

# Bioinformatic and Genetic Association Analysis of MicroRNA Target Sites in One-Carbon Metabolism Genes

Nicole Stone<sup>1‡</sup>, Faith Pangilinan<sup>1</sup>, Anne M. Molloy<sup>2</sup>, Barry Shane<sup>3</sup>, John M. Scott<sup>2</sup>, Per Magne Ueland<sup>4</sup>, James L. Mills<sup>5</sup>, Peader N. Kirke<sup>6</sup>, Praveen Sethupathy<sup>1\*</sup>, Lawrence C. Brody<sup>1\*</sup>

**1** Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **2** School of Immunology and Biochemistry, Trinity College, Dublin, Ireland, **3** Department of Nutritional Sciences and Toxicology, University of California, Berkeley, California, United States of America, **4** Section of Pharmacology, Institute of Medicine, University of Bergen and Haukeland University Hospital, Bergen, Norway, **5** Division of Epidemiology, Statistics, and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland, United States of America, **6** Child Health Epidemiology Unit, Health Research Board of Ireland, Dublin, Ireland

## Abstract

One-carbon metabolism (OCM) is linked to DNA synthesis and methylation, amino acid metabolism and cell proliferation. OCM dysfunction has been associated with increased risk for various diseases, including cancer and neural tube defects. MicroRNAs (miRNAs) are ~22 nt RNA regulators that have been implicated in a wide array of basic cellular processes, such as differentiation and metabolism. Accordingly, mis-regulation of miRNA expression and/or activity can underlie complex disease etiology. We examined the possibility of OCM regulation by miRNAs. Using computational miRNA target prediction methods and Monte-Carlo based statistical analyses, we identified two candidate miRNA “master regulators” (miR-22 and miR-125) and one candidate pair of “master co-regulators” (miR-344-5p/484 and miR-488) that may influence the expression of a significant number of genes involved in OCM. Interestingly, miR-22 and miR-125 are significantly up-regulated in cells grown under low-folate conditions. In a complementary analysis, we identified 15 single nucleotide polymorphisms (SNPs) that are located within predicted miRNA target sites in OCM genes. We genotyped these 15 SNPs in a population of healthy individuals (age 18–28, n = 2,506) that was previously phenotyped for various serum metabolites related to OCM. Prior to correction for multiple testing, we detected significant associations between TCbIR rs9426 and methylmalonic acid (p = 0.045), total homocysteine levels (tHcy) (p = 0.033), serum B12 (p < 0.0001), holo transcobalamin (p < 0.0001) and total transcobalamin (p < 0.0001); and between MTHFR rs1537514 and red blood cell folate (p < 0.0001). However, upon further genetic analysis, we determined that in each case, a linked missense SNP is the more likely causative variant. Nonetheless, our Monte-Carlo based *in silico* simulations suggest that miRNAs could play an important role in the regulation of OCM.

**Citation:** Stone N, Pangilinan F, Molloy AM, Shane B, Scott JM, et al. (2011) Bioinformatic and Genetic Association Analysis of MicroRNA Target Sites in One-Carbon Metabolism Genes. PLoS ONE 6(7): e21851. doi:10.1371/journal.pone.0021851

**Editor:** Matej Oresic, Governmental Technical Research Centre of Finland, Finland

**Received:** March 11, 2011; **Accepted:** June 7, 2011; **Published:** July 12, 2011

This is an open-access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the Creative Commons CC0 public domain dedication.

**Funding:** This study was supported by the Intramural Research Programs of the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development, and the National Human Genome Research Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: sethupathyp@mail.nih.gov (PS); lbrody@mail.nih.gov (LCB)

‡ Current address: Biomedical Sciences Graduate Program, University of California San Francisco, San Francisco, California, United States of America

## Introduction

One-carbon metabolism (OCM) comprises a set of reactions involving folate coenzymes and is critical for essential processes including DNA methylation, cell proliferation, and the synthesis of nucleic and amino acids. Insufficient folate or vitamin B12 intake, genetic variation [1], or drug interference [2] can disrupt normal OCM function. OCM dysfunction is linked to severe health complications such as cancer [3,4], anemia [5] and neural tube defects [6]. Folate-mediated OCM also influences levels of the non-protein amino acid, homocysteine. Elevated homocysteine levels have been linked to an increased risk for neural tube defects [7]. Recent studies have revealed widespread changes in gene expression under folate-deficient conditions [8]. However, the underlying molecular mechanisms of these changes are poorly understood.

MicroRNAs (miRNAs) are ~22 nucleotide non-coding RNAs that regulate eukaryotic gene expression at the post-transcriptional

level [9]. Specifically, they associate with the RNA Induced Silencing Complex (RISC) and guide it to target sites within mRNAs. Once bound to mRNA, RISC induces gene repression through a variety of mechanisms, including direct mRNA cleavage and translational inhibition [10,11]. Most of the ~950 currently known human miRNAs [12] are predicted to target hundreds of genes, and many genes are targeted by multiple miRNAs [9,13].

miRNAs have been implicated in a wide array of fundamental biological processes, such as development [14], lipid metabolism [15], response to environmental stress [16] and innate immunity [17]. Accordingly, mis-regulation of miRNA expression and/or activity has been linked to many diseases including various cancers and cardiovascular conditions [18], and is likely to underlie the molecular etiology of many other disorders.

The role of miRNAs in the modulation of folate-mediated OCM has not been extensively investigated. However, initial studies suggest that folate influences miRNA expression; human lymphoblastoid cells grown under folate-deficient conditions

exhibit significant changes in the levels of 24 miRNAs [19]. Given this observation, we examined the possibility of miRNA-mediated regulation of the OCM pathway. Specifically, we applied a computational strategy to predict whether any known human miRNAs are candidate “master regulators” of the genes most commonly associated with OCM. In a complementary analysis, we also assessed whether genetic variants within predicted miRNA target sites in OCM genes are associated with relevant metabolites.

