Title Page

Competition between DNA methylation, nucleotide synthesis and anti-oxidation in cancer versus normal tissues

Sha Cao[1,2]*, Xiwen Zhu[1,3]*, Chi Zhang[4], Hong Qian[5], Heinz-Bernd Schuttler[6], Jianping Gong[3], Ying Xu[1,7+]

[1]Computational Systems Biology Lab, Department of Biochemistry and Molecular Biology, and [2]Department of Statistics, University of Georgia, Athens, GA, [3]Department of Hepatobiliary Surgery, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China, [4]Department of Medical and Molecular Genetics, Indiana University, Indianapolis, IN, [5]Department of Applied Mathematics, University of Washington, Seattle, WA, [6]Department of Physics & Astronomy, University of Georgia, Athens, GA, [7]College of Computer Science and Technology and School of Public Health, Jilin University, Changchun, Jilin, China

+Correspondence to: Ying Xu (xyn@uga.edu)

* These authors have equal contributions to the study.

Running title: DNA hypomethylation as a result of resource competition

Keywords: DNA hypo-methylation; Nucleotide synthesis; Anti-oxidation response; Ordinary differential equations; Systems biology

Authors' additional information:

Ying Xu, Department of Biochemistry and Molecular Biology, A110 Life Science Building; 120 Green Street, University of Georgia, Athens, GA 30602-7229; 706-542-9779; Correspondence to: xyn@uga.edu

The authors declare no potential conflicts of interest.

This is the author's manuscript of the article published in final edited form as:

Cao, S., Zhu, X., Zhang, C., Qian, H., Schuttler, H.-B., Gong, J., & Xu, Y. (2017). Competition between DNA Methylation, Nucleotide Synthesis, and Antioxidation in Cancer versus Normal Tissues. Cancer Research, 77(15), 4185–4195. https://doi.org/10.1158/0008-5472.CAN-17-0262

Abstract

Global DNA hypomethylation occurs in many cancer types but there is no explanation for its differential occurance or possible impact on cancer cell physiology. Here we address these issues with a computational study of genome-scale DNA methylation in 16 cancer types. Specifically, we identified (1) a possible determinant for global DNA methylation in cancer cells, and (2) a relationship between levels of DNA methylation, nucleotide synthesis and intracellular oxidative stress in cells. We developed a system of kinetic equations to capture the metabolic relations among DNA methylation, nucleotide synthesis, and anti-oxidative stress response, including their competitions for methyl and sulfur groups, based on known information about one-carbon metabolism and trans-sulfuration pathways. We observed a kinetic-based regulatory mechanism that controls reaction rates of the three competing processes when their shared resources are limited, particularly when the nucleotide synthesis rates or oxidative states are high. The combination of this regulatory mechanism and the need for rapid nucleotide synthesis, as well as high production of glutathione dictated by cancer-driving forces, led to the nearly universal observations of reduced global DNA methylation in cancer. Our model provides a natural explanation for differential global DNA methylation levels across cancer types and support the observation that more malignant cancers tend to exhibit reduced DNA methylation levels. Insights obtained from this work provide useful information about the complexities of cancer due to interplays among competing, dynamic biological processes.

Introduction

It has been widely observed that cancer genomes tend to have increased DNA methylation levels in the promoter regions of their protein-encoding genes, but intriguingly their global methylation levels (including promoter and non-promoter regions) tend to decrease in comparison with normal tissue cells. This has been observed in a variety of cancer types, including colon, liver, gastric, ovarian, breast, thyroid, and lung cancer (1-4). Certain cancers can have over 50% reduction *vs.* their normal controls as observed in human primary GBMs and glioma cell lines (5). It has been speculated that such reduced levels of global DNA methylation might have been selected to increase opportunities for the host cancer cells to overcome or to adapt to specific stresses encountered as reduced methylations generally imply increased gene expressions (6).

A number of studies have been published aiming to explain the possible causes for the altered DNA methylation levels in cancer, predominantly on tumor suppressor genes. A popular view has been that increased methylation in the promoter regions of such genes will keep the expressions of these genes low, hence enabling the survival of the cancerous host cells (7,8). A few studies have suggested possible causes of the reduced global DNA methylation in cancer (9). One proposal is that dietary deficiency in methyl carriers (folate and methionine) could be a reason (8,10) as insufficient methyl groups in diet have been linked to hepatic steatosis, cirrhosis, and even hepatic tumorigenesis (10). It was speculated that methyl-carrying molecules entering the one-carbon metabolism might be preferentially directed towards *de novo* synthesis of thymidylate needed for nucleotide synthesis at the expense of homocysteine re-methylation, hence resulting in reduced global DNA methylation during folate deficiency (11). It is noteworthy that the one-carbon metabolism has three exits: the folate cycle, the methionine cycle and the transsulfuration pathway with the latter going to the production of glutathione (GSH), the main antioxidant in human cells. Since cancer cells generally have high oxidative states, they tend to produce more GSH molecules for anti-oxidation and survival, therefore fewer homocysteine molecules will be directed towards DNA methylation compared to normal controls, hence resulting in reduced global DNA methylations as another proposal suggests (12). All these proposals suggest that the reduced DNA methylation may be the result of competition among several processes, including nucleotide synthesis and anti-oxidation, for their shared

resources but they are all speculative without detailed mechanistic understanding that offers a reliable explanation of why other processes may outcompete DNA methylation and how this may be regulated.

We present here a computational study to address both the *why* and the *how* questions through (i) quantitative analyses of the well-established competitive relations among the three aforementioned processes using a set of ordinary differential equations; and (ii) an integrative analysis of epigenomic and transcriptomic data of cancer *vs.* control tissues of 16 cancer types to derive cancer type specific relations between DNA methylation and its competing processes. Specifically, we answer: (1) what may have caused the reduced level of global DNA methylation in cancer? And (2) why are the global DNA methylation levels distinct across different cancer types?

We first introduce some basics of the metabolic pathways under study. S-adenosyl methionine (SAM) is the most essential compound for DNA methylation, which consists of a *sulfur-containing group* and a *methyl group*. The sulfur-containing group comes from amino acid methionine, and the methyl group is from amino acid serine carried by folate. While DNA methylation consumes methyl groups, the sulfur-containing group is recycled back to methionine or converted to a GSH precursor. Figure 1 shows the detailed pathways for the methionine, folate and GSH metabolisms and their relations, through which metabolites *S*-adenosylhomocysteine (SAH), SAM, methionine and Hcy can be inter-converted. Since folate and GSH pathways do not interact directly (Figure 1), we study them separately and call the sub-system consisting of the folate and methionine pathways as F-M and the other with GSH and methionine pathways as G-M. Throughout the paper, we use transsulfuration pathway, anti-oxidation system and GSH pathways interchangeably. In addition, DNA methylation means global level DNA methylation, unless otherwise specified.

Materials and Methods

Data

DNA methylation data measured using HumanMehtylation450 arrays for 14 cancer types and 40 bisulfide sequencing data for eight cancer types are retrieved from the TCGA database. We have also used RNA-seq based

gene expression data for 14 cancer types from the TCGA database. Six of these have all three data types available and the numbers of samples available for these three data types are summarized in Supplementary Table S1.

Method for estimating array-based DNA methylation level

The methylation array data (the HumanMehtylation450) we used cover ~485,000 CpG probes. These probes fall uniquely into one of the six categories: TSS1500, TSS200, 5' UTR, first exon, gene body, and 3' UTR, as summarized in Supplementary Table S2. Approximately half of the CpG probes are in more than one category. We have estimated each sample's total methylation level in each of the six CpG categories by summarizing the beta values of all the CpG islands that are located in each of each category, where a beta value is defined as the ratio between the methylated probe intensity and the overall intensity (sum of methylated and un-methylated probe intensities).

Method for estimating sequence-based DNA methylation level

For bisulfide sequencing data, we estimated the total methylation level of the CpG regions located in both gene regions and the LINE-1 regions in the genomes of 39 cancer patients the same way as in (13). The total methylation level in genes is estimated by averaging the methylation levels across all CpG sites that are located within gene bodies. The total methylation level of gene promoter region is estimated by averaging the methylation levels across CpG sites that are located within 2Kb upstream of gene's transcription starting site. The total methylation level of the LINE-1 elements is estimated by averaging the methylation levels across all CpG sites that are located within LINE1 elements across the whole genome.

