment, located in the intron region downstream of exon A1b (Fig.5), was chosen as the 3' homologous recombination region. The 3' homology arm was generated as follows:

(1) the 2.7kb *Xho*I – *Hind*III 3'-terminal fragment of plasmid pGEMLZ-A1a was excised, gel-purified and cloned into the pBluescript II SK- vector (Stratagen) to form a intermediate plasmid pSK-XhoI-HindIII.

(2) The length of the 3'arm was increased by ligating a 2.2 kb *Hind*III-*Spe*I fragment from plasmid pGEMLZ-Ex1 that was contiguous with the 2.7kb *Xho*I –*Hind*III fragment in mouse genomic DNA (Fig. 5) into a resultant intermediate plasmid pSK-XhoI-HIII to yield a 4.85 kb 3' homology recombination region. This 4.85 kb fragment was excised from pBluescrip with *Not*I *and Xho*I to introduce a unique *Not*I site at the 3' end of the downstream recombination region. The 4.85 kb fragment with *Xho*I and *Not*I ends was then introduced into the intermediate targeting construct pKO2LX -A1a which had been digested with *Sal*I and *Not*I, resulting in plasmid pKO2LX-FPGS-A1aA1b targeting vector (16.2 kb). The final targeting vector pKO2LX-FPGS-A1aA1ab was then transformed into Stbl 2 chemically competent cells (Invitrogen). Plasmid DNA from overnight bacterial cultures of recombinant clones was purified, and analyzed with restriction enzymes *HindIII* and *BglII* digestion (Fig. 8). Clone number 10 and 12 were chosen for large scale DNA amplification. The plasmids DNA were purified by

Endofree Plasmid Maxi Kit (QIAGEN). The plasmid DNA sequence was again confirmed by sequence analysis and the sequence-confirmed targeting vector was transfected into mouse 129sv ES cells. This procedure for the introduction of the 3'-recombination region is diagrammed in Fig. 6.

Identification of homologous integration in ES clones

Mouse 129/sv derived ES cells were chosen for gene targeting, based on the ability of ES cells from this stain to support germline transmission even after extensive manipulation of the ES cells in culture (*Simpon, 1997)*. A total of 30 µg of *Not*I-linearized pKO2lx FPGS-A1aA1b were electroporated into 129/sv ES cell (5×10^7 cells), and the ES cells were plated onto a feeder layer of primary embryonic fibroblasts. After 10 days growth, G418 resistant clones were isolated and plated into two duplicate 96 well plates by Mrs. Herju Zhang in the Transgenic Core Facility at VCU.

ES clones that survive positive – negative selection are not all the desired homologous recombinants. They may have survived drug selection because the targeting vector was randomly integrated into the genome of the recipient ES cells with the retention of the positive selection cassette and loss or mutation of the *tk* gene. However, in the recipient ES cells where homologous recombination between the targeting vector and the endogenous gene has occurred, the positive selection cassette will be integrated into the target locus while the *tk* gene is lost and these cells will survive both positive G418 and negative GANC selection (Fig. 9a).

The clones harboring the desired homologous recombination event were then identified by PCR analysis. Two sets of PCR primers were designed for this; 1.) Primers +4309 and Neo 1B (Fig. 9a); The primer +4309 sequences were located at the 5' intron outside of the homologous recombination region contained in the targeting vector. The anti-sense primer Neo 1B contained *Neo* gene specific sequence. Together, these primers would amplify a 6.4kb DNA fragment if the genomic DNA of a cell contained the properly inserted homologous recombinant *fpgs* sequence. 2.) Primer-15459 and +ECB (Fig. 9a). The primer +ECB was targeting vector-specific sequence, located over the unique restriction sites (*EcoR*I *Cla*I *BamH*I) downstream of the *loxP*3. The anti-sense primer -15459 was located in the third intron of *fpgs* gene sequence, outside of the sequence used as a 3' homologous region in the targeting vector. This set of primers would amplify a 5.2 kb DNA fragment from the targeted *fpgs* genomic locus. The genomic DNA of each G418 resistant clone was isolated from each well of 96 well plates. PCR analysis was performed with primers +4309 and Neo1B, to identify clones in which a correct 5' homologous recombination event had occurred (Fig. 9A). Out of 192 clones, nine positive clones were identified by this first round of PCR screening (Fig. 9B). These positive clones were further analyzed by a second round of PCR using primer -15459 and +ECB to identify any of the clones with the correct 3' homologous recombination event. Any clones containing a correct integration event should give a 5.2 kb PCR product (Fig. 9A) in this second round of PCR. Five clones were clearly

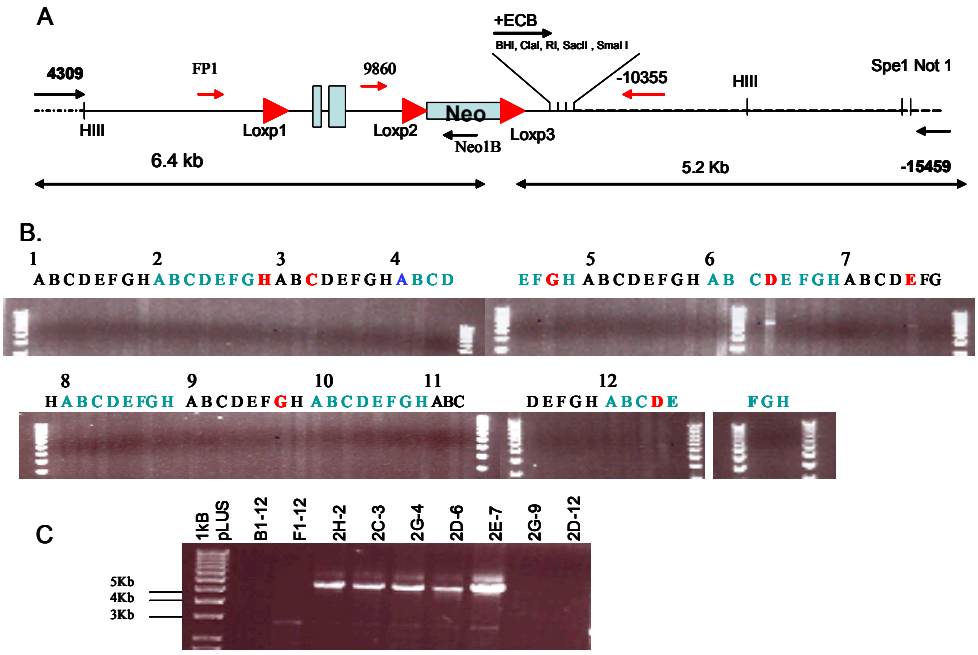


Fig. 9. PCR analysis of homologous recombinant targeted ES cell clones. (A) Schematic diagram showing the location of the primer pairs used in the PCR analysis (black arrow. (B) Results of the PCR analysis using the primers 4309/Neo1B on one of the 96 well plates. Alphabet number indicated the column number on the 94 well plates; the alphabet letter indicated the well in each column. The positive clones were indicated with red letter. (c) PCR analysis using the primers +ECB/-10355 on the clones that were identified positive from first round PCR analysis by primers 4309/Neo1B.

identified to be inserted correctly by this second round of PCR (Fig. 9C). The clones that were positive after the PCR screenings were subject to detailed sequence analysis.

Five independent recombinant ES cell clones (2.6% of the 192 total G418 resistance ES colonies originally isolated) were identified as potential homologous recombinant candidates for further analysis. To ensure homologous recombination at the FPGS locus, the PCR products were then gel purified and subjected to sequencing analysis across the loxP1, loxP2, and loxP3 regions using *fpgs* primers noted in Fig. 10 Sequencing was conducted by the VCU Massey Cancer Center DNA Core facility.

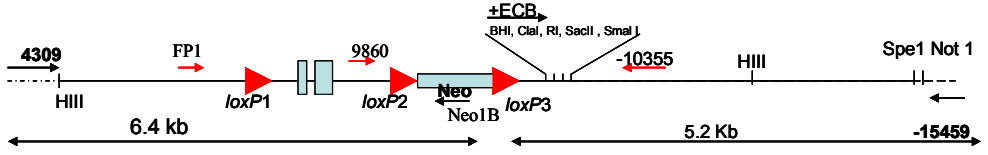
Unintended homologous integration into the exon A1a/A1b region

In our conditional gene targeting, the critical exons were flanked with two *loxP* sites in the targeting vector for deletion at a later step. This design was used to avoid the potential lethality of genomic deletion during the development of the embryo. However, it also carries a risk, since it results in the inclusion of an extra potential homologous recombination region (Fig. 6). Hence, three *fpgs* intronic primers (FP1, 9860 and -10355) that represented sequence located next to the loxP sites were designed for sequence analysis of the presence or absence of the three loxP sites, and this analysis was the final step to confirm the correct homologous recombination. Genomic DNAs were purified from the potential positive targeted ES cell clones, and the *fpgs* targeted alleles were PCR amplified with the primers FPGS+4309 and Neo 1B, which were used in the identification of 5' homologous recombination event. The PCR product of 6.4 kb

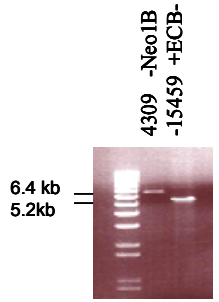
a

PCR primer sequences for homologous recombination analysis in ES cell				
Forward primer (5' to 3')		Reverse primer (5' to 3')		product size (bp)
FPGS+4309	CAGAAGTTTCTGCCACCTCCCAAT	**NEO 1B**	CATTCGACCACCAAGCGAAACATC	6.4kb
+ECB	GATCGATCCTAGAGAATTCCTAGACCG	**FPGS-15459**	ATTGCATATCTGTGGAGGTCAGGG	5.2kb
Sequencing primers for loxPs analysis				
FP1	TGAGTCAGTAGGCTCAGTGTGAGA	loxP1		
FPGS-9860	TCGCTTGTCTACCACGCACTGTAA	loxP2		
FPGS-10355	TGTTCAGGCATGTCACAGAGTCCA	loxP3		

b



c



d

ES CLONE	FP1	9860	10355
2H-2	WT seq	-	-
2C-3	*loxP*$_1$	*loxP*$_2$	ECB seq and *loxP*3
2G-4	WT seq	-	ECB seq and *loxP*3
2D-6	*loxP*$_1$	*loxP*$_2$	ECB seq and *loxP*3
2E-7	WT seq	*loxP*$_2$	ECB seq and *loxP*3
2G-9	*loxP*$_1$	-	ECB seq and *loxP*3

Fig. 10. Analysis of prospective targeted ES cell clones. (a). The DNA sequence of PCR primers and the sequencing primers is shown in the table. (b). Schematic diagram showing the location of the primer pairs used in the PCR analysis (black arrow) and the primers used in the sequencing analysis of three *loxP*s (red arrow). (c). PCR re-analysis on ES clone 2D-6 with primer set 4309/Neo1B and +ECB/-10355 (illustrated in a and b), the PCR products were on a 0.8% agarose gel. (d). The results of sequence analysis that crossed the 3 loxP regions. Primers used for sequencing were FP1, 9860 and -1035. The ES clones harboring all three *loxP* sites were indicated in red.

were detected and purified from an agarose gel, and this DNA fragment was used for sequencing analysis of *loxP*1 and *loxP*2 sites. Primer FP1 was used to sequence across the *loxP*1 region and primer 9860 was used to sequence across the loxP2 region. Primer FPGS-15459 and +ECB, which were used in the identification of the 3' homologous recombination events, were used to amplify a 5.2kb DNA band from the same genomic DNA isolated from positive targeted ES clones. This 5.2kb product was then gel purified and sequenced with primer -10355 to identify the existence of the *loxP*3 site. Five out of six were sequenced successfully and identified to contain loxP3 sites. One clone did not sequence well. Sequencing with primer 9860 indicated the presence of *loxP*2 in all three clones that sequenced well. However, only three out of six ES clones (2C-3, 2D-6 and 2G-9) were identified to contain *loxP*1 site, the other three ES clones (2H-2, 2G-4 and 2E-7) were found to have wild type sequences instead of *loxP*1. Clones 2C-3 and 2D-6 contained all three *loxP* sites, which were used for blastocyst injection. *LoxP*1 and *loxP*3 were found in Clone 2G-9, but we failed to obtain good sequence for the *loxP*2 (Fig. 10d). However, we had identified two clones that contain all three loxP sites; hence we did not spend more time on clone 2G-9, which was probably correct. The results of sequencing analysis, which across the *loxP*1, *loxP*2, and *loxP*3 region on six potential positive targeted ES clones, indicated that 50% of our targeted ES cell clones have lost the *loxP*1 site during recombination events. In these targeted ES clones, the flanked A1aA1b region was used for homologous recombination instead of our selected 5' homologous recombination region (Fig. 10d).

The position of insertion of targeting sequence in ES clones was also tested by Southern blot analysis with *Sac*II digested genomic DNAs using a probe extending across A1a and A1b exons. A new *Sac*II site was introduced in the targeting vector next to the loxP3 site (Fig. 11A), so an 11 kb *Sac*II fragment would be present in homologous recombination allele in addition to the 18kb fragment from the unchanged wild type allele using this probe (Fig. 11A). Because the introduced *Sac*II site is locate next to the *loxP*3 site, it would not be able to identify exactly where the upstream homologous recombination event happened, and recombination either in the 5' homologous recombination region or in the region meant to be floxed by *loxP*1 and *loxP*2 would produce the same *Sac*II digestion pattern. This was clearly showed in Fig. 11B. Both endogenous (18kb) and targeted (11kb) alleles were detected on the ES clones loaded in lane 2,3,4 and 5. However, only the ES clones loaded in lane 4 (2C-3) and 5 (2D-6) contain all three *loxP* sites. Hence to identify correctly targeted ES cells clones, the sequence analysis of all three *loxP* sites is still a necessary step for our designed targeting vector.