Building equations for each reaction in nucleotide synthesis, DNA methylation and GSH synthesis

For each mono- or bi-substrate reaction under consideration (Figure 1), we built a reaction equation based on its kinetics defined by the Michaelis-Menten equation (14), where reaction rate V can be written as:

$$V = \frac{V_{max}[S]}{(K_m + [S])}$$

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

or

$$V = \frac{V_{max}[S_1][S_2]}{(K_{m,1} + [S_1])(K_{m,2} + [S_2])}$$

Reactions with non-standard forms are taken from (14) and detailed in Supplementary Table S3. All the model parameters, V_{max} and K_m, are collected from (14) (see Supplementary Table S4).

We have noted that the F-M-G system falls into the category of *stiff* systems (15) as the concentrations of different molecular species differ by several orders of magnitudes. Numerical solvers to such ODEs tend to perform poorly on such systems. We have applied the following strategy in an iterative manner by alternatively working on the F-M and G-M systems separately to derive a solution to the F-M-G system, which takes advantage of the fact that there is a natural separation in time scales for the two sub-systems and neither of them is a stiff system.

We started by determining the initial values for the 15 molecular species used in the whole system. By analytically solving for the steady-state concentrations in the F, M, G systems individually using MATLAB function "solve", we noted that 10 of the 15 species could be represented as analytical functions of the rest five. Though it was not likely that we could represent the steady state concentrations of all the 15 species using one less variable (i.e. four variables), these reduced number of variables are primarily for reducing the search space for initial values and ease of computation, and theoretically shouldn't affect the identification of the steady state concentrations. We now define the steady-state concentrations of these five species in normal tissues as x_1^0, x_2^0, x_3^0, x_4^0, x_5^0. We will search the vicinity of each of these values, specifically $\frac{x_i^0}{10}$, $\frac{x_i^0}{5}$, x_i^0, $5x_i^0$, and $10x_i^0$ for $5 \geq i \geq 1$. For each of the 5^5 combinations of initial values, we calculate the corresponding values of the 10 dependent variables, and then do the following:

1. Solve the F-M system in steady state by treating the variables in the G system as constants using their current solutions;
2. Update the current solution to variables in the F-M system based on the solution derived in (1);
3. Solve the G system in steady state by treating the variables in F-M system as constants using their current solutions;

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

4. Update current solution to variables in the G system based on the solution derived in (3);
5. Repeat Steps 2-4 until the first derivatives of all the 15 variables, defined as $\frac{\mathrm{abs}(s_{t+\mathrm{d}t}-s_t)}{dt}$, as well as their relative changes, defined as $\frac{\mathrm{abs}(s_{t+\mathrm{d}t}-s_t)}{s_t}$, are within a pre-specified threshold, 0.01.

The convergence of the algorithm for finding steady state concentrations of the F-M-G system is established in Supplementary document.

Relevant enzymes in F-M-G system and marker genes for estimating the levels of folate, serine, cysteine, and methionine

We have used known marker genes and their expression levels for estimating the level of each metabolite and enzyme used in our ordinary differential equations and analysis. Specifically, we have done the following, with detailed gene list given in Supplementary Table S5:

1. For folate, we have used expressions of genes, *FPGS*, *GGH*, *SLC19A1*, encoding the folate homeostasis mediators as a measure for its concentration (16);
2. Transporters of amino acids cysteine, serine and methionine are taken from (17).

Estimating the expression level of a group of genes in a common pathway

Obtaining an overall expression of a group of genes on multiple samples, X, is to find a one-dimension representation, d, for X, which could be formulated as an optimization problem

$$\min_{\alpha,\mathrm{d}} f(\alpha, d) = ||X - \alpha d^T||_F$$

which is the largest eigen-value of the matrix

$$\sum_{i=1}^{N} x_i^T x_i$$

where x_i is the row vector of X. Note α serves as an ancillary variable in here.

Proof:

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

$$f(\alpha, d) = \sum_{i=1}^{N} (x_i - \alpha_i d^T)(x_i - \alpha_i d^T)^T = \sum_{i=1}^{N} (x_i x_i^T - 2\alpha_i d^T x_i^T + \alpha_i{}^2 d^T d)$$

$$\frac{\partial f(\alpha, \square)}{\partial(\alpha_i)} = -2d^T x_i^T + 2\alpha_i d^T d$$

Let $\frac{\partial f(\alpha,d)}{\partial(\alpha_i)} = 0$, we have $\alpha_i^* = \frac{d^T x_i^T}{d^T d}$. We replace $f(\alpha, d)$ with α_i^*

$$f(\alpha, d) = \sum_{i=1}^{N} (x_i x_i^T - \frac{d^T x_i^T x_i d}{d^T d})$$

To minimize $f(\alpha, d)$ is equivalent to maximizing

$$\sum_{i=1}^{N} \frac{d^T x_i^T x_i d}{d^T d}$$

And that would be the eigenvector corresponding to the largest eigen-value of the matrix

$$\sum_{i=1}^{N} x_i^T x_i$$

Results

Reduced global DNA methylation in cancer *vs.* normal control

We have examined DNA methylation data of 5,219 tissue samples of cancer *vs.* control tissues of 16 cancer types, out of which 5,179 samples are measured using the array technology and 40 samples measured using the bi-sulfide sequencing technology. The two datasets each cover different cancer types, with the array and sequencing data covering 14 and 8 cancers types, respectively, giving rise to a total of 16 distinct cancer types (see **Materials and Methods**). The array data consist of methylation data of 486,428 CpG islands, covering 99% of RefSeq genes, with an average of 17 CpG sites per gene distributed across its entire promoter region (i.e., TSS1500, TSS200, 5' UTR, and the first exon) and the gene body regions (i.e., gene body and the 3' UTR). The total methylation levels of promoter and gene body regions' CpG islands across different cancers are summarized

in Supplementary Figure S1. Details regarding how the methylation level of each CpG category is estimated can be found in the **Materials and Methods** section.

We noted that cancer tissues tend to have significantly increased methylation levels in promoter regions (Figure S1A) and substantially reduced methylations in gene bodies (Figure S1B), which account for more than 80% of the CpG islands considered here. Analyses of bisulfide sequencing-based methylation data gives rise to the same conclusion that gene bodies tend to be hypo-methylated (Figure S1C) and promoter regions have increased methylation levels. In addition, we have also done similar analyses on DNA methylation in transposable elements, specifically LINE-1, which is accepted as a reliable measure for estimating the DNA methylation level at the genome scale (18); and observed similar decrease in the methylation level in this region (Figure S1D). We noted that among the 16 cancer types under study, two cancer types PRAD and BRCA have increased levels of methylations than their matching normal tissues.

In the following sections, we built mathematical models to study the detailed reasons and associated mechanisms for the observed hypo-methylation at the global scale in cancer. The units for concentration and time are uM and hr respectively.

The methionine cycle: dependencies of DNA methylation on folate and transsulfuration pathways

We have built a system of kinetic equations based on the known pathway models shown in Figure 1, to describe the reaction rates associated with 15 key molecular species in the folate, methionine and GSH (F-M-G) cycles, which consists of 24 enzyme-catalyzed reactions and four transporters, shown in Table 1 with detailed information of how each reaction equation is derived given in Supplementary Table S3.

We have examined the relationship between the DNA-methylation level and each of its four parameters, namely the levels of folate, serine, cysteine and methionine uptakes, respectively, according to this system of equations with all the four parameters having their values sampled from the normal ranges of their respective physiological values collected from (14). Specifically, we have derived the numerical solutions of the fixed points of this system of equations for each combination of the parameter values uniformly sampled from the given ranges, using the MATLAB ode solver (see **Materials and Methods**). Figure 2 shows the level of relevant

reactions or metabolites changing as a function of the levels of folate and serine concentrations (Figure 2A-2C) and as a function of the levels of cysteine and methionine concentrations (Figure 2D-2F).