Cre excision in targeted ES cell

The expression of the neomycin phosphotransferase gene driven by the phosphoglycerate kinase I promoter (*PGK-Neo*) is widely used as a selectable marker for homologous recombination in ES cells. Previous studies have shows that retention of the *PGK-Neo* selection cassette in targeted loci may yield unexpected phenotypes in knockout mice due to the altered expression of neighboring genes (*Olson, 1996)*. To

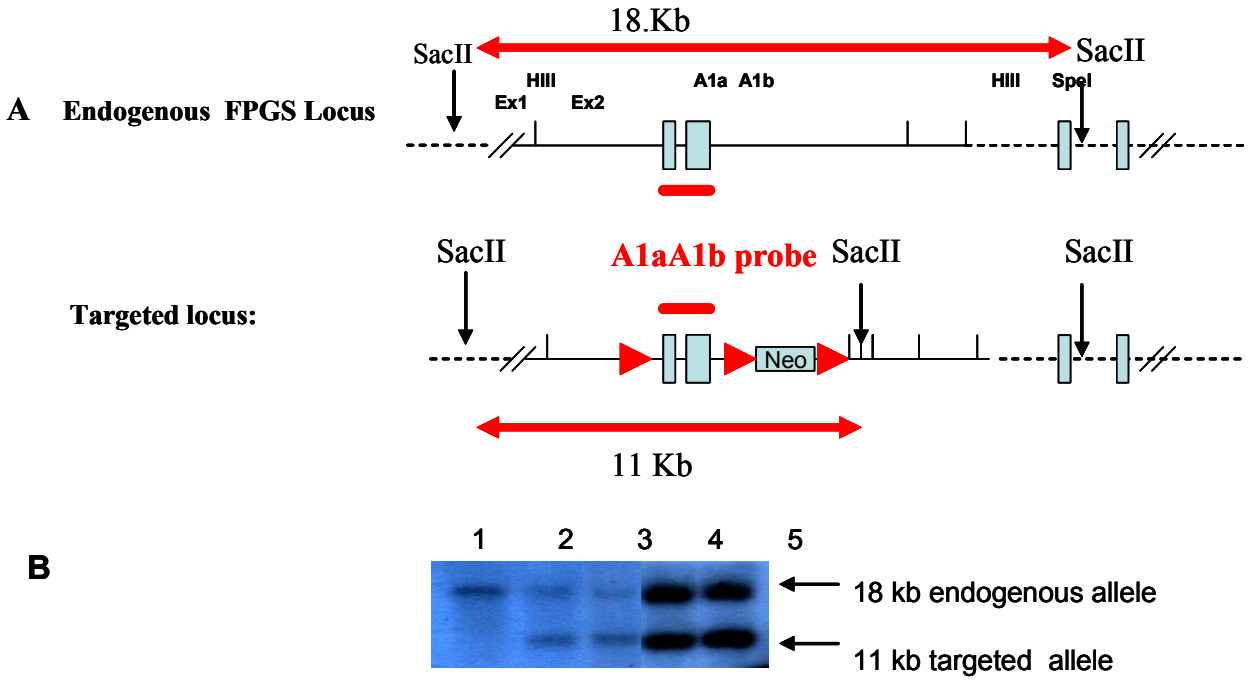


Fig 11. Southern Analysis of Homologous recombination. A: The diagram shows the *Sac*II sites at endogenous *fpgs* locus and targeted locus with the A1aA1b probe. The Targeted ES cells have a 11 kb *Sac*II –digested band, in addition to an 18 kb wild type band, following hybridization with the A1aA1b probe. B: Southern blot analysis of the ES cell clones. Clone 1-wild-type ES clones, clone 2, 3,4,5 were targeted ES cells, but clone 2 and 3 lost loxp1 site. (the genomic DNA were isolated from low density cells, so the concentration is much lower in Lane 1,2,3)

eliminate the potential interference of the *PGK-Neo*, the pCMV-*Cre* plasmid** was introduced into the correctly targeted ES clone 2D-6 by electroporation. The expression of *Cre* recombinase in ES cell that contain targeted loci with *loxP* sites will result in three possible *Cre*-mediated deletions (Fig. 12B): 1.) The *Cre* recombinase catalyzed recombination between the first and the third loxP sites resulting in deletion of upstream promoter P1 and A1a/A1b exons and the *Neo* cassette (complete deletion). This ES cell would be useful for generation of P1 null mice (-/-). 2.) The *Cre* recombinase catalyzed recombination between the second and the third loxP sites; this will result in a deletion of *PGK- Neo* cassette, leaving two *loxP* sites flanking upstream promoter P1, A1a and A1b exons. This ES clone could be used to generate conditional knockout mice. 3.) In the last alternative, the *Cre* recombination could catalyze recombination between the first two *loxP* sites, resulting in the deletion of exons A1a and A1b and retention of the *Neo* cassette. The cells with the complete deletion and the conditional allele are not viable under *G418* selection, allowing identification of the useful alleles (Fig. 12).

After transfection, the cells were cultured for two days and spread into 100mm dishes by serial dilution (500-2000 cells per dish) to form colonies. About twenty-four clones were picked after one-two week culture and seeded in duplicate plates. One set was grown in selection media with 180 ug /ml G418. (This was done by Mrs. Herju Zhang in the Transgenic Core Facility at VCU). The duplicate plate was used to expand the non-surviving clones from selection foe genotyping of targeted allele by PCR. Fourteen out of twenty-four *G418* sensitive clones contained *Cre* mediated deletion

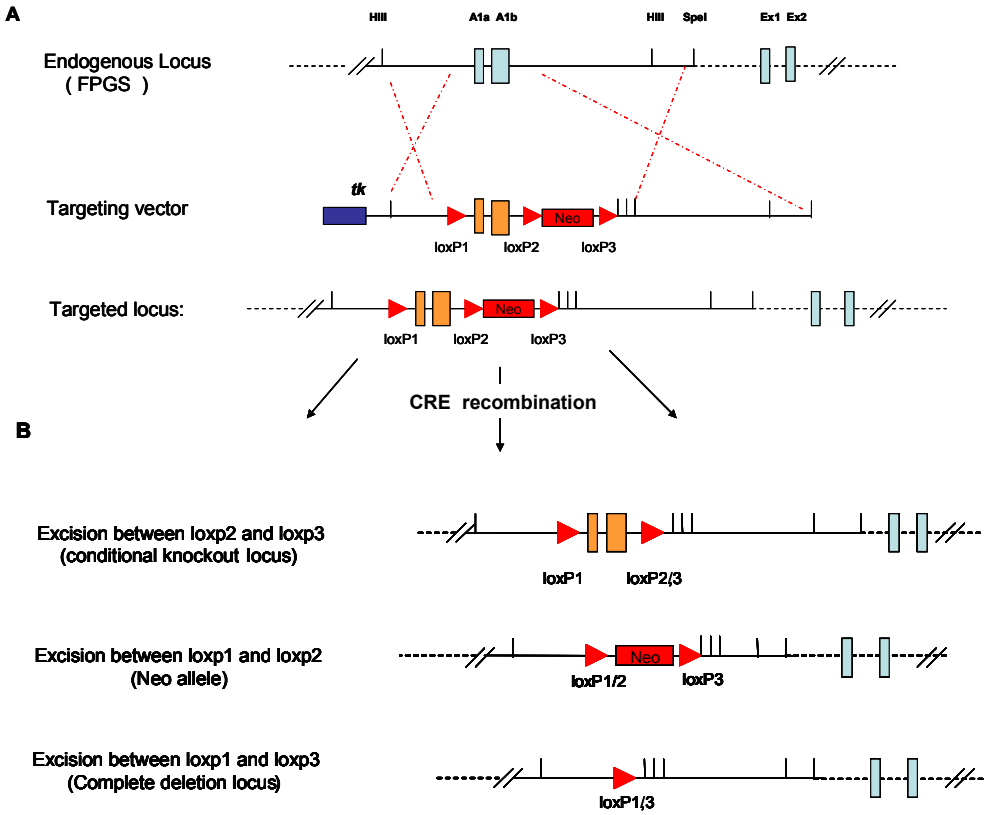


Fig 12. Gene targeting in ES cell and *Cre* mediated deletion of the A1aA1b region and the selection marker. (A) Schematic diagram showing homologous recombination of the targeting vector with endogenous *fpgs* locus in ES cells. (B) *Cre*-mediated recombination products. Recombination between *lox*P2-*lox*P3 sites (large filled triangles) that flanked the *Neo* marker, leads to the production of a conditional knockout locus. Recombination between *lox*P1 and *lox*P3 lead to the deletion of the upstream promoter and A1aA1b exon, leaving one *loxP*1,3 site in the genome, a complete deletion allele. Recombination between *loxP*1 and *loxP*2 would delete P1 and exons A1a and A1b, but would leave behind a *Neo* casstte.

alleles of these six ES clones contained the alleles that recombined between *loxP*2 and *loxP*3 (Fig. 13), and eight clones had recombined between *loxp*1 and *loxP*3 (complete deletion). Hence, transient expression of *Cre* recombinase in targeted ES cell 2D-6 was sufficient to induce recombination in more than 50% of the cell clones studied. The G418 sensitive ES cell clone #10 contained a complete A1aA1b knockout allele (KO-10) and clone #18 contained a *loxP* flanked A1aA1b allele (conditional knockout CKO-18), these clones were chosen for C57BL/6 blastocyst injection to create chimeric mice.

Generation of homozygous knockout mice

An ES cell clone that had undergone correct homologous recombination and *Cre*-deletion of the *Neo* gene (conditional knockout clone #18, CKO-18) and a second targeted ES cell clone with Cre-mediated complete deletion of P1 promoter and exons A1a and A1b (knockout clone #10, KO-10) were injected into C57BL/6 blastocysts by Mrs. Herju Zhang of the VCU-transgenic core facility. Injected blastocysts were transferred into three pseudo-pregnant mice. The chimeric nature of resulting offspring was determined by coat color. Three KO-10 chimera males and nine CKO-18 chimera males were identified that had 50% - 95% agouti coat color. These males were bred to C57BL/6 females to obtain germline transmission of the targeted allele.

** Note that the pCMV-Cre plasmid contains a neo-resistance cassette, so that any subclone that stably integrated pCMV-Cre could be discarded during cellular screening for Neo-resistant subclones.

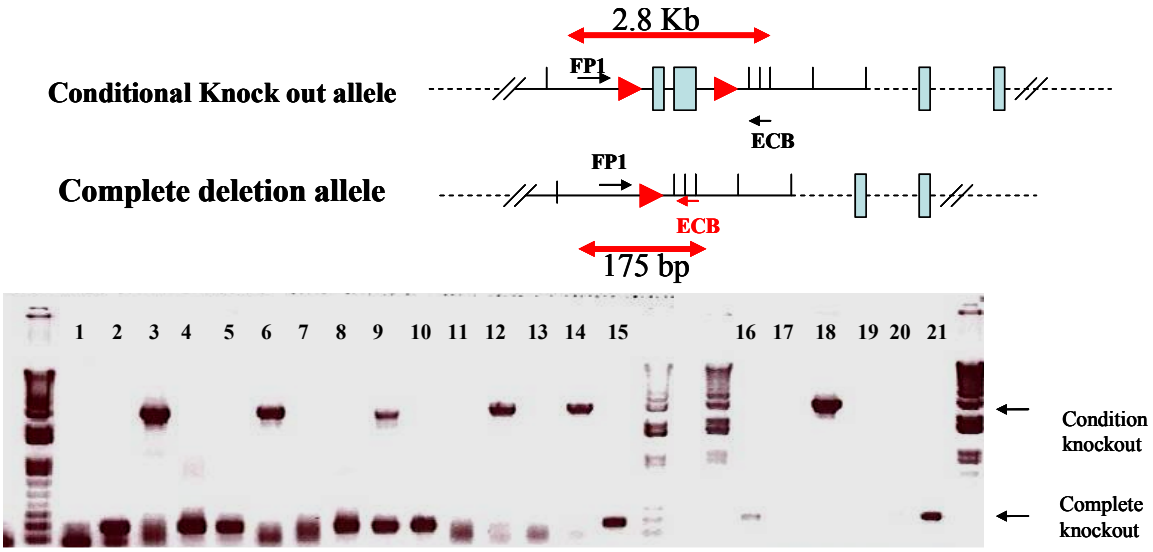


Fig 13. In-vitro *Cre* recombination. A, Diagram of alleles in targeted ES cell that resulted from *Cre*-mediated deletion of the A1aA1b Exons and *PGK-Neo* cassettes. Arrows indicate the location of PCR primers (FP1/-ECB) and the size of the PCR product. B, the results of PCR analysis using the primer set FP1/-ECB (illustrated in A) in order to determine Cre recombination events. PCR products were analyzed in a 1% agarose gel.

One KO-10 chimera male (#3) was sterile (Table 2). The other two chimeras, both of which contained higher percentage of agouti coat color, produced forty-three pups, 75% of these pups had agouti coat color. Forty-one pups from CKO18 chimera were all agouti. The agouti coat color from the 129/*sv* ES cell is dominant over C57BL/6 black coat color, which facilitates the identification of pups derived from the 129/*sv* ES cells. Based on the results of PCR genotyping of tail DNA (see Methods and Fig 10) of each animal (Table 2), we found that half of the F1 agouti offspring were heterozygous for the targeted alleles (complete knock out or conditional knock out). Heterozygous F1 mice (+/-), which carried the FPGS-P1 deletion allele, were mated to generate homozygous knockout F2 mouse with a mixed 129*sv*/C57BL/6 genetic background. Seventy-seven F2 mice resulting from heterozygous F1 intercrosses were genotyped and assigned to one of the three possible categories: wild type (+/+), heterozygous (+/-) and homozygous mutant (-/-) mice. The genotypic analysis confirmed the presence of complete knockout pups, as well as wild type and heterozygous individuals. The ratio of homozygous: heterozygous : wild-type mice agreed with 1:2:1 Mendelian frequency (Table 3). No obvious phenotype differences were observed between animals with the three genotypes. Fewer females were obtained than expected in both in the homozygous KO genotype and wild-type genotypes (Table 3). This first data set suggested that the targeted deletion of the FPGS P1 promoter was compatible with an apparently normal embryonic development, an observation that surprised us. The homozygous P1 knockout (-/-) mouse exhibited an apparently normal phenotype, showing no gross

Table 2 Chimera offspring germline transmission summaries

Mouse strain	Cell Line	Chimera	Total Pups	Agouti Pups	Percentage	Genotype A1aA1b
Complete P1-A1aA1b knockout	KO-10	3 Agouti male				
		KO10-1 M 95% agouti	18	9	50%	3 +/- 6 +/+
		KO10-2 M 70% agouti	25	25	100%	12 +/- 13 +/+
		KO10-3 M 50% agouti	21	0	0	
Conditional knockout	CKO-18	9 agouti male				
		CKO18-2 100% agouti	15	15	100%	2 +/c 7 +/+
		CKO18-9 98% agouti	18	18	100%	5 +/c 9 +/+
		CKO18-10 75% agouti	8	8	100%	3 +/c 8 +/+

morphological defects, and were not distinguishable from their wild type or heterozygous littermate. Moreover, both male and female mutant mice were fertile and both were used to establish subsequent homozygous mutant mouse colonies from F2 individual of the two mixed genetic background -129sv/C57BL/6 (Table 4, 5).

Breeding targeted alleles into the C57BL/6 strain.

Our targeted ES cell lines were derived from the 129/sv strain of mice. Many studies have shown that the 129/sv strain has significant biological limitations that could interfere with phenotypic analysis of a targeted allele, because 129 strains are actually a diverse and complex family of sub-strains *(Crawlel, et al. 1997).* C57BL/6 mice breed well and are commonly used in the production of transgenic mice. In order to obtain a C57BL/6 congenic knock-out strain, the heterozygous F1 mice were backcrossed to C57BL/6 for six generations before being interbred to produce the homozygous P1-A1aA1b knockout and wild-type littermate mice used in this study. Two sets of oligonucleotide primers were designed for PCR analysis to distinguish wild-type, heterozygous, and homozygous mice of complete P1-A1aA1b knockout (Fig. 14A, 15). Primer FPGS 9047 and FPGS -9322 were located in the exons A1a and A1b and are designed to detect the existance of wild-type alleles. PCR amplification on the genomic DNA that was purified from the wild-type or heterozygous KO mouse with primers 9047-9322 would produce a 275 bp DNA band (Table 1 and Fig. 14, 15). Targeting vector specific primer "-ECB" and *fpgs* intronic primer FP1, were designed to

Table 3. Genotype of C57BL/6 backcross progeny –KBF6

Genetic background C57BL/6	**Heterozygous Knockout** +/-	**Wild-type** +/+	**Total**	**Mendelian frequency**
Total expected	78	78	156	1:1
Total observed	78	78	156	1:1
Male observed	44	43	78	1:1
Female observed	34	35	78	1:1

Table 4. Genotype of progeny of F1 or F6 heterozygous intercrosses

Genetic background	**Wild-type** +/+	**Heterozygous Knockout** +/-	**Homozygous knockout** -/-	**Total**	**Mendelian frequency** +/+ : +/- : -/-
(C57BL/6)F6 -2KBF1-2 expected	37.25	74.5	37.25	149	1 : 2 : 1
Total observed	47	67	32	149	1 : 1.4 : 0.7
Male observed	24	33	19	76	1 : 1.4 : 0.8
Female observed	23	34	13	70	1 : 1.5 : 0.5
(129sv/C57BL6)F1-2KOF1 Expected	18	35	18	72	1 : 2 : 1
Total observed	19	34	19	72	1 : 1.8 : 1
male observed	13	16	12	41	1 : 1.2 : 0.9
Female observed	6	18	7	31	1 : 3 : 1.1

amplify a 175 bp DNA band from the mouse genomic DNA that contain the P1-A1aA1b knockout allele.