We have observed that as the levels of serine (a methyl donor) and folate (a methyl carrier) increase, the DNA methylation level, nucleotide synthesis rate go up (Figure 2A-2B); and the system's total methyl amount increases with increase in folate, but not with increase in serine uptake (Figure 2C), this indicates that folate is the key factor in determining the total methyl pool, not serine. As the level of methionine (a sulfur donor) increases, the DNA methylation level, GSH synthesis and total sulfur concentration all go up (Figure 2D-2F). It is noteworthy that changes in the rate of cysteine uptake do not have a significant impact on the DNA methylation rate, since it goes into the downstream of the methionine cycle and it does not directly contribute to the sulfur supply for SAM used in DNA methylation. Here, the DNA-methylation rate, nucleotide synthesis rate and GSH synthesis rate are estimated as the reaction velocities catalyzed by enzymes *DNMT*, *TS* and *GS*, respectively (see Supplementary Table S3); the total methyl and the sulfur levels are estimated as the sum of concentrations of all metabolites that carry methyl and sulfur groups, respectively.

This analysis confirms that our system of equations captures the intuition that the level of DNA methylation should be an increasing function of the levels of methyl and sulfurs, provided by folate and serine, and methionine, respectively. In the following sections, we will study how this relationship is affected by other processes when they compete for methyl and sulfurs.

The F-M system: a model of competition for methyl between nucleotide synthesis and DNA methylation

As the ultimate donor of methyl, serine reacts with THF to generate 5,10-methylene-THF, during which methyl in serine is converted to a methylene group. Then this methylene group will go to one of three places: (1) as the methyl group of dTMP by the reaction catalyzed by thymidylate synthase (*TS*); (2) as the methyl group of 5mTHF catalyzed by *MTHFR* and ultimately for DNA methylation; and (3) to THF or 1,10-CH=THF. In our differential equations representing the F-M system, we have used the (steady state) levels of 5mTHF and DHF as estimates for the levels of methyl groups going to DNA methylation and nucleotide synthesis, respectively, since they are the immediate downstream metabolites of 5,10-methylene-THF going into the two processes. Here, we

address the following question: when competing for methyl groups between DNA methylation and nucleotide synthesis, does one process have an encoded priority over the other?

We have assessed whether the two processes may have a competitive relationship as defined by our differential equations under a condition that nucleotide synthesis must be done at a rapid rate to mimic the typical situation in cancer. Specifically, we have checked how the DNA methylation rate changes when the rate coefficient V_{max} of *TS* increases, with *TS* being the key enzyme leading to nucleotide synthesis.

Figure 3 shows that as the need for nucleotide synthesis goes up, reflected by *TS*'s V_{max} value, methyl groups going to DNA methylation decreases (Figure 3A) while those going into nucleotide synthesis increase (Figure 3B). So do the DNA methylation rate (Figure 3C) and the nucleotide synthesis rate (Figure 3D), respectively. More specifically, as V_{max} increase from the low to the high end of its normal range, methyl going to DNA methylation is decreased by 4.8% while that going to nucleotide synthesis is increased by 40.6%. Correspondingly, the DNA methylation rate is decreased by 0.5% while the nucleotide synthesis rate is increased by 36.5%. We predict that this is the result of a regulatory mechanism which controls the competition between the two processes.

To elucidate this regulatory mechanism, we have conducted the following analysis focused on three enzymes *SHMT*, *TS* and *MTHFR* forming a Y shaped branch structure in the folate cycle with *SHMT* catalyzing the reaction leading to 5,10-m THF, which then branches out to nucleotide synthesis catalyzed by *TS* and to methionine cycle by *MTHFR* (see Figure 1). In cancer tissues, the expression level and hence the V_{max} of *TS* tend to increase substantially. We first examined the reaction rate constants of the three enzymes: V_{max}= 5200 uM/hr and K_m = 600 uM for *SHMT*; V_{max}= 5000 uM/hr and K_m = 6.3 uM for *TS*; and V_{max}= 5300 uM/hr and K_m = 50 uM for *MTHFR*. Hence, the reaction rate of *TS* is close to two orders of magnitude higher than that of *SHMT*, ~91 times higher to be more exact, which will increase as the V_{max} of *TS* increases. We also noted that the substantial increase in the V_{max} of *TS* as done in the above illustrative example leads to only 2.9% increase in the concentration of THF (see Supplementary Table S6), indicating that the increase in the reaction rate of *SHMT* is limited by this number, regardless of the level of increase in the serine supply. Because of the tiny increase in this

reaction rate, we assume, for the simplicity of discussion, that the rate remains unchanged. This immediately implies that the increased reaction rate of *TS* will take away a portion of the flux to 5mTHF to meet the need of the increased *TS* reaction rate; and the higher the *TS* reaction rate is increased, the higher proportion of the flux to 5mTHF will be diverted towards the reaction catalyzed by *TS*, hence more reduced level of DNA methylation.

In sum, it is the combination of the reaction rate constants in the folate cycle, particularly of three enzymes *SHMT*, *TS* and *MTHFR* along with their relative expression levels that play the key regulatory role in governing the competition for methyl groups between nucleotide synthesis and DNA methylation.

The G-M system: a sulfur-redistribution model: redox balance *vs.* DNA methylation

SAM serves a unique role in the system under study, as it not only contains a methyl group but also a sulfur group that carries the methyl molecule. As discussed earlier, the sulfur group in homocysteine can be recycled back to the methionine cycle and further to SAM, or it can go to the GSH metabolic pathway, indicating that DNA methylation also needs to compete with the GSH pathway for sulfur, in addition to its competition with nucleotide synthesis. This pathway starts with a reaction between homocysteine and serine (Figure 1), leading to the generation of cystathionine that is then cleaved by cystathionine lyase to generate α-ketobutyrate and cysteine, which is then used for GSH production. Here we study how GSH production and DNA methylation may compete for sulfur, which ultimately affects the level of DNA methylation. Similar to the previous section, we use the (steady state) levels of two immediate downstream metabolites, methionine and cystathionine, respectively, to estimate the level of sulfurs going to DNA methylation and GSH synthesis, respectively.

We have examined how the two processes change their activity levels as the level of intracellular H_2O_2 goes up, where H_2O_2 is used to represent the oxidative state since it is the most abundant reactive oxygen species (ROS) in cancer in general (19). Generally, as the H_2O_2 level goes up, the host cells will increase their GSH production to naturalize the excess H_2O_2 to keep the oxidative stress under control, which will consume sulfurs. Here, we show how an increased demand for sulfur by GSH production affects the level of DNA methylation. We have observed as the need for anti-oxidation and hence GSH synthesis goes up (reflected by the H_2O_2 level), sulfurs going to DNA methylation decrease (Figure 3E) while sulfurs going to GSH production increase (Figure

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

3F). So do the DNA methylation rate (Figure 3G) and GSH synthesis rate (Figure 3H). More specifically, sulfur going to DNA methylation is decreased by 13.6% and those to GSH production is increased by 6.2%. Correspondingly, DNA methylation rate is decreased by 1.9% and the GSH synthesis rate is increased by 2.6%. As in the previous section, we predict that this is the result of a regulatory mechanism that controls the flux of sulfur to different branches when they are limited.

We have conducted an analysis similar to that in the previous section on four enzymes: *SAHH*, *BHMS*, *MS* and *CBS* (Figure 1), which play key roles in the regulatory mechanism under investigation. As before, we noted: V_{max}= 320 uM/hr and K_m = 6.5 uM for *SAHH;* V_{max}= 2160 uM/hr and K_m = 12 uM for *BHMT;* V_{max}= 500 uM/hr and K_m = 1 uM for *MS;* and V_{max}= 700000 uM/hr and K_m = 1000 uM for *CBS*. From the above example, we observed: as the H_2O_2 level goes up to the high end we set, the Hcy concentration goes up by 6% (see Supplementary Table S7). For the simplicity of discussion, we assume that there is no change in the Hcy concentration considering its tiny increase. All these reveal that the reaction rate of *CBS* goes up substantially and the Hcy flux into MET will go down. Knowing that the reaction rate of *CBS* is significantly higher than those of *BHMT* and *MS* based on their rate constants, we predict that the Hcy flux into MET and hence SAM will go down substantially, which is consistent with the observed change of MET in the above example, as detailed in Supplementary Table S7. In addition, H_2O_2 is known to have an inhibitory role on *MS* and *BHMT* (Figure 1). Hence the reaction rates of both *MS* and *BHMT* will go down as the H_2O_2 level goes up.