Out of 156 KBF6 (KO #10 germline backcrossing into the C57BL6 strain for six generations) mice, 78 were male and 78 were female, 44 male and 34 female mice contained the heterozygous P1 KO allele (Table 3). The ratio of mice with genotype +/+ to mice with genotype +/- was 1:1, and the ratio of male to female with genotype +/- was 1.2:1. This gender distribution was found to differ from Mendelian transmission of genotypes from which an equal number of males and females were expected. However, when compared with wild type mice and confirmed with χ^2 value (= 1.05) the deviation was not statistical significant $P > 0.05$.

Phenotype of complete knockout mice

Crossing of heterozygous mice produced viable homozygous offspring in both the C57BL/6 (Table 5) and 129SV/C57BL/6 mixed backgrounds (Table 4). Breeder pairs of *fpgs* heterozygous males and females generated normal litter sizes (average 8 pups/litter) at the expected frequency (every 3-4 weeks) (Table 4). One hundred forty nine pups were produced from intercrossing of heterozygous mice with C57BL/6 background. The genotyping analyses of these pups confirmed the presence of homozygous knockout mice (-/-), as well as heterozygous mutant (+/-) and wild type (+/+) at the ratio of 1:1.4:0.7, which was different from the expected Mendelian ratio of 1:2:1 (Table 4). In addition, the ratio of male and female mice did not deviate from

Table 1 PCR primer sequences for the targeting allele analysis in ES clones					
Forward primer (5' to 3')		Reverse primer (5' to 3')		product size (bp) CDKO	KO
FPGS+4309	CAGAAGTTTCTGCCACCTCCCAAT	NEO 1B	CATTCGACCACCAAGCGAAACATC	6400 bp	
+ECB	GATCGATCCTAGAGAATTCCTAGACCG	FPGS-15459	ATTGCATATCTGTGGAGGTCAGGG	6400 bp	
FPGS-9047	TGCGAACCTCTCAGGTAAACACTC	FPGS-9322	CCTGAAAGGAGGCAGTCTTAGCTT	275 bp	
FP1	TGAGTCAGTAGGCTCAGTGTGAGA	-ECB	GCGGTCTAGGAATTCTCTAGGATCG	4664 bp	175 bp
FPGS-9860	TCGCTTGTCTACCACGCACTGTAA	FPGS-10355	TGTTCAGGCATGTCACAGAGTCCA	2250 bp	576 bp

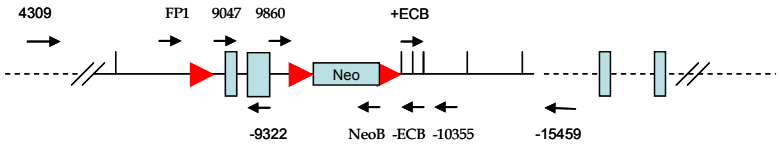


Fig14. Schematic diagram showing the location of the primer pairs (black arrow) used in PCR for mouse *fpgs* genotyping. The sequences of primers were shown in the table1. CDKO: The size of the PCR product generated from a conditional knockout allele. KO: the size of the PCR product generated from complete P1 Aa1aA1b knockout allele.

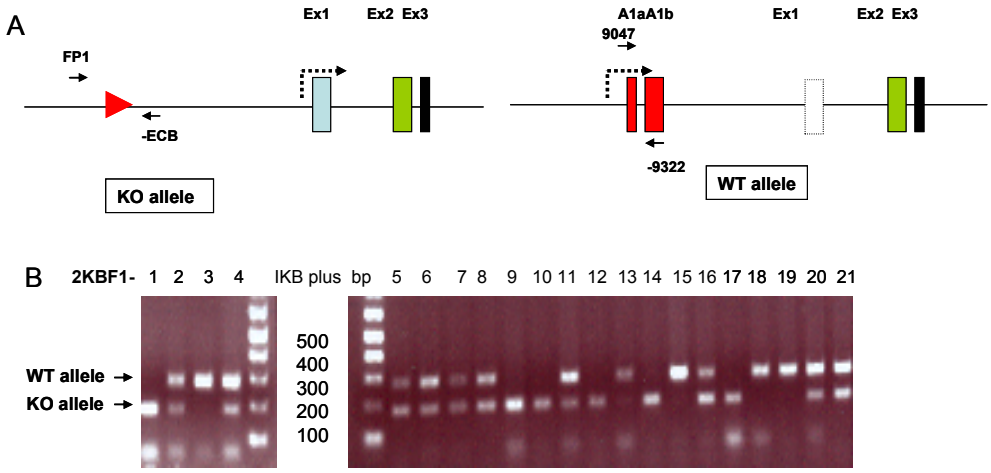


Fig 15. Genotyping of C57BL/6 backcrossing offspring. (A) Structure of the wild-type and knockout *fpgs* alleles. The start sites of the two transcript *fpgs* are shown by the dotted arrow. The primers used in the analysis are indicated as short arrows. (B) PCR genotyping of mouse tail DNA. The wild-type allele produces a band at approximately 295bp with primers 9047 and 9322, while the mutant allele produces a band at approximately 175bp with primers FP1 and -ECB. The wild-type (lanes 3, 15, 18, 19) mice show only one band at 295bp, the heterozygous mice (lanes 2,4,5,6,7,8, 11,13,14,16, 20, 21) show both 295- and 175-bp bands and the homozygous KO mice (lanes 1,9,10,14,17) show only one band at 175 bp.

expected for the homozygous -/- mice. Homozygous (-/-) males and females were obtained at the ratio of 1.4:1, which was not statistically significant different from the expected 1:1, and also confirmed by χ^2 value (=1.12) $P > 0.05$. No differences were observed between animals derived from two genetic backgrounds, 129sv/ C57BL/6 mix and C57BL/ 6 genetic backgrounds in terms of sex ratios (Table 4). N6F1 homozygous (-/-) mice with C75BL/6 background, developed normally and do not showed any obvious alteration in morphology. There are no differences with mice body length or body weight between knockout mice and wild-type mice (Fig. 16). Moreover, both male and female 2KBF1 -/- mice were fertile and was used to produce subsequent homozygous knockout mouse colonies (2KBF2). The absence of apparent overt phenotypic differences in homozygous KO mice was further investigated. The growth curves of 2KBF1 knockout mice vs. wild -type litter mates were analyzed from the onset of weaning age (4 weeks old), fed either with normal diet (1300 ug folate /kg diet, Test Diet) or low folate diet (63 ug folate/kg diet, Test Diet). All three genotypes of mice (+/+, +/-, and -/-) were observed for growth, each group contained at least six mice, that were produced from several independent litters. Body weight was recorded twice a week from 4 weeks of age to 14 weeks of age. Body weight increased with aging and was higher in males than in females. No significant weight differences were detected between the mice fed with normal diet and low folate diet (Fig 19, 20). All genotypes (+/+, +/-, -/-) appeared to grow and gain weight at similar rates, and follow comparable

Table 5. Homozygous KO mouse reproduction and lifespan

Filial generation number	Number of Mouse produced	Number of female	Number of male	Time observed	Number of litter produced
F1 +/-	10	6	2		
2KOF2	19	7	12	11	
2KOF3	39	15	24	13	5
2KOF4	62	28	33	9	9
2KOF5	23	10	16	7	4
2KOF6	10	7	2	4	2
2KOF7	5	2	3	3	1
TOTAL	158	69	90		21

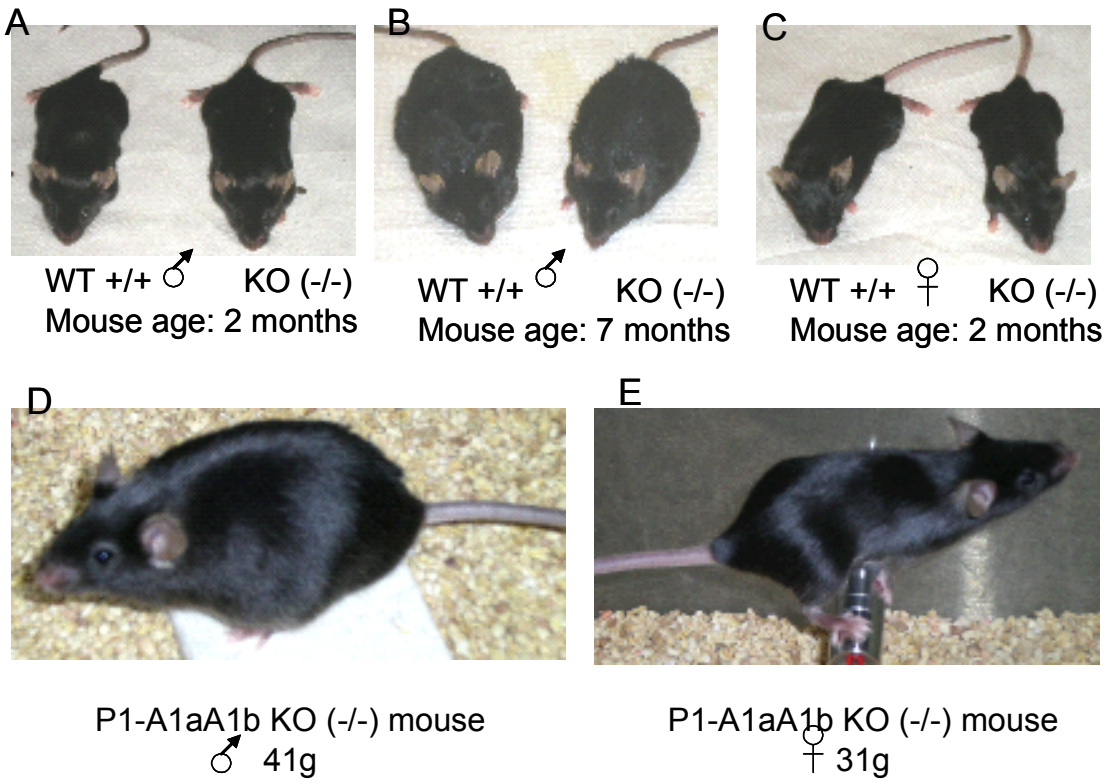


Fig16. Comparison of homozygous P1 KO mice with its wild-type litter mate. (A). 8 weeks old adult male KO and WT mouse. (B). 7 month old male KO mice and WT male mice. (C) 8 weeks old adult female KO mice and WT mice. Mice in A,B and C were under anesthesia when the pictures were taken for better comparison. (D,E) KO mice with out anesthesia.

growth curves. Differences in weight between wild type and KO mice of similar age were not statistically significant. (Fig. 17, 18) More detailed analysis over extended periods of time is currently underway to uncover any potentially subtle differences in A1aA1b knockout versus wild type mice.

Effect of P1 deletion on *fpgs* gene expression in mouse liver

To define the effects of deletion of P1 and exons A1a and A1b on mouse liver *fpgs* gene expression, total RNA was extracted from adult liver from *fpgs*-P1 homozygous knockout (-/-), heterozygous -knoockout (+/-) and wild-type (+/+) mice, and mRNA was reverse-transcribed into cDNA using an oligo dT primer (Invitrogen). Real-time reverse transcription polymerase chain reaction (RT-qPCR) experiments, based on SYBR Green detection, were conducted to measure the expression of mRNA transcript levels from *fpgs* promoter P1 and P2. Two sets of *fpgs* specific primers were designed to detect the expression level of *fpgs* mRNA species; 1.) Primer A1aA1b-A was located in the sequences of exon A1a (5'-TGT GCA CAG GAT GGA AAT GCG AAC C-3'), and its anti-sense primer –Ex3a (5'- TGT TTA GCC GGT TCA AGT CCT CCA – 3') was located in exon 3. Together, these primers would amplify a 340bp cDNA fragment from P1 *fpgs* mRNA species. 2.) Primer Ex1 (5'-CTA GCT TGC GGT GGC ATT AAG ACT- 3') was located in the sequence of exon 1(Fig. 21), which was used in the transcript made from P2. Primer Ex1 and anti-sense primer -Ex3a amplify a 343 bp cDNA fragment from P2 *fpgs* mRNA species. The real time PCR results were

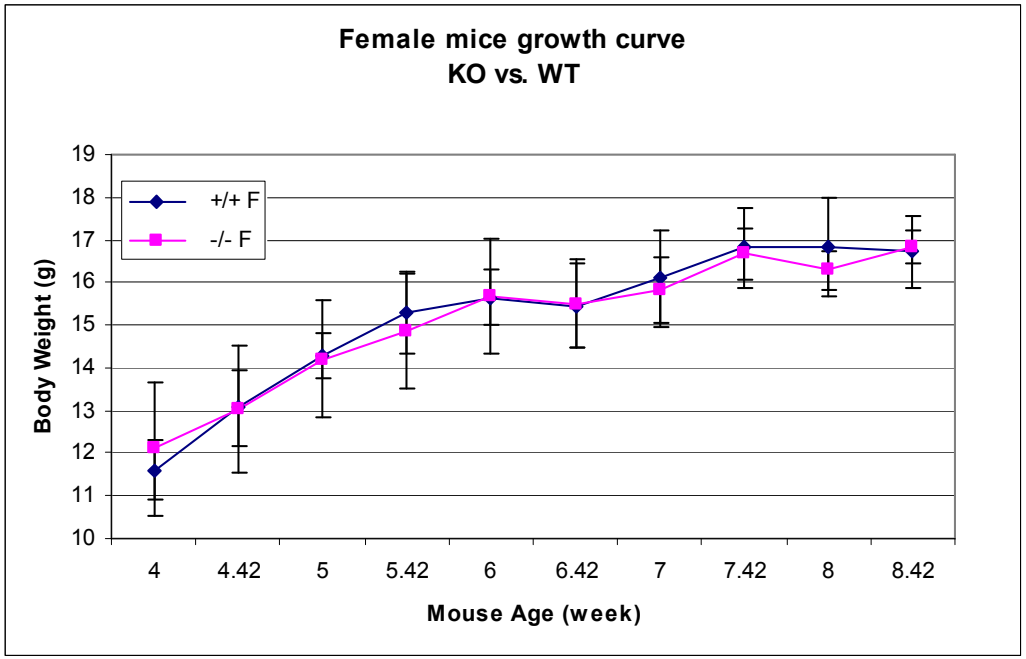


Fig17. Body weight increase in weaning mice of wild-type and KO female mice. Each group contain six weaning age mice with same genetic background (C57BL/6 F6). The body weight were measured twice a week start at weaning age (4 weeks old). The wild-type mice show in close diamond and A1aA1b KO mice show in square.

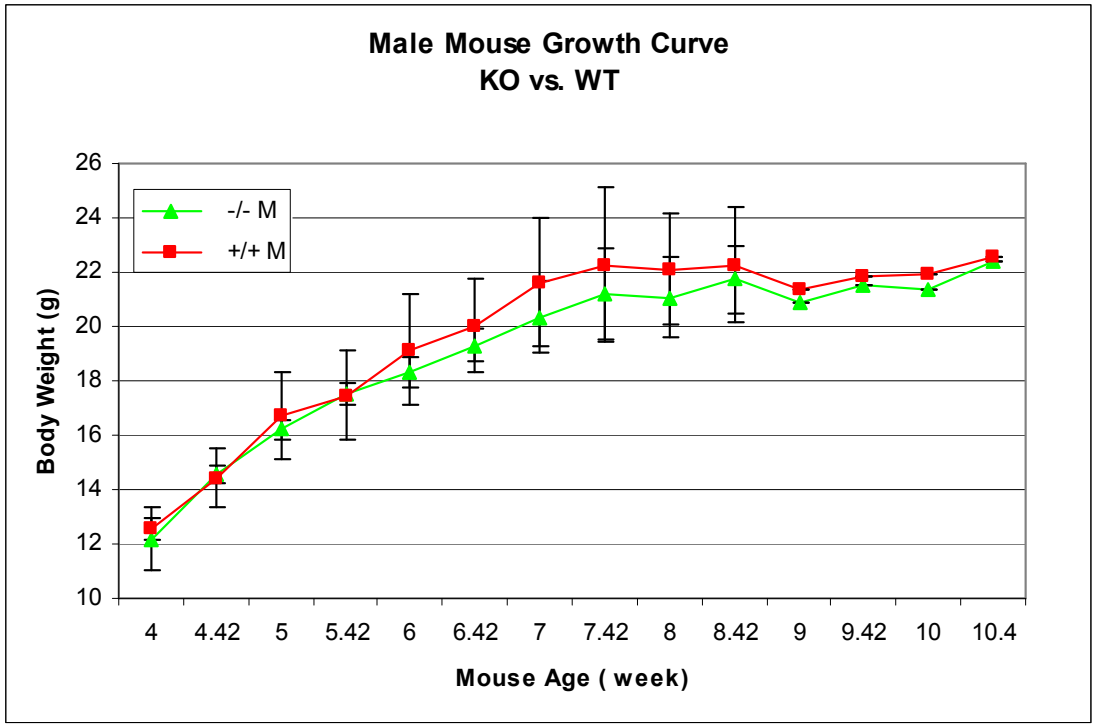


Fig. 18. Body weight increase in weaning mice of wild-type and KO male mice. Each group contain six weaning age mice with same genetic background (C57BL/6 F6). The body weight were measured twice a week start at weaning age (4 weeks old). The wild-type mice show in close square and A1aA1b KO mice show in triangle.