Based on the above, our prediction of the regulatory mechanism is: when the H_2O_2 concentration is not too high, its their inhibitory roles on *MS* and *BHMT* will slow down the flux towards DNA methylation, leading to the accumulation of Hcy. As the H_2O_2 concentration further goes up, Hcy concentration continues to increase. Once once the Hcy concentration is close to or exceeds the K_m value of *CBS*, the enzyme will instantly dump all the Hcy into the GSH synthesis pathway due to the very high reaction velocity of *CBS*. Overall, it is the combination of the reaction constants of the methionine cycle, particularly those of the four enzymes discussed here and the inhibitory role of H_2O_2 that controls the competition between DNA methylation and GSH synthesis.

We have also conducted a simulation analysis of the F-G-M system as a whole by systematically going through each of the 117 kinetic parameters encoded in the system by individually changing the value of each

parameter, specifically through multiplying its default value by 0.1, 0.2, 0.5, 0.6, 0.8, 1.0, 1.2, 1.5, 1.6, 1.8, 2, respectively. The goal is to determine which of these parameters are most impactful on the competitions under study. As expected, the kinetic parameters associated with *DNMT* and the uptake of folate, serine and methionine, respectively, are the most impactful, all resulting in at least 10% change in the DNA methylation level through the above parameter manipulation. Interestingly, other most impactful parameters are those associated with *MTHFR*, *MS*, *CBS* and *BHMT*, which are at the core positions of the whole system, where redistributions of methyl and sulfur happen, each of which leads to at least 5% change in the level of DNA methylation.

In sum, our analyses revealed that there are regulatory mechanisms, largely encoded in the relative levels of their enzymes' reaction rate constants, that determine how the three processes compete for two shared resources: methyl and sulfurs.

We have previously demonstrated that the rate of nucleotide synthesis in cancer is dictated by the level of cytosolic Fenton reactions, which also largely determines the level of oxidative stress (20) while the level of cytosolic Fenton reactions is predominantly determined by concentrations of H_2O_2 and iron, for whose accumulation chronic inflammation is largely responsible (21). Hence we predict that the observed genome-scale hypo-methylation in cancer is the result of these encoded regulatory mechanisms and the urgent need for rapid nucleotide synthesis and GSH synthesis, dictated by the level of Fenton reactions.

With this established framework, we address why different types of cancers may have different levels of global DNA methylation using cancer tissue gene-expression data.

Cancer specific DNA methylation

To understand why different cancer types may have different levels of hypo-methylation *vs.* their normal controls, we have conducted cancer-specific analyses of the integrated model of the above three subsystems, collectively referred to as the *F-M-G system*, through applying the observed gene-expression levels of the relevant enzymes, normalized with respect to their corresponding normal controls. Supplementary Figure S2 shows differentially expressed genes across the F-M-G system. Note that for cases where multiple genes encoding one

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

enzyme or transporter, we estimated the integrated expression level of the gene group as a whole, using a method given in **Materials and Methods**.

We have developed a model for predicting the DNA-methylation level of a given cancer tissue based on the expression levels of selected enzymes in the F-M-G system. We used the Michaelis-Menten equation to capture how the reaction velocity *V* depends upon the concentrations of the main substrate *S* and the catalyzing enzyme *E*:

$$V = \frac{V_{max}[S]}{K_m + [S]} = \frac{k_{cat}[E][S]}{K_m + [S]}$$

Here, we can reasonably assume that the relevant reaction rate constant K_m for each enzyme is the same across different cancer types. Hence, the rate of each reaction under consideration is entirely determined by $[S]$ and $[E]$. The cancer specific enzyme concentration $[E]$ would be that of normal condition multiplied by the fold change in the enzyme's gene expression levels in cancer *vs.* control tissues. For each cancer type, we estimated an "average" fold-change in expression levels of genes encoding the relevant enzymes/transporters between cancer and control samples, as described in **Materials and Methods**. Figure S2 shows changes in concentrations of the relevant enzymes and transporters across different cancers *vs.* corresponding controls. Clearly, so estimated enzyme concentrations will give rise to different steady-state concentrations of each metabolite and reaction rates, as shown in the 4th column of Table 2.

Our cancer specific DNA methylation prediction is calculated for ten cancer types, which have both methylation array data and RNA-Seq gene-expression data for cancer and control samples. For each cancer type, we have estimated the steady-state concentrations of all the relevant metabolites, particularly the DNA-methylation level using the reaction rate catalyzed by enzyme DNMT. As shown in Table 2, our predictions of the DNA methylation levels are highly consistent with the experimental data in the eight cancer types considered. For the two cancer types where our predictions are not consistent with experimental data, namely BRCA and THCA, we believe that the reason for the hypo-methylation prediction in BRCA and hyper-methylation for THCA is due to the possibility that certain factors that may also contribute to DNA methylation are not included in our model.

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

We have then studied how DNA methylation levels differ when patients have different levels of GSH synthesis and nucleotide synthesis rates. The analysis is done on the same ten cancer types as above. We noted that samples with higher nucleotide synthesis or GSH synthesis levels tend to have lower DNA methylation rates, consistent with our model that DNA methylation is at an inferior position when competing for shared resources with nucleotide synthesis and GSH synthesis. This is the case for seven out of ten cancer types but not for KIRC, PRAD and THCA as shown in Table 3. These three cancer types clearly warrant further studies in order to understand why they behave differently from the other seven cancer types. Note that the levels of GSH synthesis and nucleotide synthesis are estimated using genes involved in glutathione synthesis and RNA polymerases pathways, respectively, with details given in **Materials and Methods**.

We further studied whether DNA methylation is indeed competing with nucleotide synthesis and anti-oxidation system for methyl and sulfur. Particularly, we are interested in the relationships between DNA methylation and nucleotide synthesis/anti-oxidation capacity when methyl/sulfur is limited. To accomplish this, we have introduced a measure of the *average methyl/sulfur availability*: the ratio between nucleotide synthesis/anti-oxidation capacity and methyl/methionine availability with the property: the lower the ratio is, the lower the average availability of methyl/methionine is for the nucleotide synthesis/anti-oxidation capacity. For methyl compound, we have observed significant correlations between DNA methylation and nucleotide synthesis in those samples with high and low average methyl availability, respectively; and similarly for sulfur, we have also observed significant correlations between DNA methylation and anti-oxidant capacity in those samples with high and low average sulfur availability, respectively. Here, the levels of methyl and methionine, anti-oxidation capacity and nucleotide synthesis are estimated using genes involved in folate and methionine transporters (Supplementary Table S5), glutathione synthesis and RNA polymerases pathways, respectively, with details given in **Materials and Methods**.

As shown in Table 4, when the average methyl availability is low, most cancer types showed significant negative correlations (*p*-value cutoff: 0.05) between DNA methylation and nucleotide synthesis, except for KIRC, PRAD and THCA; and when the average sulfur availability is low, most cancer types showed significant negative correlations (p-value < 0.05) between DNA methylation and the anti-oxidant capacity, except for BRCA, KIRC,

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

PRAD and THCA. These data strongly suggest competitive relations between DNA methylation and nucleotide synthesis/anti-oxidation capacity for methyl and sulfur. It has also explained: (1) why our prediction of the methylation level in BRCA is not accurate shown in Table 2; and (2) why the DNA methylation levels do not depend on nucleotide synthesis or anti-oxidation capacity as shown in Table 4, as methyl and sulfur may be not limited resources in these cancer types, i.e., not rate-liming factor in the cancerous cell division; and DNA methylation does not need to compete for the two resources with other processes. We have also noticed that even when the average methyl/sulfur availability is high, the negative correlations are also significant (p-value < 0.05) for some cancer types. Hence, we posit that even though the average availability of methyl/sulfur is high in thee samples compared to other samples, these resources are still limited with respect to their cell division rates dictated by cytosolic Fenton reactions (see Discussion), and hence competitions are still there.

Discussion

We have previously developed a model proposing that Fenton reactions, $Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + \bullet OH$, in cytosol and mitochondria may represent key drivers of cancer initiation and progression at a more basic level than the previously proposed drivers such as genomic mutation (22), epigenomic alteration (23) and metabolic reprogramming (24). Fenton reactions have been found to take place when concentrations of Fe^{2+} and H_2O_2 are sufficiently high in the same location without involvement of any enzyme. When there are plentiful reducing elements near the reaction sites, such as sulfur, NADH or superoxide, Fe^{3+} can be reduced to Fe^{2+}, hence enabling the reaction to continue indefinitely, which is also referred to as the Harbor-Weiss reaction (25). Then the reaction can be rewritten as: $O_2^- + H_2O_2 \rightarrow \cdot OH + OH^- + O_2$ with Fe^{2+} as a catalyst since it is not consumed by the (continuous) reaction; and superoxide as the reducing element as our analysis revealed that O_2^- is the most commonly used reducing element in cancer (26). We have demonstrated statistically that all cancers in the TCGA database have Fenton reactions in their cytosol, mitochondria and extracellular matrix and space (27). An important implication of this model is that cytosolic Fenton reactions drive *de novo* nucleotide synthesis (and glycolytic ATP production) to produce net protons (H^+) at rates comparable to those of OH^--producing Fenton

reactions, hence keeping intracellular pH stable. More specifically, it is the rates of cytosolic Fenton reactions that dictate the rate of nucleotide synthesis and hence the rates of DNA synthesis and cell division in cancer.

This model, in conjunction with the discovery made here, strongly suggests that there is an encoded regulatory mechanism that determines the winners in competition for methyl and sulfurs when they are of limited availability, and it offers a natural explanation of why cancers in general have reduced genome-scale DNA methylation as well as why different cancers tend to have distinct levels of global DNA methylation, hence addressing an important and open question in cancer biology.

In addition, we have also provided an explanation as why certain cancers tend to have more reduced DNA methylations, i.e., those with higher levels of nucleotide synthesis and oxidative stresses, hence having established that the global DNA methylation level could be used as a predictor for more aggressive cancer types.

Further extension of the current study will include detailed metabolisms of serine, methionine and H_2O_2 to make the model more realistic: concentrations of the latter compounds are presently treated as input parameters, rather than treating them in a more realistic manner via explicitly modeling the ways that they are actually brought into cancer cells. In addition, the observation on BRCA and THCA were that the global methylation levels in their tumor samples do not differ significantly from their normal samples, while our model predicts that BRCA is hypo-methylated, and THCA is hyper-methylated. We believe that these discrepancies are due to the assumption in our model that only two factors contribute to the global DNA methylation level, namely the competitions for sulfur and for methyl groups with two other processes. Careful inspection of the data in Table 4 revealed that the global methylation levels have no significant negative correlation with neither of the two competing processes, namely nucleotide synthesis and anti-oxidation system in BRCA and THCA, suggesting that other factors, such as hormones, may also affect the global methylation level of DNA. We examined the global methylation levels of different subtypes of breast cancer, and noted that triple-negative breast cancers (free of hormone regulation) are significantly hypo-methylated in tumor (p-value= 0.0129), consistent with our model prediction. This suggests one possible direction for further development of our model in the future. We believe that our study here offers a good example for studying complex, non-linear relationships among multiple players involved in specific

biological processes, leading to novel understanding about previously made perplexing observations, and can be applied to study a suite of such problems in cancer research.

Acknowledgements

We thank colleagues in the UGA Computational Systems Biology Lab especially Mr. Yi Zhou, and Professor Jonathan Arnold of the UGA Genetics Department for their help throughout this project.

References

1. Ehrlich M. DNA hypomethylation in cancer cells. Epigenomics **2009**;1:239-59
2. Li J, Huang Q, Zeng F, Li W, He Z, Chen W, *et al.* The prognostic value of global DNA hypomethylation in cancer: a meta-analysis. PloS one **2014**;9:e106290
3. Kuchiba A, Iwasaki M, Ono H, Kasuga Y, Yokoyama S, Onuma H, *et al.* Global methylation levels in peripheral blood leukocyte DNA by LUMA and breast cancer: a case-control study in Japanese women. British journal of cancer **2014**;110:2765-71
4. Nojima M, Matsui T, Tamori A, Kubo S, Shirabe K, Kimura K, *et al.* Global, cancer-specific microRNA cluster hypomethylation was functionally associated with the development of non-B non-C hepatocellular carcinoma. Molecular cancer **2016**;15:31
5. Cadieux B, Ching TT, VandenBerg SR, Costello JF. Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation. Cancer research **2006**;66:8469-76
6. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer cell **2014**;26:577-90
7. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene **2002**;21:5427-40
8. Crider KS, Yang TP, Berry RJ, Bailey LB. Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role. Advances in nutrition **2012**;3:21-38
9. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene **2002**;21:5400-13
10. Davis CD, Uthus EO. DNA methylation, cancer susceptibility, and nutrient interactions. Experimental biology and medicine **2004**;229:988-95
11. Field MS, Kamynina E, Agunloye OC, Liebenthal RP, Lamarre SG, Brosnan ME, *et al.* Nuclear enrichment of folate cofactors and methylenetetrahydrofolate dehydrogenase 1 (MTHFD1) protect de

novo thymidylate biosynthesis during folate deficiency. The Journal of biological chemistry **2014**;289:29642-50

12. Tsun ZY, Possemato R. Amino acid management in cancer. Seminars in cell & developmental biology **2015**;43:22-32
13. Warden CD, Lee H, Tompkins JD, Li X, Wang C, Riggs AD, *et al.* COHCAP: an integrative genomic pipeline for single-nucleotide resolution DNA methylation analysis. Nucleic acids research **2013**;41:e117-e
14. Reed MC, Thomas RL, Pavisic J, James SJ, Ulrich CM, Nijhout HF. A mathematical model of glutathione metabolism. Theoretical biology & medical modelling **2008**;5:8
15. Higham DJ, Trefethen LN. Stiffness of odes. BIT Numerical Mathematics **1993**;33:285-303
16. Figueiredo JC, Levine AJ, Lee WH, Conti DV, Poynter JN, Campbell PT, *et al.* Genes involved with folate uptake and distribution and their association with colorectal cancer risk. Cancer causes & control : CCC **2010**;21:597-608
17. Sahoo S, Aurich MK, Jonsson JJ, Thiele I. Membrane transporters in a human genome-scale metabolic knowledgebase and their implications for disease. Frontiers in physiology **2014**;5:91
18. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic acids research **2004**;32:e38
19. Bienert GP, Schjoerring JK, Jahn TP. Membrane transport of hydrogen peroxide. Biochimica et Biophysica Acta (BBA) - Biomembranes **2006**;1758:994-1003
20. Richter Y, Fischer B. Nucleotides and inorganic phosphates as potential antioxidants. Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry **2006**;11:1063-74
21. Imlay JA, Chin SM, Linn S. Toxic DNA damage by hydrogen peroxide through the Fenton reaction in vivo and in vitro. Science **1988**;240:640-2
22. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, *et al.* Patterns of somatic mutation in human cancer genomes. Nature **2007**;446:153-8

23. De Carvalho DD, Sharma S, You JS, Su S-F, Taberlay PC, Kelly TK, *et al.* DNA methylation screening identifies driver epigenetic events of cancer cell survival. Cancer cell **2012**;21:655-67

24. Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell **2008**;134:703-7

25. Mello Filho AC, Meneghini R. In vivo formation of single-strand breaks in DNA by hydrogen peroxide is mediated by the Haber-Weiss reaction. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression **1984**;781:56-63