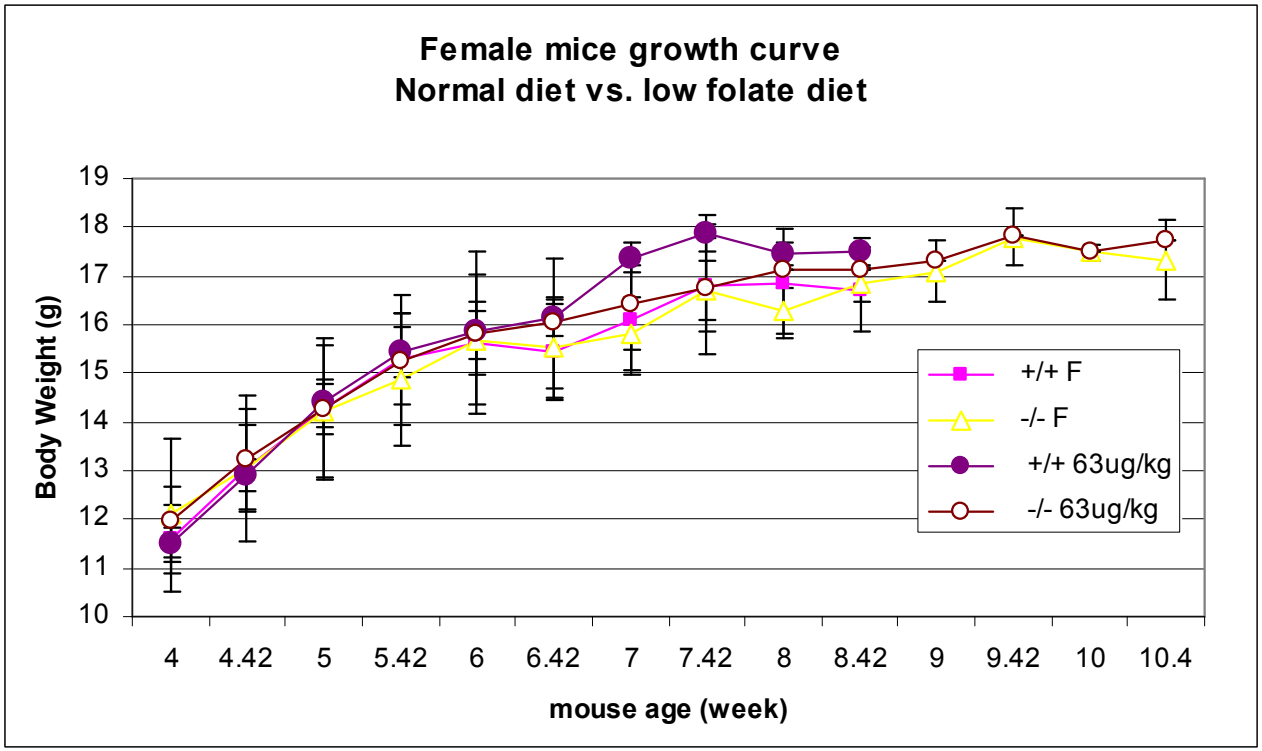


Fig. 19. Comparison of female mice growth pattern of wild-type and KO that fed with the diet contained either 1300 ug/kg or 63 ug/kg folate. Each group contain six weaning age mice with same genetic background (C57BL/6 F6). The body weight were measured twice a week start at weaning age (4 weeks old). The genotype of mice and folate contains of each curve represent are indicated in the box above.

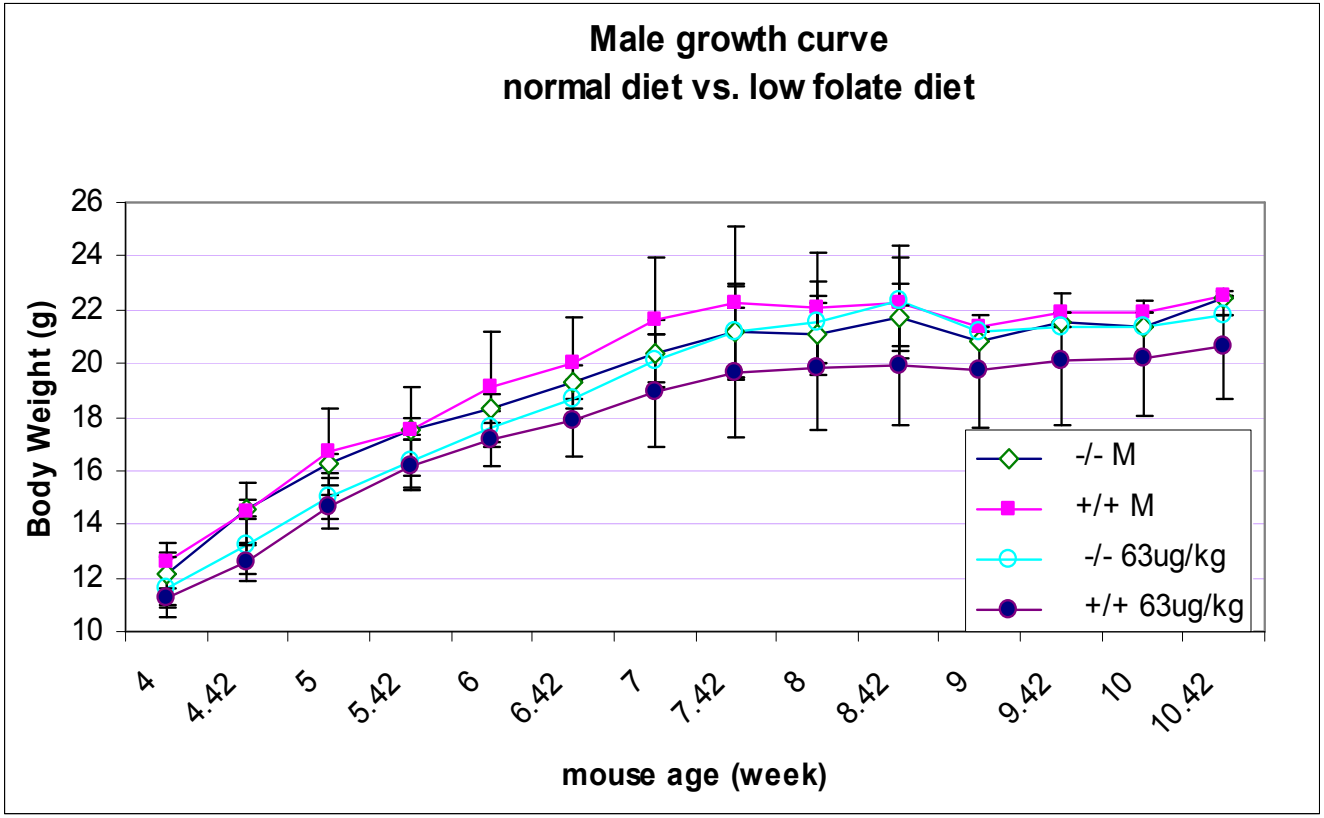


Fig. 20. Comparison of male mice growth pattern of wild-type and KO that fed with diet contained either 1300 ug/kg or 63 ug/kg folate. Each group contain six weaning age mice with same genetic background (C57BL/6 F6). The body weight were measured twice a week start at weaning age (4 weeks old). The genotype of mice and folate contains of each curve represent are indicated in the box above.

normalized to the mouse *actin* gene to control for the cDNA loading in the reaction and calculated as $2^{\Delta Ct}$ method using reference gene (see Method), which allows us to directly compare quantitative data from two separate experiments. All qPCR reactions were in duplicate or triplicate. To make the comparison among each qPCR run reliable, the analysis of gene expression of each genotype were done with RNA from six individual mice. Our results revealed asignificant changes in *fpgs* gene expression following the deletion of upstream promoter P1 and exons A1a and A1b. As expected, P1 transcription was not detectable in homozygous P1 knockout mouse liver, and was 2.4 fold reduced in heterozygous knockout mouse liver compared to the expression in RNA from livers of wild-type mice (Fig. 22). This result indicated that the two *fpgs* alleles are equally expressed in the mouse liver. Deletion of one allele will reduce the total amount of gene transcription by about ½ if both alleles are expressed independently. In contrast, after deletion of upstream promoter P1 and exons A1a and A1b, the expression of the P2 transcript was up-regulated in mouse liver. As shown in Fig. 23, the P2 transcription level increased 4.6 fold in homozygous P1 knockout mice and 3 fold in heterozygous mice. The relative expression levels of P1 to P2 transcription in mouse liver were compared in Fig. 24; RT-qPCR data analysis indicated a 480 fold higher P1 transcript level than P2 transcript level in wild-type mouse liver. Interestingly, in wild-type mouse liver there was a 100 fold higher *fpgs* transcript level, almost all from P1, than the transcript level in the KO mouse, which only expressed *fpgs* mRNA from P2 (Fig 24,).

PCR primer sequences for the targeting allele analysis in ES clones				
Forward primer (5' to 3')		Reverse primer (5' to 3')		product size (bp)
A1aA1b-A	5'-TGTGCACAGGATGGAAATGCGAAC C-3'	-Ex3a	5'- TGTTTAGCCGGTTCAAGTCCTCCA – 3'	340
Ex1	5'-CTAGCTTGCGGTGGCATTAAG ACT-3'	-Ex3a	5'- TGTTTAGCCGGTTCAAGTCCTCCA – 3'	343
Ex3F	5'-TCATGTCACTGGGACCAAAGGGAA-3'	Ex6R	5'-AGACATGGCTGTCGTCCTTGAACT-3'	218

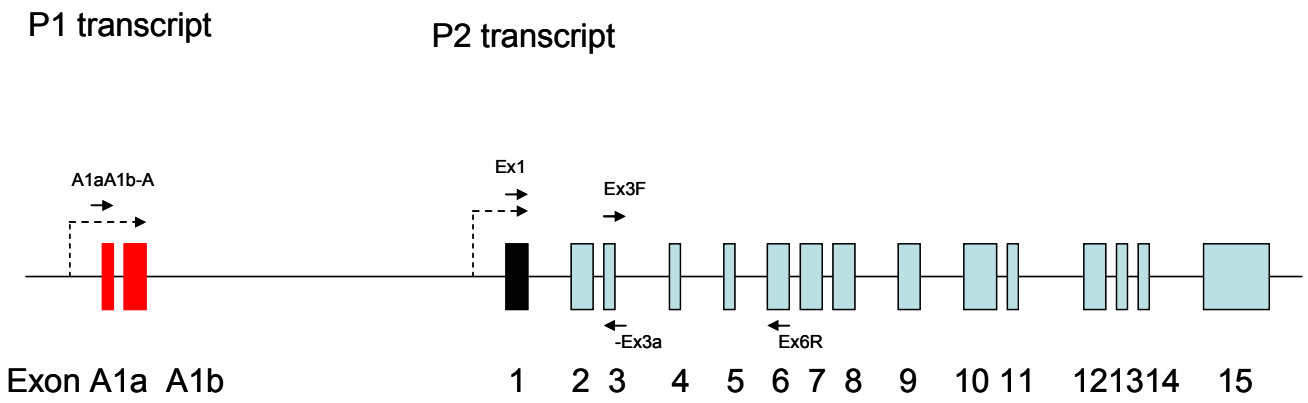


Fig. 21. Schematic diagram showing the location of the primer pairs (black arrow) used in RT-qPCR for *fpgs* gene expression analysis. The sequences of primers and size of PCR product were show in the table1above.

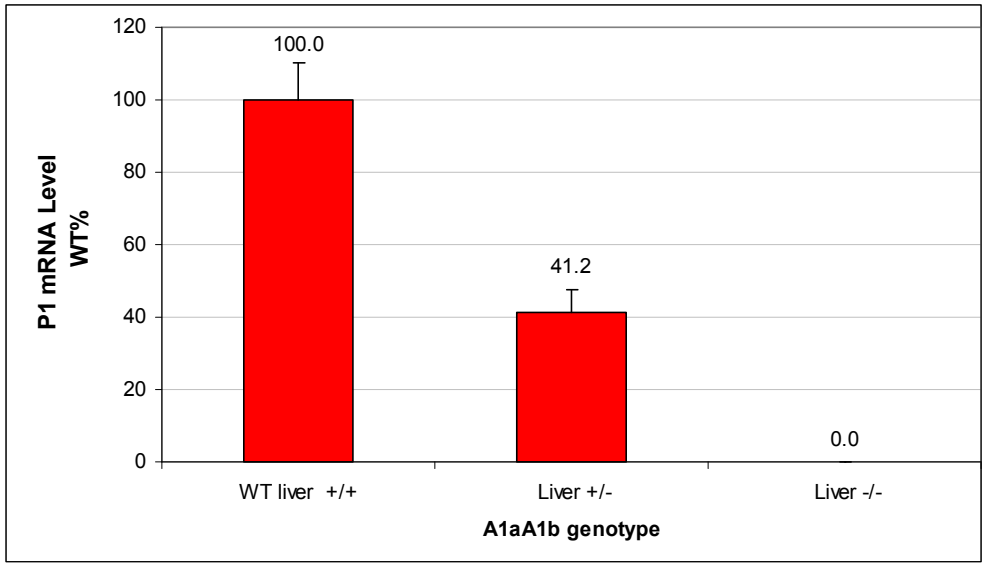


Fig 22. Relative *fpgs* P1 expression change after deletion of upstream promoter P1 and exons A1aA1a in mouse liver determined by real-time PCR. The *fpgs* transcript from P1 was completely abolished in homozygous A1aA1b knockout (-/-) mice liver, and the P1 transcription level was 2.5 fold reduced in heterozygous knockout (+/-) mice liver compared to wild type (WT, +/+). Bars show the mean ±STD from six mice (3male + 3 female) livers.

Q-PCR has a sensitivity that makes it possible to robustly detect and quantify low-level gene expression from RNA in tissue. However, qPCR analysis of a gene expression is based on the PCR amplification by the specific primer set, so that potential bias may be introduced when comparing data across primer sets by differences in template preference, length and variations of PCR amplification efficiency. To make the comparison of the relative abundance of the mRNAs for the two *fpgs* isoforms reliable, primers Ex3F (5'-TCA TGT CAC TGG GAC CAA AGG GAA-3') and Ex6R (5'-AGA CAT GGC TGT CGT CCT TGA ACT-3') (Fig. 21) were designed and used to amplify a 218bp cDNA fragment from exon3 to exon 6 of *fpgs* cDNA. Since both isoforms of *fpgs* mRNA sequences differ only on the N-terminal exon, and have identical sequences after exon 2, the *fpgs* mRNA transcripts detected by Ex3F and Ex6R primer would represent the *total fpgs* transcripts in a tissue. It was used as an internal control in RT-qPCR gene expression analysis. The relative expression levels of P1 to P2 detected by Ex3F-Ex6R primers largely agreed with the previous data determined with A1aA1b-A and Ex1 primers; the total mRNA for *fpgs* was much lower in KO than in wild type liver, but the ratio was only 15-fold different. This result confirmed that the *fpgs* P1 transcript in wild-type mouse liver was between 15-fold to 100 fold higher than the P2 transcript in knockout mouse liver (Fig. 25).

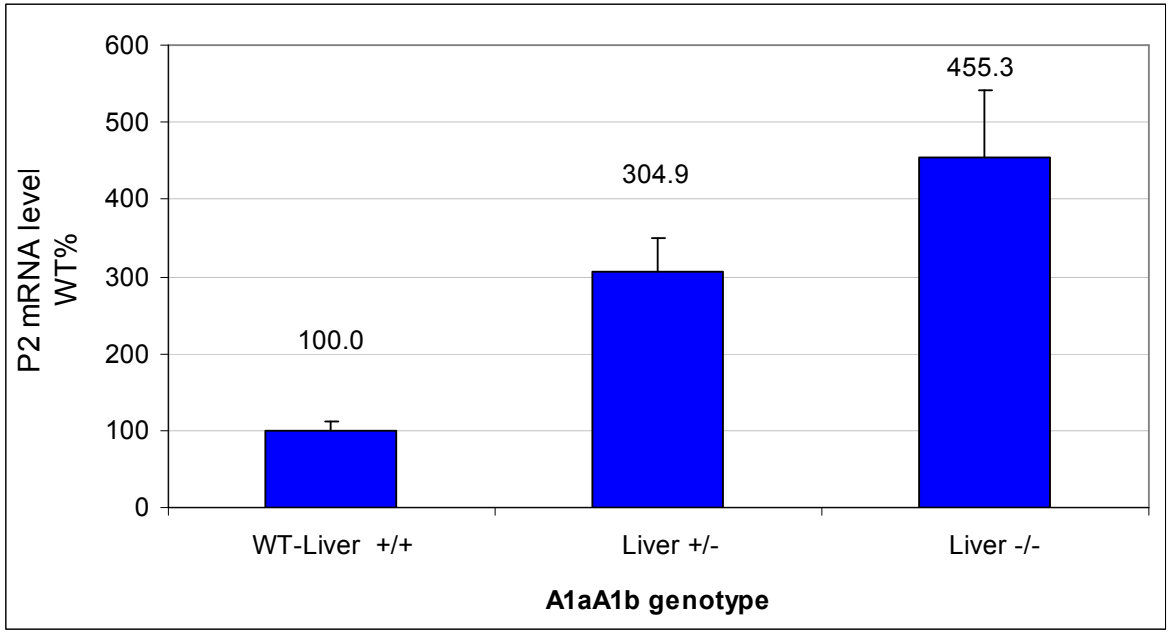


Fig 23. Relative *fpgs* Ex1 isoform expression upon deletion of P1 in mouse liver determined by realtime PCR. Deletion of upstream promoter P1 and exons A1a and A1b up-regulated the downstream P2 transcription by 4.6 fold in KO (-/-)mouse liver, and 3 fold increase in heterozygous A1aA1b KO (+/-) mice liver. Bars show the mean ±STD from six mice (3male + 3 female) livers.

	+/+	+/-	-/-
Liver P1	100	41.18	0
Liver P2	0.21	0.62	0.95

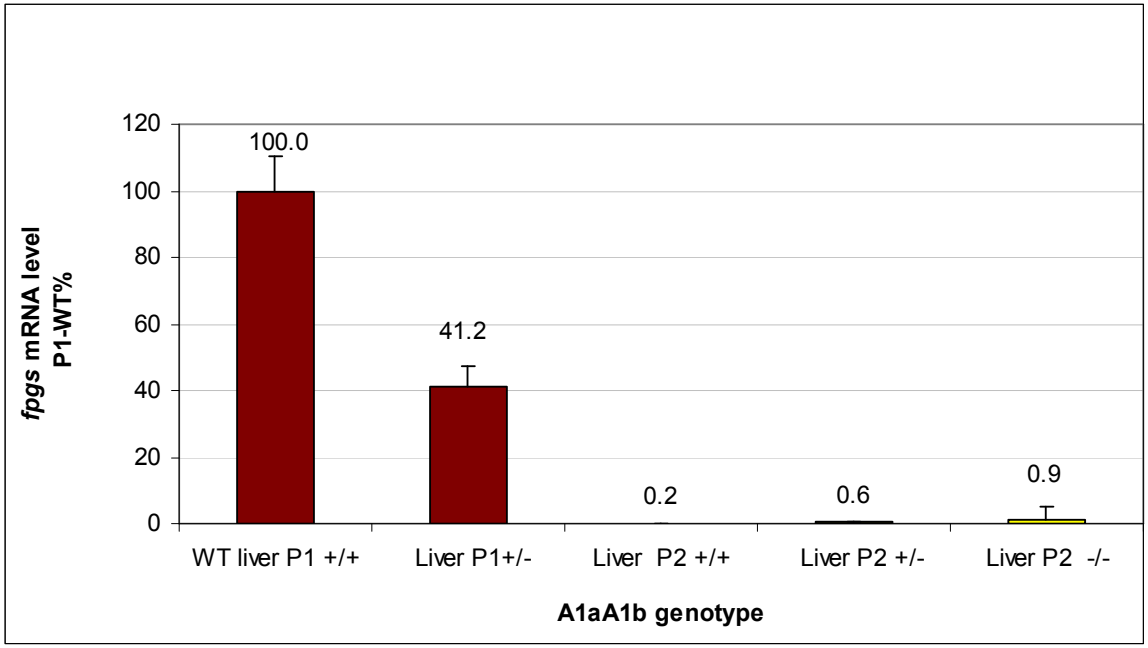


Fig. 24. Comparison of *fpgs* P1 and P2 expression in mouse liver. The *fpgs* P2 transcript level were 99 fold lower in P1 knockout mice (-/-) compare to *fpgs* P1 transcript in wild type mouse (+/+) liver. Bars show the mean ±STD from six mice (3male + 3 female) liver.

Effect of P1 deletion on *fpgs* gene expression in mouse kidney

The results of RNA Protection Assay (RPA) experiments from previous studies in this laboratory by Fiona Turner indicated that mouse kidney expresses transcripts from both P1 and P2 promoters in a ratio of about 6-10 to 1 (*Turner and Moran 1999*). To define the effect of P1 deletion on *fpgs* gene expression from P2 in mouse kidney, real time qPCR analysis were performed on the cDNA reverse-transcribed from the total RNA isolated from mouse kidney from mice with different genotypes. Notably, the expression of P1 mRNA in the kidney was abolished in the homozygous P1 KO mice kidney, and was reduced by half in heterozygous mouse kidney compared to the level expressed in the wild-type kidney. The decrease in P1 expression in heterozygous mouse kidney resulted from the deletion of one of two *fpgs* allele (Fig. 26) and was expected. In contrast, the *fpgs* transcripts originating at P2 were up-regulated 1.5 fold in both heterozygous and homozygous P1 knockout mouse compared to kidney from wild-type mice (Fig. 27). The ratio of transcripts from P1 and P2 level was calculated from qPCR analysis data. Comparing P1 and P2 expression in kidneys from wild type mice, qPCR suggested that P2 transcripts were 28-fold lower than P1 transcripts. P1 expression in wild-type mouse kidney was 18 fold higher than P2 expression in the homozygous P1 knockout mouse kidney (Fig. 28). The 18-fold difference of P1 and P2 expression were also detected by qPCR with internal gene expression primers Ex3F and Ex6R. (Fig 29)

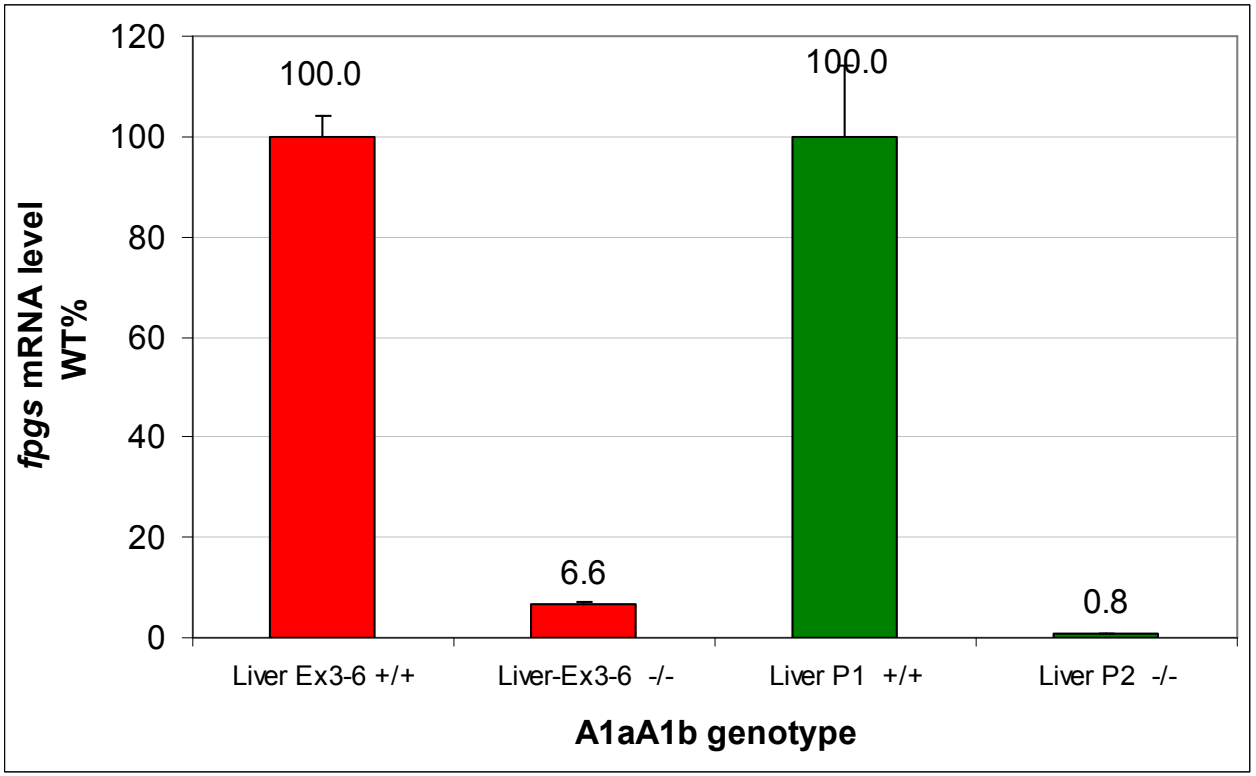


Fig. 25. Comparison of *fpgs* mRNA expression in mouse liver. The expression level of *fpgs* P1 transcript in wild-type (+/+) mouse liver and P2 transcripts in P1 KO mouse (-/-) liver were compared with total *fpgs* transcript level Ex3-6, exons used by both *fpgs* species, P1 and P2. Bars show the mean ±STD of duplicated RT-qPCR analysis of 1 female .

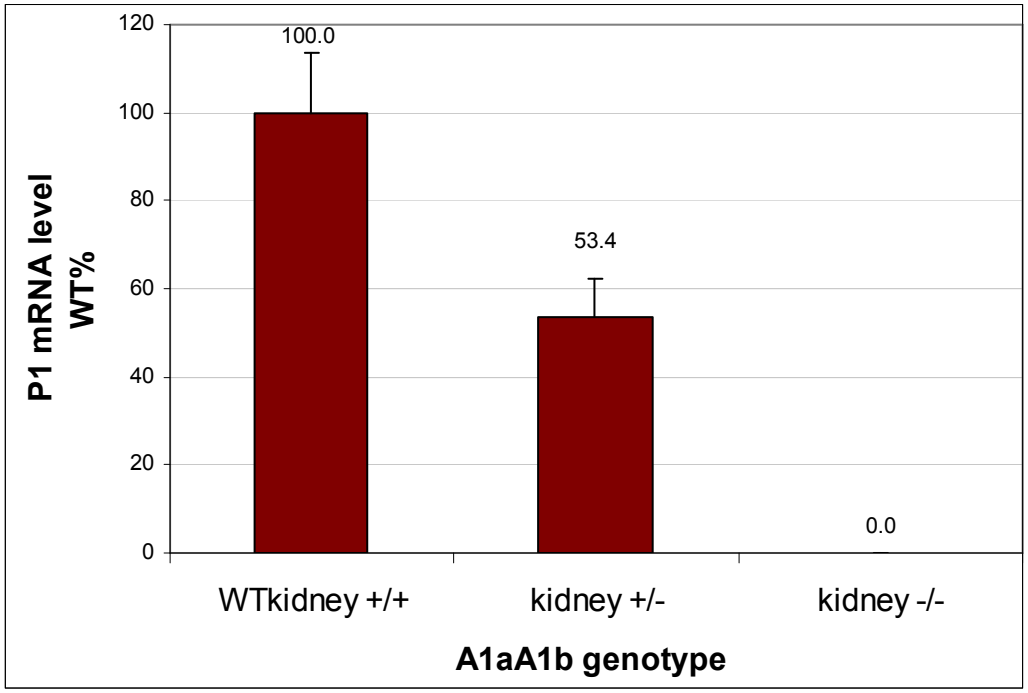


Fig. 26. Relative *fpgs* P1 expression change after deletion of upstream promoter P1 and exons A1aA1a in mouse kidney determined by real-time PCR. The *fpgs* transcript from P1 was completely abolished in homozygous A1aA1b knockout (-/-) mice kidney, and the P1 transcription level was 2.5 fold reduced in heterozygous knockout (+/-) mice kidney compared to wild type (WT, +/+). Bars show the mean ±STD from six mice (3male + 3 female) kidney.

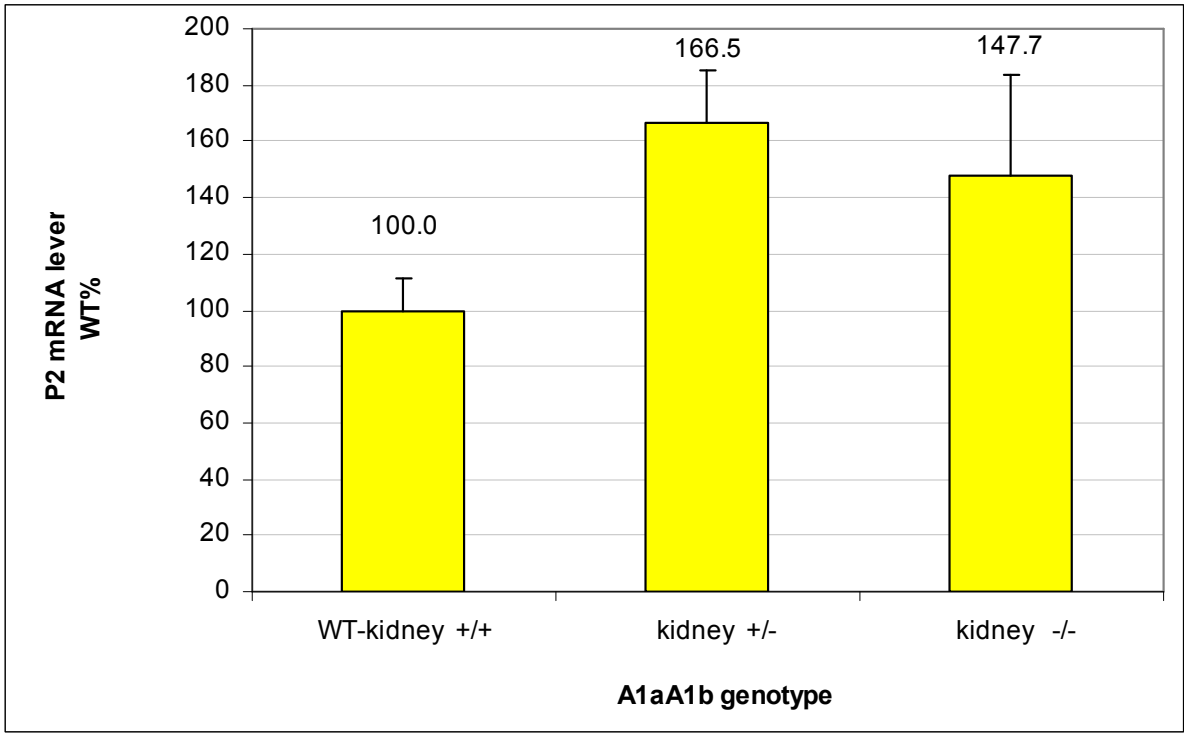


Fig. 27. Relative *fpgs* P2 expression upon deletion of P1 in mouse kidney determined by realtime qPCR. Deletion of upstream promoter P1 and exons A1a and A1b up-regulated the downstream P2 transcription by 0.47 fold in KO (-/-)mousekidney, and 0.66 fold increase in heterozygous A1aA1b KO (+/-) mice kidney. Bars show the mean ±STD from six mice (3male + 3 female) kidney.