26. Ray G, Husain SA. Oxidants, antioxidants and carcinogenesis. Indian journal of experimental biology **2002**;40:1213-32

27. Sun H, Zhang C, Dong N, Sheng T, Xu Y. Targeting Fenton's reaction and its role in the regulation of cancer tissue's pH level: a computational approach (In preparation). http://csblbmbugaedu/~zhangchi/FentonReaction/ **2016**

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Tables:

Table 1: Reaction rates associated with all the metabolites involved in folate, methionine and GSH metabolisms, with the detailed information for each rate along with an explanation, given in Supplementary Table S3.

$$\frac{d[DHF]}{dt} = V_{TS}([dUMP],[CH2F]) - V_{DHFR}([DHF],[NADPH])$$

$$\frac{d[5mTHF]}{dt} = V_{MTHFR}([CH2F],[NADPH],[SAM]) - V_{MS}([5mTHF],[HCY],[H_2O_2])$$

$$\frac{d[THF]}{dt} = V_{FTD}([10fTHF]) + V_{MS}([5mTHF],[HCY],[H_2O_2]) + V_{PGT}([10fTHF],[GAR]) + V_{ART}([10fTHF],[AICAR]) - V_{FTS}([THF],[HCOOH]) - V_{SHMT}([SER],[THF],[GLY],[CH2F]) - V_{NE}([THF],[H_2C=O],CH2F) + V_{DHFR}([DHF],[NADPH])$$

$$\frac{d[CH2F]}{dt} = V_{SHMT}([SER],[THF],[GLY],[CH2F]) + V_{NE}([THF],[H_2C=O],CH2F) - V_{TS}([dUMP],[CH2F]) - V_{MTHFR}([CH2F],[NADPH],[SAM]) - V_{MHD}([CH2F],[CHF])$$

$$\frac{d[CHF]}{dt} = V_{MHD}([CH2F],[CHF]) - V_{MCH}([CHF],[10fTHF])$$

$$\frac{d[10fTHF]}{dt} = V_{MCH}([CH2F],[10fTHF]) + V_{FTS}([THF],[HCOOH]) - V_{PGT}([10fTHF],[GAR]) - V_{ART}([10fTHF],[AICARP]) - V_{FTD}([10fTHF])$$

$$\frac{d[MET]}{dt} = V_{BHMT}([HCY],[BET],[SAM],[SAH],[H_2O_2]) + V_{MS}([5mTHF],[HCY],H_2O_2) + V_{bMetIn}([bMET],[MET]) - V_{MATI}([MET],[SAM],[GSSG]) - V_{MATIII}([MET],[SAM],[GSSG])$$

$$\frac{d[SAM]}{dt} = V_{MATI}([MET],[SAM],[GSSG]) + V_{MATIII}([MET],[SAM],[GSSG]) - V_{GNMT}([SAM],[SAH],[5mTHF],[GLY]) - V_{DNMT}([SAM],[SAH],[DNA])$$

$$\frac{d[SAH]}{dt} = V_{GNMT}([SAM],[SAH],[5mTHF],[GLY]) + V_{DNMT}([SAM],[SAH],[DNA]) - V_{SAAH}([SAH],[HCY])$$

Author Manuscript Published OnlineFirst on June 13, 2017; DOI: 10.1158/0008-5472.CAN-17-0262
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

$$\frac{d[HCY]}{dt} = V_{SAAH}([SAH],[HCY]) - V_{CBS}([HCY],[SAM],[SAH],[SER],[H_2O_2]) - V_{BHMT}([HCY],[BET],[SAM],[SAH],[H_2O_2]) - V_{MS}([5mTHF],[HCY],H_2O_2)$$

$$\frac{d[GSH]}{dt} = V_{GS}([GLY],[GLC],[GSH]) - (2)V_{GPX}([GSH],[H_2O_2]) + (2)V_{GR}([GSSG],[NADPH]) - d_1[GSH] - V_{bGSHOut}([GSH])$$

$$\frac{d[GSSG]}{dt} = V_{GPX}([GSH],[H_2O_2]) - V_{GR}([GSSG],[NADPH]) - d2[GSSG] - V_{bGSSGOut}([GSSG])$$

$$\frac{d[GLC]}{dt} = V_{GCS}([CYS],[GLU],[GSH],[GLC],[H_2O_2]) - V_{GS}([GLY],[GLC],[GSH])$$

$$\frac{d[CYS]}{dt} = V_{CTGL}([CYT]) - V_{GCS}([CYS],[GLU],[GSH],[GLC],[H_2O_2]) + V_{bCYSIn}([bCYS]) - 0.35 * \frac{[CYS]^2}{200}$$

$$\frac{d[CYT]}{dt} = V_{CBS}([HCY],[SAM],[SAH],[SER],[H_2O_2]) - V_{CTGL}([CYT])$$

Table 2: Cancer specific DNA methylation: observed *vs.* predicted levels. Column 2 is the observed DNA methylation level changes calculated using methylation array data with *p*-values of hyper-, no change or hypo-methylation shown in the third column. Columns 4 and 5 are predicted steady-state DNA methylation levels for cancer and control tissues, where a cancer type is predicted to have hypo-methylation if the predicted level of DNA methylation is lower in cancer compared to that in controls. A prediction is considered to be consistent with experimental data if they both show hypo- or hyper-DNA methylation.

Cancer type	Observed methylation changes	*p*-value	Predicted cancer (μM/hr)	Predicted control (μM/hr)
BLCA	Hypo	1.33E-08	81.95	94.71
BRCA	No change	5.13E-01	90.27	94.71

COAD	Hypo	2.51E-02	86.99	94.71
HNSC	Hypo	8.39E-04	89.61	94.71
KIRC	Hypo	7.25E-13	82.48	94.71
LIHC	Hypo	5.10E-13	85.76	94.71
LUAD	Hypo	4.11E-02	92.59	94.71
LUSC	Hypo	7.63E-11	85.09	94.71
PRAD	Hyper	1.00E-04	99.83	94.71
THCA	No change	9.13E-02	106.49	94.71

Table 3: Comparisons of DNA methylation levels between samples with different levels of anti-oxidation and nucleotide synthesis rates for 10 cancer types. The second column shows *p*-values of Wilcoxin tests for the null hypothesis that DNA methylation levels of patient samples with high GSH synthesis rate is lower than those with relatively lower GSH synthesis activities; the third columns shows *p*-values of Wilcoxin tests for the null hypothesis that the DNA methylation levels of patient samples with high nucleotide synthesis rate is lower than those with relatively lower nucleotide synthesis rates.

Cancer type	Anti-oxidation	Nucleotide synthesis
BLCA	**2.86E-03**	**1.25E-03**
BRCA	1.62E-01	5.54E-02
COAD	**2.60E-02**	**1.21E-02**
HNSC	**2.44E-04**	**1.38E-05**
KIRC	9.58E-01	5.97E-01
LIHC	1.68E-01	**9.69E-06**
LUAD	**8.40E-04**	**6.65E-05**

LUSC	**1.40E-02**	**2.90E-05**
PRAD	5.98E-01	6.94E-01
THCA	1.00E+00	9.34E-01

Table 4: The significances of observed negative correlations between: (1) DNA methylation and nucleotide synthesis when high (Methyl_H) and low (Methyl_L) level of methyl is available; and (2) DNA methylation and anti-oxidation capacity when high (Sulfur_H) and low (Sulfur_L) level of sulfur is available. Significant negative correlations (p-value < 0.05) are marked bold.