The percentage of wild-type fpgs P1 expression

	+/+	+/-	-/-
Kidney P1	100	53.41	0
Kidney P2	3.81	6.35	5.63

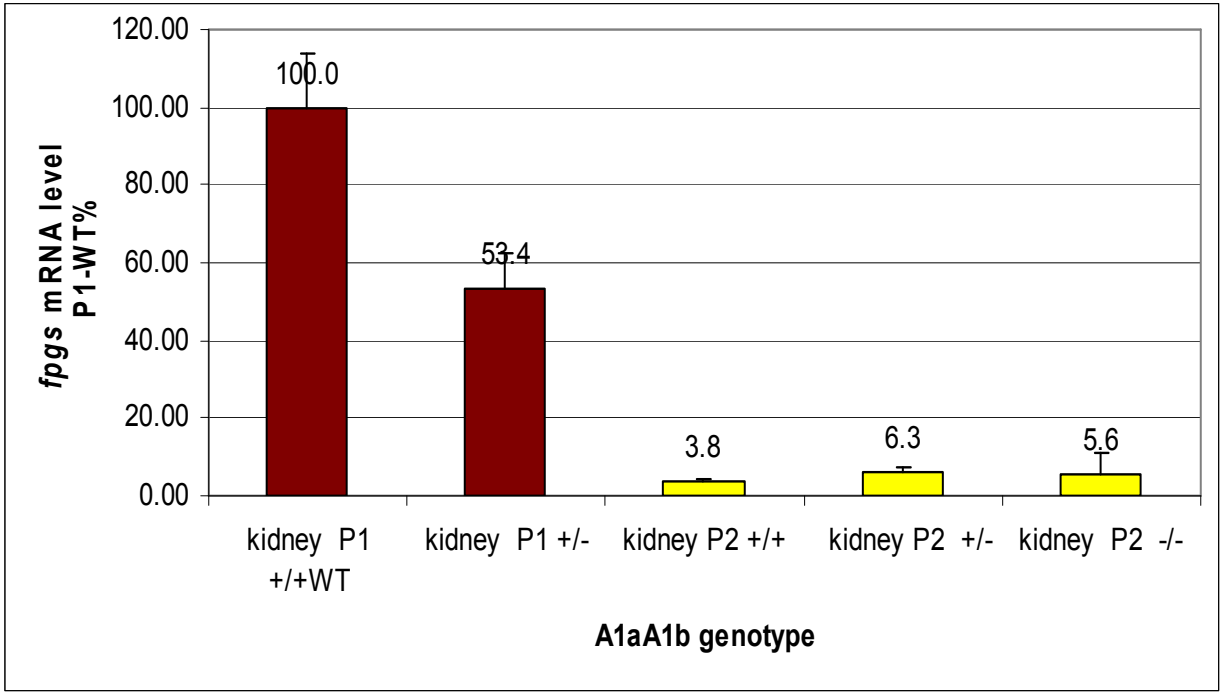


Fig 28. comparison of *fpgs* P1 and P2 expression in mouse kidney. The *fpgs* P2 isoform expression in A1aA1b knockout mice were 17.8 fold lower compared to fpgs A1aA1b isoform expressed in wild type mouse kidney. Bars show the mean ±STD from six mice (3male + 3 female) kidneys.

Comparison of *fpgs* gene expression in male vs. female mouse

To get further insight into the implication of the real time qPCR data, RT-qPCR analysis was applied to groups of male and female mice. The expression of P1 mRNA in wild-type mouse liver is 30% higher in male than in female (Fig. 30), but, in heterozygous P1 knockout mice, P1 transcription level was not significantly different between male and female (Fig. 30). In contrast, the P2 transcription level in homozygous knockout male mice liver is 78% higher than in homozygous knockout female mice liver (Fig. 31). Transcription from P2 in heterozygous mice liver is 50% higher in males than in females, and in wild-type mice liver is 30% higher in males than in females (Fig 31).

The gender differences in *fpgs* gene expression were also observed in mouse kidney. The expression of P1 mRNA in heterozygous KO mice kidney is 78% higher in male than female, but no differences were observed in wild-type mice (Fig 32). P2 mRNA expression level in homozygous knockout mice kidney is 78% higher in male than in female. However, in both heterozygous knockout mice and wild type mice kidney, this difference is not as significant as in homozygous knockout mice (Fig 33). In conclusion, the overall expression of *fpgs* in both liver and kidney was substantially higher in male mice than in female mice.

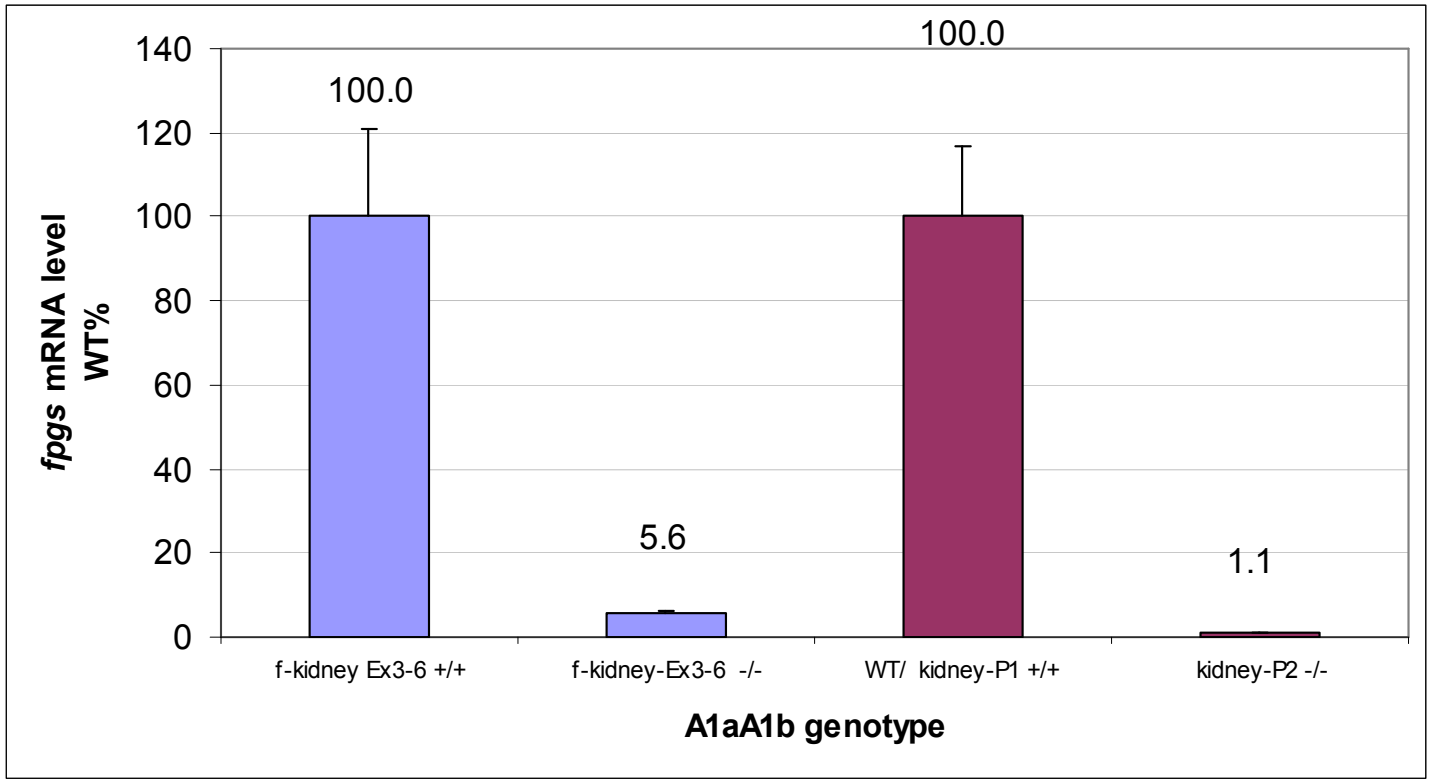


Fig 29 Comparison of *fpgs* mRNA expression in mouse kidney. The level of *fpgs* P1 transcript expressed in wild-type mouse kidney and P2 transcripts expressed in P1 KO mouse kidney were compared by qPCR analysis, and the total *fpgs* transcription level using Ex3-6 primer (exon 3-6 were common used by two *fpgs* species, P1 and P2) were also compared by real time qPCR. Bars show the mean ±STD of duplicated RT-qPCR analysis of 1 female (2KBF1).

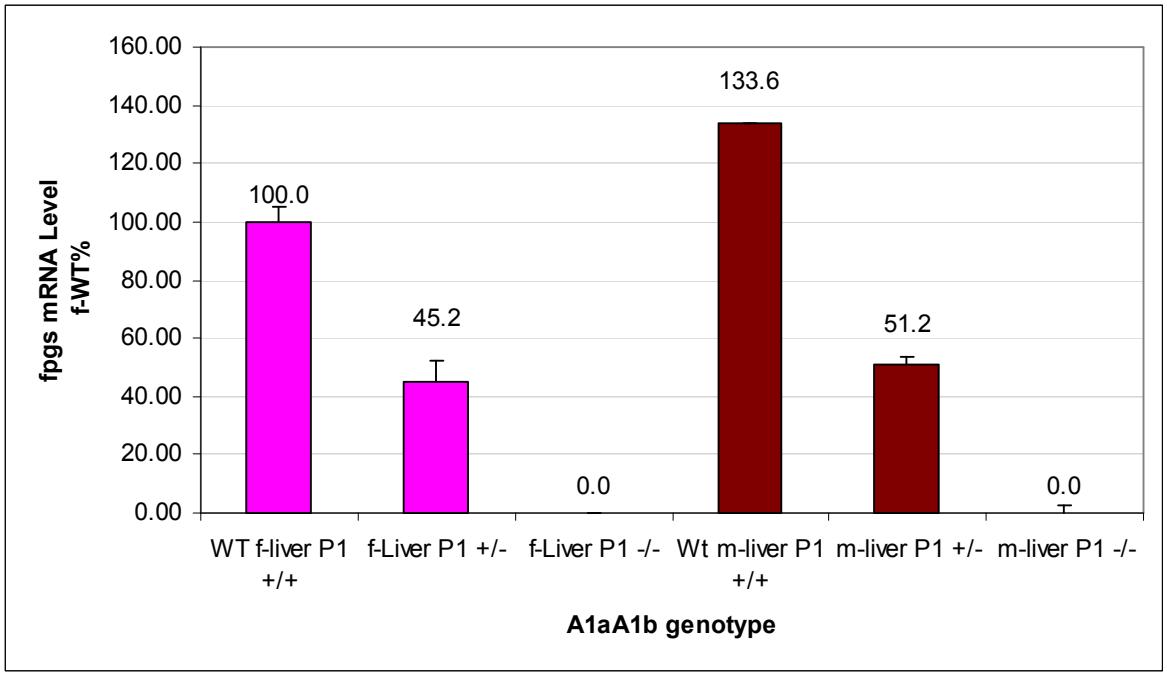


Fig. 30. Comparison of relative *fpgs* P1 expression in male and female mice liver of three A1aA1b genotypes by real-time qPCR. the P1 transcription level was 0.33 fold higher in wild-type male than female mice, 0.12 fold higher in heterozygous knockout (+/-) male than female mice. Bars show the mean ±STD from 3 mice (3male or 3 female) liver.

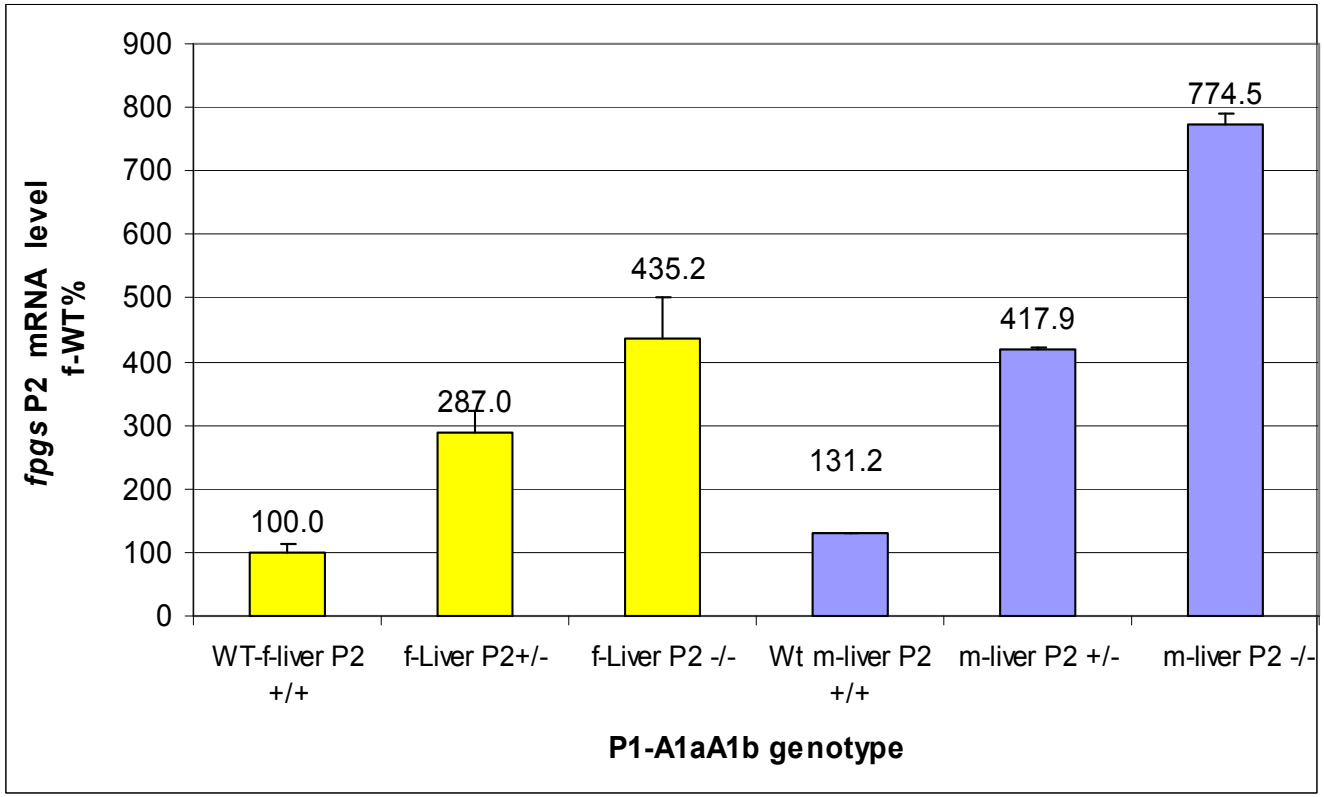


Fig. 31. Comparison of relative *fpgs* P2 expression in male and female mice liver of three A1aA1b genotypes by real-time qPCR. the P2 transcription level was 0.33 fold higher in wild-type male than female mice, 0.45 fold higher in heterozygous knockout (+/-) male mice liver than female, and was 0.77 fold higher in homozygous KO male mice liver than female. Bars show the mean ±STD from 3 mice (3male or 3 female) kidney.

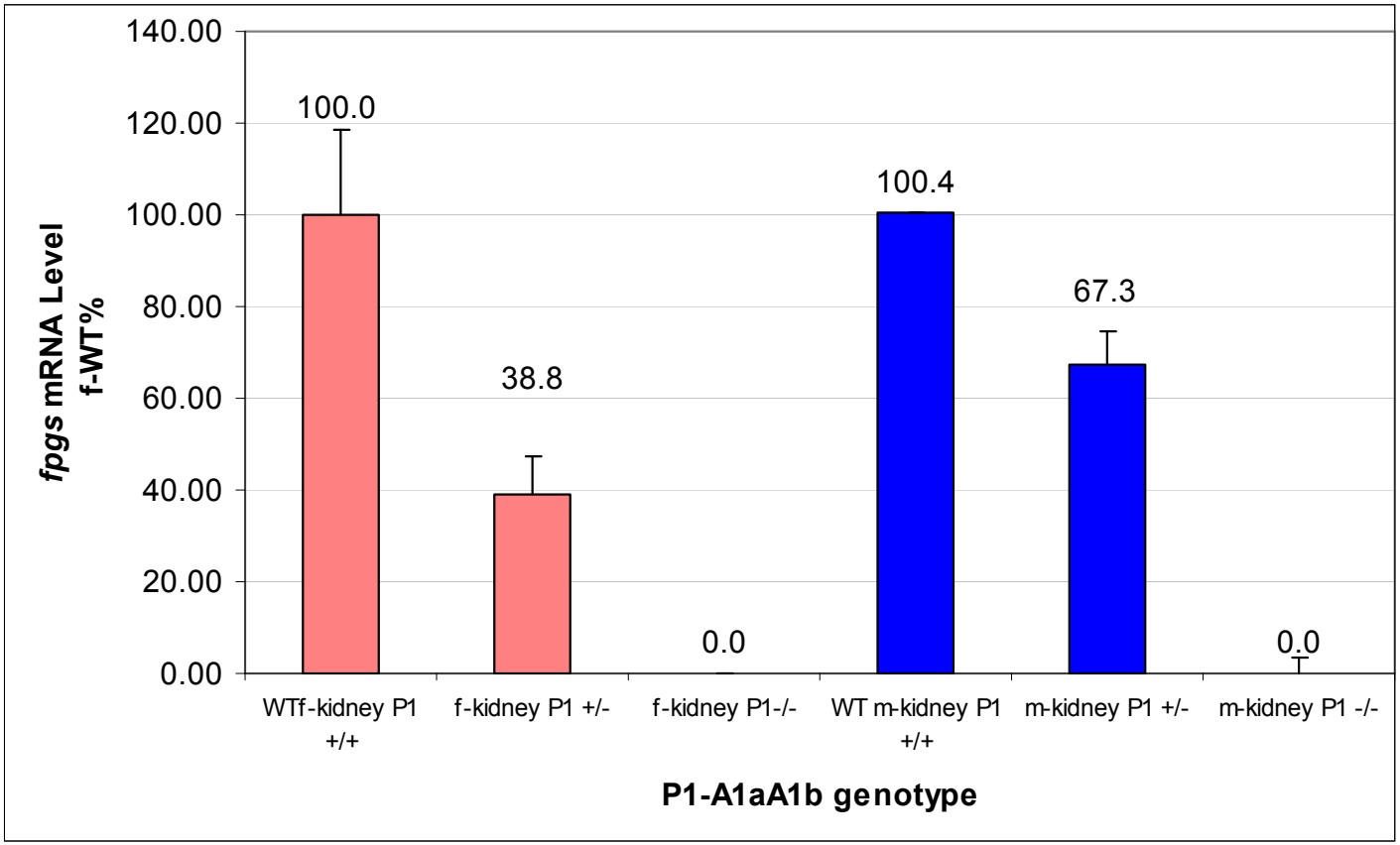


Fig. 32. Comparison of relative *fpgs* P1 expression in male and female mice kidney of three A1aA1b genotypes and determined by real-time qPCR. the P1 transcription level was same in wild-type male than female mice, 0.73 fold higher in heterozygous knockout (+/-) male than female mice. Bars show the mean ±STD from 3 mice (3male or 3 female) kidney.

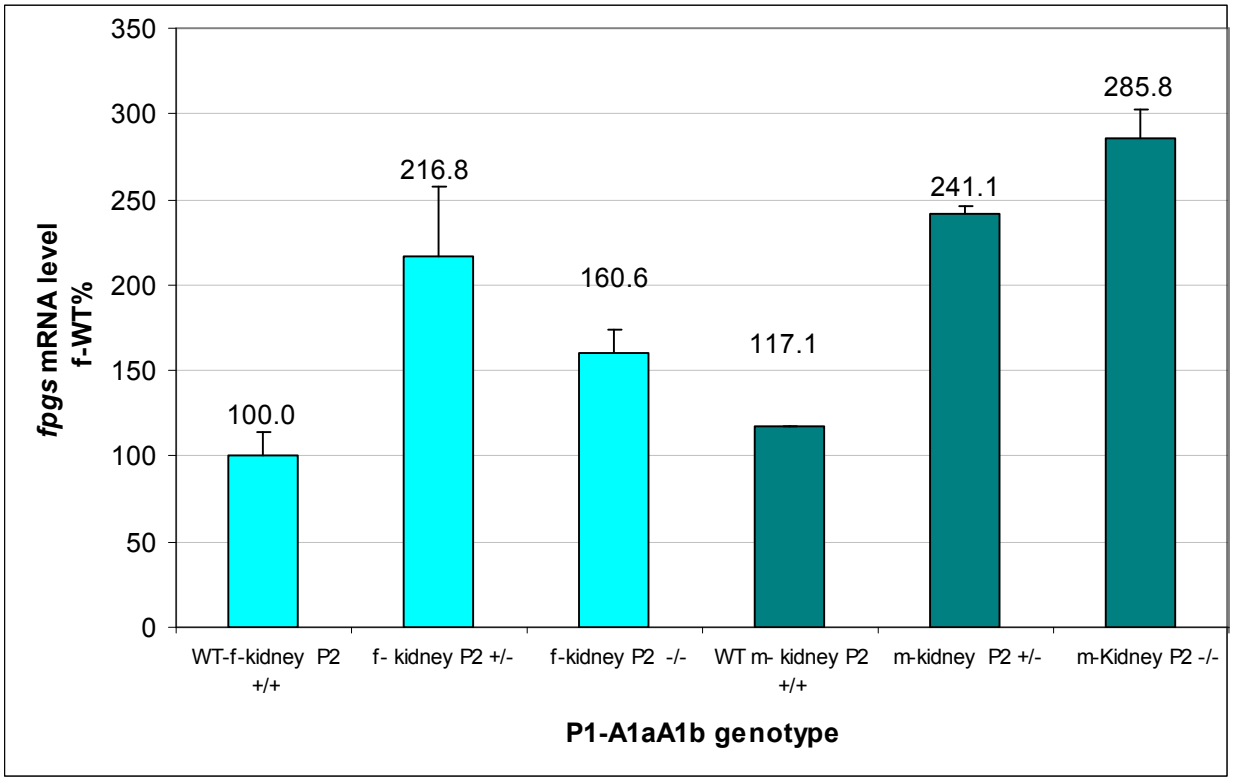


Fig. 33. Comparison of relative *fpgs* P2 expression in male and female mice kidney of three A1aA1b genotypes by real-time qPCR. the P2 transcription level was 0.17 fold higher in wild-type male than female mice, 0.11 fold increase in heterozygous knockout (+/-) male mice kidney than female, and was 0.77 fold higher in KO male mouse kidney than female. Bars show the mean ±STD from 3 mice (3male or 3 female) kidney.

Effect of *fpgs* P1 promoter deletion on liver FPGS enzyme activity

The level of *fpgs* gene expression was dramatically decreased in homozygous knockout mouse liver and kidney compared to the wild-type mouse. However, the mice had survived with this major modification. To determine whether deletion of P1 and activation of P2 would produce sufficient functional FPGS for survival, FPGS enzyme activity was measured with an assay developed in this laboratory. The mouse livers were perfused with 1x PBS and a 100,000xg supernatant produced from livers. The glutamic acid present in the liver tissue was removed with a Sephadex G-50 spin column, and the supernatant fraction of liver homogenate was then used as a source of FPGS. About 13 ug of crude lysate was incubated with a saturated amount of H4PteGlu and H^3-glutamic acid substrates. The C57BL/6 mouse liver crude lysate was also included in the experiment and use as the assay control. Our result indicated that the FPGS enzyme activity in our produced wild-type littermates mouse is similar to its breeding strain C57BL/6, and the FPGS enzyme activity was about 3 fold lower in P1 homozygous knockout mouse liver compared to wild type mouse liver (Fig 34) and 2 fold lower in the heterozygous knockout mice. But, the survival and normal growing curve of the knockout mice indicated that these amounts of FPGS enzyme were enough for the survival of the mouse.

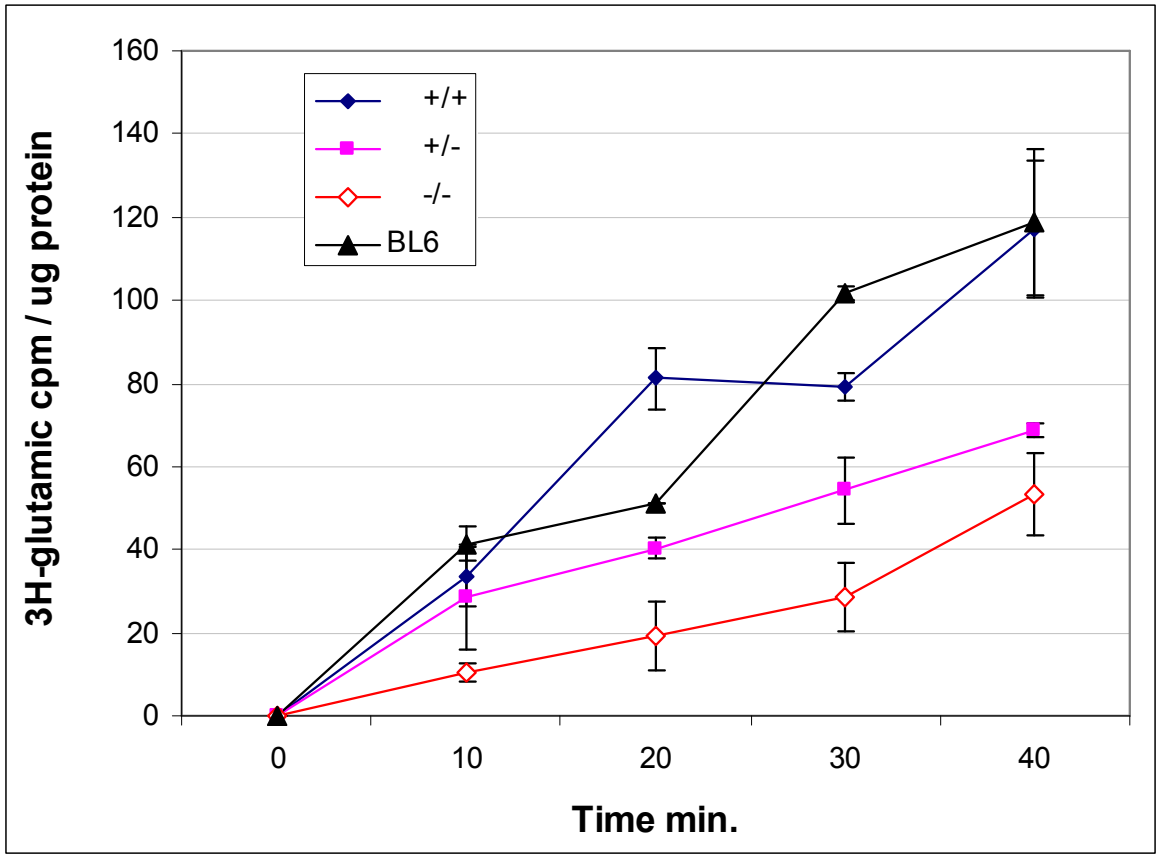


Fig 34. Comparison of liver FPGS enzyme activity in three genotypes (+/+, +/- and -/-) . FPGS enzyme activity was determined by micro-assay with 13ug liver crude lysate protein and saturating H_4PteGlu and L-[H^3]-Glutamic acid substrate. The mouse genotype were indicated in the box. Crude lysate of C57BL/6 liver were used as wild type control.

CHAPTER 4: Discussion

Gene targeting is the introduction of a modification at a specific location in the genome by homologous recombination. The frequency of homologous recombination is a linear function of length of homology between DNAs *(Fujitani,1995)*. Many research studies have shown the direct relationship between homology length and targeting frequency, and found that homology of only 1 kb is sufficient *(Kamisugi, 2005)*. In the conventional gene targeting method, the vector contains two regions of homology to the target gene located on either side of a positive selectable marker replacing the intermediate exon sequence in between. In our conditional gene targeting, the vector was designed similar to the conventional gene targeting strategy, but, instead of deletion of promoter and exons during construction of a targeting vector, the critical promoter P1 and exons A1a and A1b were flanked with two loxP sites in the targeting vector for deletion at a later step. This is a very powerful approach for potentially lethal genomic deletion. However, it also carries a risk, since it results in the inclusion of an extra 2.6 kb potential homologous recombination region. The results of sequencing analysis across the loxP1, loxP2, and loxP3 region on six potential positive targeted ES clones indicated that 50% of our targeted ES cell clones lost loxP1 site during recombination events. In these targeted ES clones, the flanked region (P1, A1a and A1b) was used for homologous recombination instead of our selected 5' homologous recombination region (Fig. 10d). The competition between the designed homologous recombination region and the flanking gene sequences for homologous recombination was obvious in retrospect, but initially was an unpleasant surprise. Over 1kb long of flanking target

sequence in the condition knockout targeting vector construct certainly created a potential risk for an unwanted homologous recombination event, resulting in the loss of *LoxP*1(Fig 10). To eliminate such incidence the detailed sequencing analysis across all three *loxP* sites of targeted locus in potential targeting ES cell clone was a necessary step, although only *loxP*1 poses a serious threat of unintended loss.

The mouse expresses two FPGS isoforms; the P1 transcript was expressed only in a few differentiated tissue, normally liver and kidney, whereas the P2 transcript was expressed in normal and malignant dividing tissues. This dual promoter mechanism directing expression of mouse isoforms is not conserved in humans. The human *fpgs* P2 transcript constitutes ~ 98% of *fpgs* transcripts in differentiated tissues and, as in the mouse, everything in dividing cells *(Turner and Moran. 2000)*. In this study, we attempted to determine why the two promoter system is used in the mouse. Will the deletion of the upstream P1 promoter activate or up-regulate the P2 transcription in mouse liver and kidney? Is the two promoter system essential for mouse development? Here, we report the deletion of *fpgs* P1 *in vivo*. We have generated a mouse that only expresses P2 *fpgs* transcripts in all tissues, similar to the situation in humans. The previous study of *fpgs* gene transcription in mouse tissues using RNAse protection assays had demonstrated that the P2 transcript was present in most mouse dividing tissues and the *fpgs* transcript found in mouse liver was almost exclusively transcribed from the P1 promoter, with an estimated 3-5% of *fpgs* transcripts initiating from the downstream promoter. In this work, we refine this estimate to <1%. Hence, two

possible outcomes were considered when we started this work; 1.) The deletion of P1 promoter and its exons may cause embryonic lethality, because the deletion of the upstream promoter and exons may not be sufficient to activate the P2 promoter. 2.) The deletion of P1 promoter and its exons may release the promoter interference previously shown to operate in this gene (*Racanelli et al. 2008*), hence, activating the P2 promoter enough that the P2 mRNA expression level will reach the level of that is expressed in mouse dividing cells. To our surprise, homozygous P1 knockout mice are viable and develop to adult mice and do not show any obvious impairments compared to wild type littermates with respect to growth, weight, fertility and longevity. However, *fpgs* gene expression from P2 in P1 knockout mouse liver demonstrated up-regulation of P2 transcription only by 4.6 fold compared to P2 expression in wild-type mouse. It was surprising that the mice survived.