	Methyl_H	Methyl_L	Sulfur_H	Sulfur_L
BLCA	**5.50E-05**	**5.26E-03**	8.45E-02	**2.20E-02**
BRCA	4.02E-01	**1.96E-02**	2.47E-01	2.10E-01
COAD	6.21E-01	**3.14E-04**	1.03E-01	**2.89E-03**
HNSC	**6.00E-04**	**2.25E-05**	1.62E-01	**1.10E-02**
KIRC	4.69E-01	7.26E-01	9.99E-01	5.14E-01
LIHC	**5.63E-04**	**2.84E-05**	9.32E-01	**5.00E-02**
LUAD	1.02E-01	**3.51E-05**	**1.15E-02**	**2.08E-04**
LUSC	**1.59E-06**	**1.08E-03**	1.18E-01	**1.36E-02**
PRAD	6.90E-01	4.73E-01	9.64E-02	1.82E-01
THCA	9.10E-01	9.60E-01	9.41E-01	9.97E-01

Figure legends:

Figure 1: The one-carbon metabolic pathway (adapted from (14)) consisting of the folate, the methionine and the transsulphuration pathways. All the reaction substrates are in upright letters and the catalyzing enzymes are in italics. Substrates in purple, green and red represent metabolite variables in three subsystems in our model, and all the other substrates are treated as constants. Metabolites in light blue are those that could activate or inhibit certain reactions. bMET and bCYS represent methionine and cysteine up-taken from the blood circulation, respectively.

Figure 2: Under normal cytosolic conditions, the DNA methylation rate (A), nucleotide synthesis rate (B) and methyl concentration (C) (*z*-axis) in steady states as a function of the serine (*x*-axis) and folate (*y*-axis) concentrations, respectively; and DNA methylation rate (D), GSH synthesis rate (E) and total sulfur concentration (F) (*z*-axis) as a function of the cysteine (*x*-axis) and methionine (*y*-axis) influx from blood circulation, respectively.

Figure 3: (A) The level of 5mTHF (*y*-axis), reflecting the level of methyl going to DNA methylation as a function of the *TS*'s V_{max} value (*x*-axis); (B) The level of DHF (*y*-axis), reflecting the level of methyl going to nucleotide synthesis as a function of the *TS*'s V_{max} value (*x*-axis); (C) the DNA methylation rate (*y*-axis) as a function of the *TS*'s V_{max} value; (D) the nucleotide synthesis rate (*y*-axis) as a function of the *TS*'s V_{max} value; (E) The level of methionine (*y*-axis), representing the level of sulfur going to DNA methylation as a function of the H_2O_2 concentration (*x*-axis); (F) The level of cystathionine (*y*-axis), reflecting the level of sulfur going to GSH synthesis as a function of the H_2O_2 concentration (*x*-axis); (G) The DNA methylation rate (*y*-axis) as a function of the H_2O_2 concentration; (H) The GSH synthesis rate (*y*-axis) as a function of the H_2O_2 concentration.

Figure 1

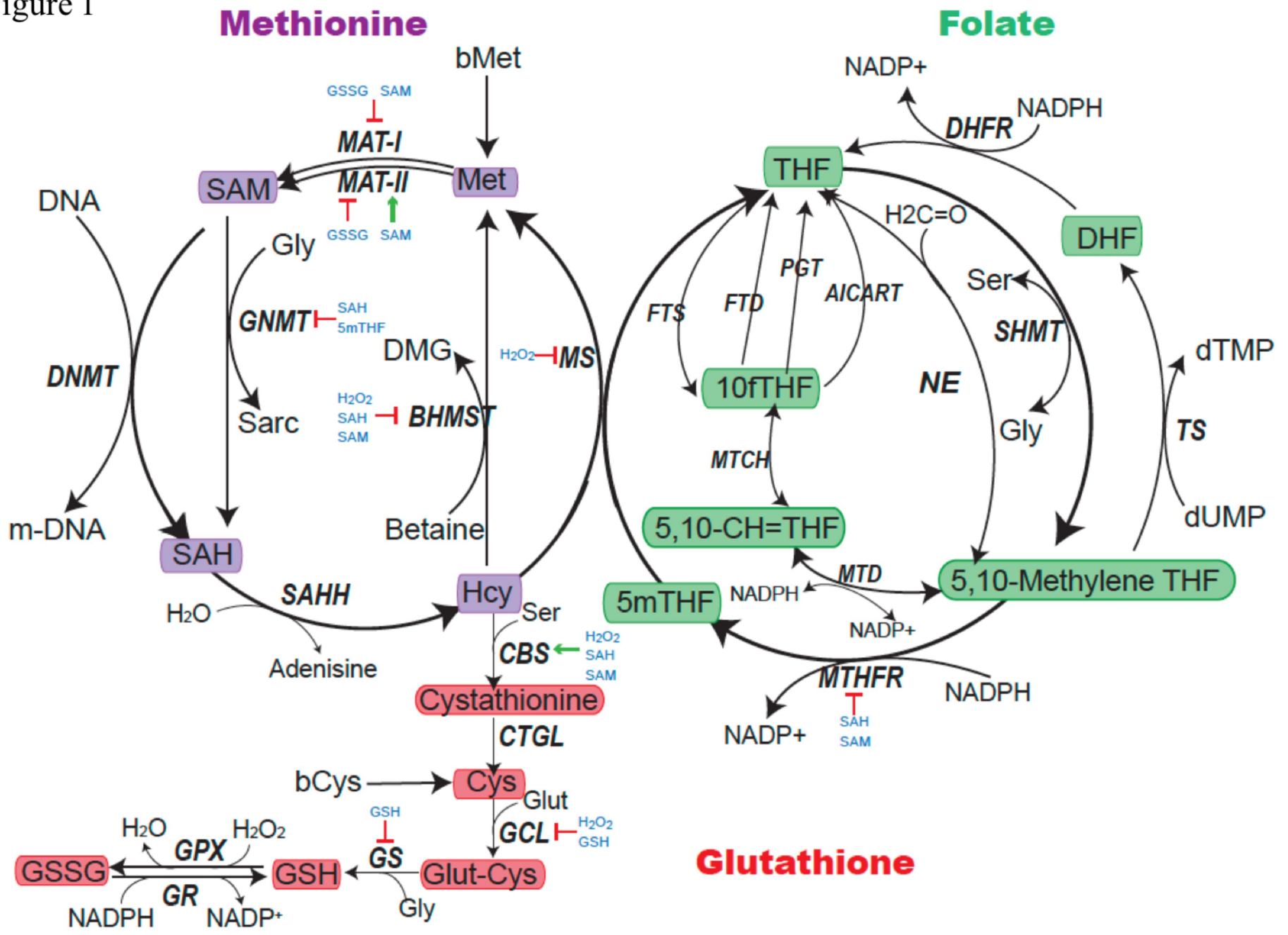

Methionine
Folate
bMet
GSSG
SAM
MAT-I
MAT-II
SAM
Met
DNA
Gly
GNMT
SAH
5mTHF
DMG
H2O2
MS
DNMT
Sarc
H2O2
SAH
SAM
BHMST
m-DNA
Betaine
SAH
SAHH
Hcy
H2O
Ser
Adenisine
CBS
H2O2
SAH
SAM
Cystathionine
CTGL
bCys
Cys
Glut
GSH
GCL
H2O2
GSH
H2O
H2O2
GPX
GSSG
GSH
GS
Glut-Cys
GR
NADPH
NADP+
Gly
NADP+
DHFR
NADPH
THF
H2C=O
DHF
PGT
AICART
FTS
FTD
Ser
SHMT
dTMP
10fTHF
NE
Gly
TS
MTCH
dUMP
5,10-CH=THF
MTD
5,10-Methylene THF
5mTHF
NADPH
NADP+
MTHFR
NADPH
NADP+
SAH
SAM
Glutathione