The previous study of chromatin immunoprecipitation mapping over the two mouse *fpgs* promoters in our laboratory in different tissues had demonstrated that the mouse *fpgs* gene used epigenetic marking of the P1 promoter and occlusion of the down stream promoter to ensure the tissue specific usage of each of two *fpgs* isoforms (*Racanelli et al. 2008*). In mouse liver, a promoter interference effect was found to result from the activation of P1 in mouse liver which prevented the transcriptional activation of P2. This effect appeared to be due to the residence of RNAPII elongation complexes covering the nucleosome-deficient DNase hypersensitive region before P2, hence, physically occluding initiation at the P2 promoter (*Racanelli et al. 2008*). The

results of this study also suggest an interaction between the regulation of the two *fpg*s gene promoters; deletion of upstream promoter P1 and exons A1a and A1b released the apparent promoter interference mechanism and activated the downstream promoter. The level of P2 *fpgs* transcription in homozygous P1 knockout mouse liver was found to be 15-100 folds lower than P1 transcription in liver of the wild type mouse. Despite the low *fpgs* mRNA expression level in liver and kidney, the homozygous knockout mouse shows no phenotypic alteration compared with its wild-type littermates and there were no major developmental aberrations noted thus far. How do the homozygous knockout mice survive under the circumstance of such a low level of fpgs gene expression in liver? FPGS enzyme activity was measured in liver of knockout mice, heterozygous knockout mice and wild-type mice. Interestingly, FPGS protein activity appeared to be reduced 30% in the heterozygous knockout and more than 50% reduced in homozygous knockout compared to wild-type mice (Fig. 34), however, abundant enzyme activity remained. These results may indicate that *fpgs* mRNA expression is not correlated quantitatively to its protein expression in mouse liver, given that the P1 mRNA was expressed in wild-type mice liver at a much higher level (100 fold) compared to P2 mRNA expressed in knockout mice liver. Compared with this 15-100-fold difference in mRNA levels, FPGS enzyme activity in this tissue appeared to be only 2 fold lower than those in the liver of P1 knockout mice than in wild type controls (Fig. 24, 41). These differences may be caused by the artifacts in the enzyme assay. However, alternative promoter usage resulting in alternative leader exons is known to affect the

level of gene expression, the stability of the mRNA, the translational efficiency of the mRNA, and affect the structure of the amino terminus of the protein encode by the gene and consequence affect the protein stability (*Torik. 1996).* All these circumstances may also be the explanation of unexpected FPGS enzyme activity in the knockout mouse liver and the survival of homozygous knockout mice. These hypotheses remain to be adequately tested. However, the survival of P1 knockout mice may also be the consequence of the high folate content of the diet used in animal facility. A folate depletion study is currently ongoing.

In summary, we have successfully generated a mouse reflected human folate metabolism much closer than that in wild-type mice. It may become a useful tool for the investigation of the function of human FPGS that plays a critical role in the folate metabolism and anticancer therapeutics. The FPGS- humanized mouse liver model would also be an appropriate *in vivo* tool for the study of the antifolate drug toxicity and inhibition.

Reference

Adjei AA. Pemetrexed (2003): a novel multitargeted antifolate agent. Expert Rev Anticancer Ther; 3: 145–155.

Alcántara R., Ast V., Axelsen K.B., Darsow M., de Matos P., Ennis M., Morgat A. and Degtyarenko K. (2007) IntEnz Nucleic Acids Res. Molecular Biology Database Collection entry number 508.

Andreassi, J. L., II, and R. G. Moran. (2002) Mouse folylpoly-gamma-glutamate synthetase isoforms respond differently to feedback inhibition by folylpolyglutamate cofactors. Biochemistry 41:226–235

Anguera,M.C. et al. (2006) Regulation of folate-mediated one-carbon metabolism by 10-formyltetrahydrofolate dehydrogenase. J. Biol. Chem., 281:18335–18342.

Antonsson B, Barredo J, Moran RC: (1990) A microassay for mammalian folylpolyglutamate synthetase. Anal Biochem 186:8,

Antony, AC. (1992). The biological chemistry of folate receptors. Blood 79: 2807– 20

Antony AC. (1996). Folate receptors. Annu. Rev. Nutr. 16: 501– 21.

Appling, D. R. (1991). Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. FASEB Journal, 5: 2645–2651.

Assaraf, Y. G. (2006) The role of multidrug resistance efflux transporters in antifolate resistance and folate homeostasis. Drug Resistance Updates, 9:227–246.

Bradford MM: (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248,

Crawley, J N ., Belknap, J K., Collins, A ., Crabbe, J C., Frankel, ., Henderson, N., Hitzemann, R J., Maxson, S C., Miner, L L., Silva, A J., Wehner, J M., Wynshaw-Boris, A., and Paylor , R., (1997), Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. Psychopharmacology 132 :107-124

Christine T.N. Phan, Derbra M. Maclvor, Bruce A . Hug, Jonathan W.Heusel, and Timothy J. Ley. (1996) Long-range disruption of gene expression by a selectable marker cassette. Proc. Natl. Acad. Sci. Vol.93:13090-13095,

Deng,C and Capecchi, M.R. (1992) Reexamination of gene targeting frequency as a function of the extent of homology between the vector and the target locus. Mol. Cell. Biol. 12: 3365-3371

Donehower La, Harvey M, Vogel H, et al.(1995) Effects of genetic background on tumorigenesis in p53-deficient mice. Mol Carcinog. Sep; 14 (1): 16-22.

Eveland, S. S., Pompliano, D. L. & Anderson, M. S. (1997). Conditionally lethal Escherichia coli murein mutants contain point defects that map to regions conserved among murein and folylpoly-g-glutamate ligases: identi®cation of a ligase superfamily. Biochemistry, 36 : 6223-6229.

Festing, M., (1996). Inbred strains of mice. Mouse Genome 94 : 523–677

Fleischmann, A., Darsow, M., Degtyarenko, K., Fleischmann, W., Boyce, S., Axelsen, K.B., Bairoch, A., Schomburg, D., Tipton, K.F. and Apweiler, R. (2004) IntEnz, the integrated relational enzyme database. Nucleic Acids Res. 32: D434-D437.

Freemantle, S. J.; Taylor, S. M.; Krystal, G.; Moran, R. G. (1995) Upstream organization of and multiple transcripts from the human folylpoly-gamma-glutamate synthetase gene. J. Biol. Chem. 270: 9579-9584,

Fujitani, Y., Yamamoto, K. and Kobayashi, I. (1995) Dependence of Frequency of Homologous Recombination on the Homology Length. Genetics, Vol 140 : 797-809

Gerlai, R., 1996. Gene-targeting studies of mammalian behavior: is it the mutation or the background genotype? Trends Neurosci. 19 : 177–181

Gossen M, Bujard H: (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci 89: 5547–5551.

Henderson GI, Perez T, Schenker S, Mackins J, Antony AC. 1995. Maternal-to-fetal transfer of 5-methyltetrahydrofolate by the perfused human placental cotyledon: evidence for a concentrative role by placental folate receptors in fetal folate delivery. J. Lab. Clin. Med. 126: 184– 203

Kamisugi, Y., Cuming, A. C., Cove, D. J. (2005). Parameters determining the efficiency of gene targeting in the moss Physcomitrella patens. Nucleic Acids Res 33: 173-173

Kilby, N.J., Snaith, M.R., Murray, J.A., (1993). Site-specific recombinase: tools for genome engineering. Trends Genet. 9 : 413–421

Kim YI. (2004). Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? Cancer Epidemiol Biomarkers Prev.;13:511–519.

Lynn B. Bailey (1994) Folate in Health and Disease. CRC Press,

Lobe, C. and Nagy, A. (1998) Conditional genome alteration in mice. BioEssays 20 : 200–208

Marion C. Dickson, Julie S. Martin , Frances M. Cousins, Ashok B. Kulkarni, Stefan Karlsson and Rosemary J. Akhurst, (1995) Defective haematopoiesis and vasculogenesis in transforming growth factor-b1 knock out mice. Development 121: 1845-1854

Marth, J. D., (1996) Recent advances in gene mutagenesis by site-directed recombination.J. Clin. Invest. 97 : 1999–2002.

Martin J. Tymms, Ismail Kola. Gene knockout protocols, Humana press. Methods in Molecular Biology Volume 158

Metzger, D. and Chambon, P.(2001) Site- and Time – specific gene targeting in the mouse. Methods 24:71-80

McCarrick , J.W. III, Parnes, J R. Seong, R.H. (1993) Positive-Negative selection gene targeting with the diphtheria toxin A-chain gene in mouse embryonic stem cells. Transgenic Res. 2: 183-190

Moran RG, Werkheiser WC, Zakrzewski SF. (1976) Folate metabolism in mammalian cells in culture. I. Partial characterization of the folate derivatives present in L1210 mouse leukemia cells. J Biol Chem; 251: 3569–75.

Morrison, R.S., Wenzel, H.J., Kinoshita, Y., Robbins, C.A., Donehower, L.A. and Schwartzkroin, P.A., (1996) Loss of the p53 tumor suppressor gene protects neurons from kainate-induced cell death. J. Neurosci. 16: 1337–1345

Orlando C, Pinzani P & Pazzagli M 1998 Developments in quantitative PCR. Clinical Chemistry Laboratory Medicine 36:255–269

Olson, E. N., Arnold, H.-H., Rigby, P. W. J. & Wold, B. J. (1996) Role of a locus control region in the mutually exclusive expression of human red and green cone pigment genes. Cell 84: 1–4.

Perry J & Chanarin I (1973) Formylation of folates as a step in physiological folate absorption. British Medical Journal ii, 58–59.

Postic, C and Magnuson, M.A., (2000) DNA Excision in Liver by an Albumin-Cre Transgene Occurs Progressively With Age genesis 26:149–150

Qi, H., Atkinson, I., Xiao, S., Choi, Y., Tobimatsu, T., and Shane, B. (1999) Folylpoly-gamma-glutamate synthetase: generation of isozymes and the role in one carbon metabolism and antifolate cytotoxicity. AdV. Enz. Regul. 39:263- 273.

Racanelli AC., Turner, FB., Xie, LY., Taylor, S M., and Moran, R G.: (2008) A mouse gene that coordinates epigenetic control and transcriptional interference to achieve tissue-specific expression. Mol. Cell. Biol. Vol.28 : 836-848,

Rajewsky, K., Gu, H., Kühn, R. et al. (1996) Conditional gene targeting.J. Clin. Invest. 98 : 600–603.

Roy, K., Mitsugi, K, and Sirotnak, F. M. (1997) J. Biol. Chem. 272 : 5587-5593.

Sauer, B., and N. Henderson. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proc. Natl. Acad. Sci. 85 : 5166–5170.

Schalkwyk, L.C. et al. (2007) Interpretation of knockout experiments: the congenic footprint. Genes, Brain and behavior. 6: 299-303.

Schirch V, Strong WB. (1989) Interaction of folylpolyglutamates with enzymes in one-carbon metabolism. Arch Biochem Biophys; 269: 371–80.

Selhub J, Dhar GJ & Rosenberg IH (1983) Gastrointestinal absorption of folates and antifolates. Pharmacology and Therapeutics 20 : 397–418.

Sierra, E. E., & Goldman, I. D. (1998). Characterization of folate
transport mediated by a low pH route in mouse L1210 leukemia cells with defective reduced folate carrier function. Biochemical Pharmacology, 55 : 1505–1512.

Shih C, Habeck LL, Mendelsohn LG, Chen VJ, Schultz RM. (1998) Multiple folate enzyme inhibition: mechanism of a novel pyrrolopyrimidine-based antifolate LY231514 (MTA). Adv Enzyme Regul; 38 : 135–152

Simpson, E. M., Linder, C. C., Sargent, E.E.,(1997) Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. Nat. Gene. 16 : 19-27

Sirotnak FM, Tolner B. (1999). Carrier-mediated membrane transport of folates in mammalian cells. Annu. Rev. Nutr. 19 : 92–122

Smith C.M. (2000). Technical Knockout. The Scientist 14(15): 32

Subramanian VS, Marchant JS, Said HM. (2004). Apical membrane targeting and trafficking of the human proton-coupled transporter in polarized epithelia. Am J Physiol Cell Physiol 294:C233–C240,

Sun X. , Bognar A.L. , Baker E.N. , Smith C.A. (1998) Structural homologies with ATP- and folate-binding enzymes in the crystal structure of folylpolyglutamate synthetase. Proc. Natl. Acad. Sci. 95 : 6647-6652

Sun, X., Cross, J. A., Bognar, A. L., Baker, E. N., and Smith, C.A. (2001) Folate-binding Triggers the Activation of Folylpolyglutamate Synthetase J. Mol. Biol. 310 : 107-1078.

Synold, TW; Willits, EM; Barredo, JC. (1996) Role of folylpolygutamate synthetase (FPGS) in antifolate chemotherapy; a biochemical and clinical update. Leuk Lymphoma. pp 21:9–158.

Takimoto, C.H. (1996) New antifolate: pharmacology and clinical applications. The Oncologist 1:68-81

Tan, X. J., and Carlson, H.A.,.(2005) Docking Studies and Ligand recognition in Folypolyglutamate Synthetase. J. Med. Chem., 48: 7764-7772

Torik A. Y. Ayoubi and Wim J.M. Van DeVen . (1996) Regulation of gene expression by alternative promoters. FASEB Vol. 10 March

Turner, F. B., Andreassi, J. L., II, Ferguson, J., Titus, S., Tse, A., Taylor, S. M., and Moran, R. G. (1999) Tissue-specific Expression of Functional Isoforms of Mouse Folylpoly-g-glutamate Synthetase: A Basis for Targeting Folate Antimetabolites Cancer Res. 59 : 6074-6079

Turner FB, Taylor SM, Moran RG: (2000) Expression patterns of the multiple transcripts from the folylpolyglutamate synthetase gene in human leukemias and normal differentiated tissues.J Biol Chem, 275(46) :35960-35968.

Ulrike Muller. (1999) Ten years of gene targeting : targeted mouse mutants, from vector design to phenotype analysis. Mechanisms of Development , 82, issues 1-2 :3-21

Waldman, A.S. and Liskay, R.M. (1988) Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. Mol. Cell. Bio. 8:5350-5357

Wong, G.T. (2002) Speed congenics: applications for transgenic and knock-out mouse strains. Neuropeptides 36 (2-3): 230-236

Wu, L. L.; Wu, J.; Hunt, S. C.; James, B. C.; Vincent, G. M.; Williams, R. R.; Hopkins, P. N. (1994.), Plasma homocyst(e)ine as a risk factor for early familial coronary artery disease. Clin. Chem. 40: 552-561,

Zhao, R., Gao, F., Wang, Y., Diaz, G. A., Gelb, B. D., &Goldman, I. D. (2001). Impact of the reduced folate carrier on the accumulation of active thiamine metabolites in murine leukemia cells. Journal of Biological Chemistry, 276 :1114–1118.

APPENDIX A

Schematic diagram of pKO2lxFPGS-A1aA1b conditional KO targeting vector construction