Figure 2

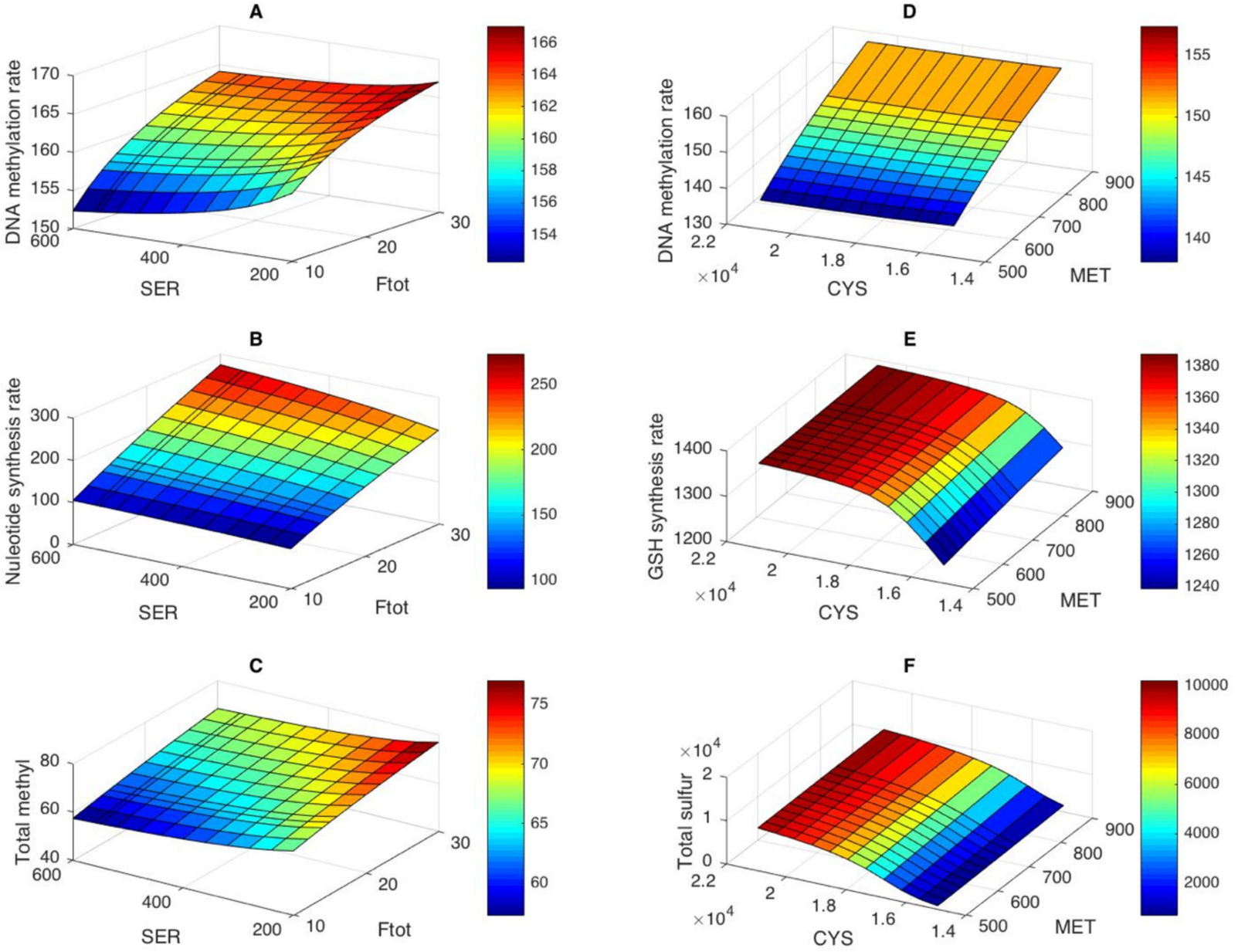
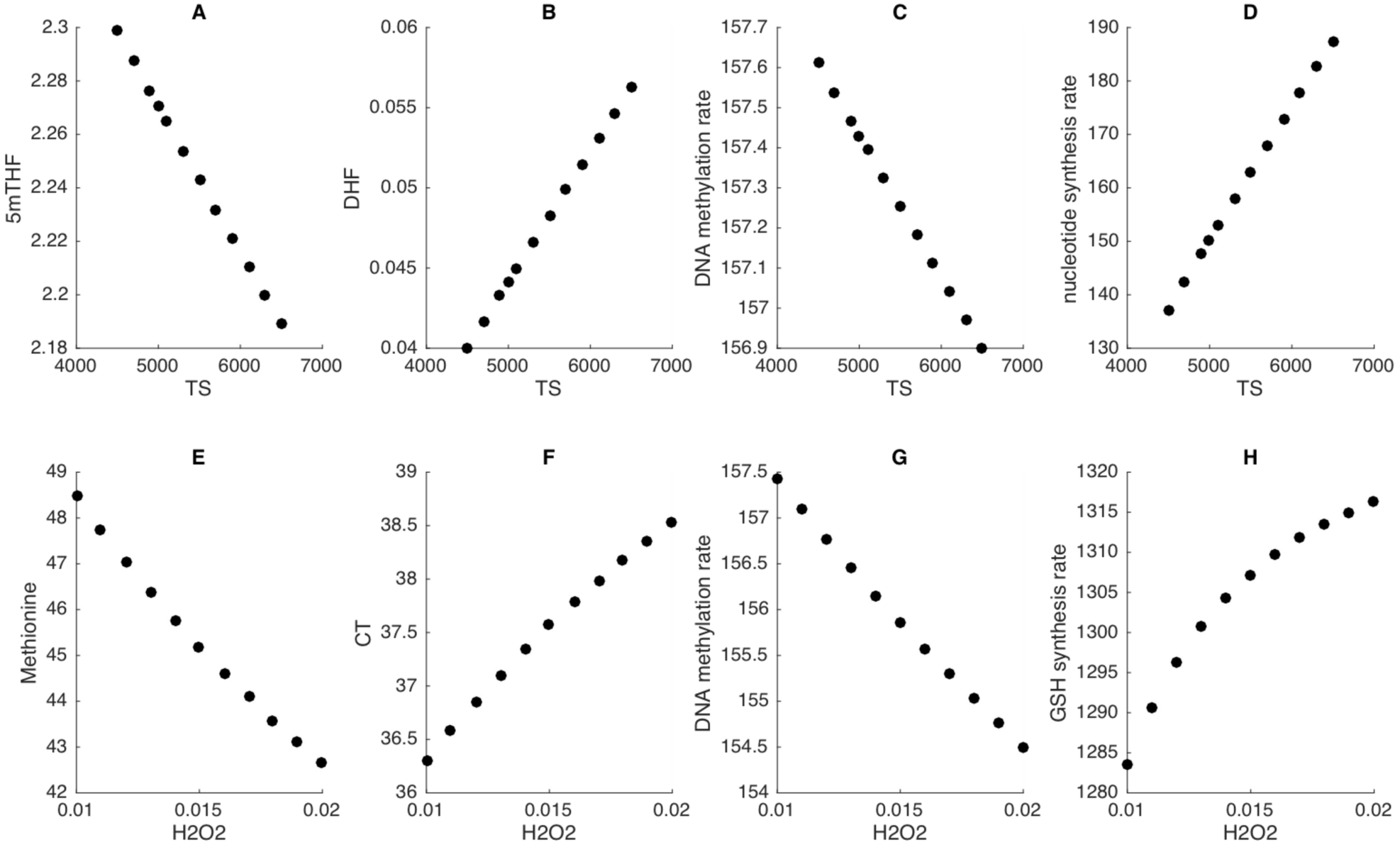

A
DNA methylation rate
170
165
160
155
150
600
400
200
SER
10
20
30
Ftot
166
164
162
160
158
156
154
D
DNA methylation rate
160
150
140
130
2.2
2
1.8
1.6
1.4
×10^4
CYS
500
600
700
800
900
MET
155
150
145
140
B
Nuleotide synthesis rate
300
200
100
0
600
400
200
SER
10
20
30
Ftot
250
200
150
100
E
GSH synthesis rate
1400
1300
1200
2.2
2
1.8
1.6
1.4
×10^4
CYS
500
600
700
800
900
MET
1380
1360
1340
1320
1300
1280
1260
1240
C
Total methyl
80
60
40
600
400
200
SER
10
20
30
Ftot
75
70
65
60
F
Total sulfur
×10^4
2
1
0
2.2
2
1.8
1.6
1.4
×10^4
CYS
500
600
700
800
900
MET
10000
8000
6000
4000
2000

Figure 3


A
5mTHF
TS
B
DHF
TS
C
DNA methylation rate
TS
D
nucleotide synthesis rate
TS
E
Methionine
H2O2
F
CT
H2O2
G
DNA methylation rate
H2O2
H
GSH synthesis rate
H2O2