## Methods

### miRNA target predictions

We used the TargetScanS 5.1 algorithm [20] for genome-wide miRNA target prediction. Target predictions were filtered further based on conservation of the miRNA and the 6–8 nt target site. For our primary analyses, we defined conservation as precisely shared nucleotide content between human and two of three other mammalian species (rat, mouse and dog).

### Discovery of candidate master miRNA regulators of OCM genes

We implemented a Monte-Carlo algorithm to identify miRNAs that are predicted to target 42 OCM genes significantly more than expected by chance. First, we randomly selected 42 genes from the human genome under the condition that they have at least one predicted miRNA target site. Then we calculated the number of genes from this randomly selected set that each known human miRNA is predicted to target. By repeating this simulation 10,000 times, we generated a background distribution of the number of predicted target genes for each miRNA, which we then used to calculate an empirical p-value for the number of predicted OCM target genes. We increased the number of simulations (up to 10,000,000) in specific instances in order to obtain a non-zero empirical p-value. To account for differences in the average 3' UTR length between OCM genes and the randomly selected genes in each simulation, we normalized the number of predicted target genes to the average 3' UTR length.

### The Trinity Student Study (TSS) Cohort

A population of 2,506 healthy, ethnically Irish individuals consisting of students attending University of Dublin, Trinity College and aged between 18 and 28 years old was recruited over a period of one academic year (TSS cohort). Each participant completed a questionnaire regarding intake of relevant supplements and fortified foods, and gave 30 mL of blood. Genomic DNA was extracted from all samples using a Qiagen QIAamp DNA Blood Mini Kit or a Qiagen DNeasy Kit (Qiagen, UK). Serum folate and vitamin B12 were measured by microbiological assays as previously described [21,22]. Holotranscobalamin II (holoTC) was directly measured with a commercial, monoclonal technique (Abbott AxSYM, Microparticle Enzyme Immunoassay); total transcobalamin II (total TC) was measured with the same technique after first saturating the protein with cyanocobalamin. Levels of total homocysteine (tHcy), methionine, and methylmalonic acid (MMA) were measured at Bevitall AS ([www.bevital.no](http://www.bevital.no)). Ethical approval was obtained from the Dublin Federated Hospitals Research Ethics Committee, which is affiliated with TCD and reviewed by the Office of Human Subjects Research at the National Institutes of Health. Written informed consent was obtained from the participants before commencing the study.

### SNP selection

To identify single nucleotide polymorphisms (SNPs) that might influence miRNA targeting efficiency, we cross-referenced the

locations of all SNPs with a minor allele frequency  $> 0.05$  from HapMap Phase III [23] and dbSNP v131 [24] against the locations of all predicted miRNA target sites in the 3' UTRs of 42 OCM-related mRNAs. This yielded 21 SNPs; however, due to primer incompatibility with multiplex design or assay failure of the original SNP, proxies ( $r^2 > 0.95$ ) were selected for 3 of these SNPs. In addition to these 21 SNPs in miRNA binding sites we also tested two SNPs with previously reported functional relevance: TCblR E88del [25] and MTHFR C677T [1].

### Genotyping

Genotypes for all 21 variants were determined by allele-specific extension product mass, detected via matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF) mass spectrometry after undergoing iPLEX assay chemistry (Sequenom, San Diego, CA, USA). Primer sequences and assay conditions are available upon request.

Out of the 21 variants, 19 were genotyped successfully, 17 passed quality control criteria, and 15 were confirmed as polymorphic in our sample set and were tested for metabolite associations. Individual assay quality was assessed based on call rate, re-genotyping concordance and population fit to Hardy-Weinberg equilibrium (HWE). Genotyping call rates (i.e., success rates) averaged 97% for the TSS samples, and were at least 95% for all accepted variants. More than 10% of TSS samples were repeated with  $\geq 98\%$  concordance for all accepted variants. Only one SNP (MTHFR rs1537514) was out of HWE ( $p < 0.01$ ). We repeated the assay for rs1537514 using different extension primers, and also separately genotyped a highly linked SNP ( $r^2 = 0.973$ ; MTHFR rs1537516), and found that both were out of HWE ( $p = 2.9 \times 10^{-5}$  and  $p = 1.9 \times 10^{-4}$ , respectively). While genotyping error is generally the standard explanation for a SNP not adhering to HWE, it is unlikely in this case due to the consistency across three separate assays. These three assays displayed similar measures of call rates and re-genotyping concordance; therefore, MTHFR rs1537514 was retained for further analysis.

### Statistical analysis

Analysis of variance (ANOVA) was used as the primary test to evaluate each polymorphism for association with metabolite levels. For each SNP, only metabolites that we predicted may be affected based on the known biological function of the gene harboring the SNP were tested. In cases where one genotype was too rare in our sample set to analyze by ANOVA, the Student's t-test was used to compare the two genotypes present. Although some metabolic datasets were not normally distributed, our large sample sizes allowed the use of parametric tests as a screening tool. All significant associations via parametric tests were recapitulated with non-parametric tests (data not shown).

## Results

We selected 42 genes (Table 1) known to be involved in the OCM pathway to test *in silico* for miRNA targeting. In order to minimize the rate of false positive miRNA target site predictions, we used the TargetScanS 5.1 algorithm [20] which was recently identified as one of the top performing prediction strategies [26], and then further filtered the predictions to ensure that both the miRNA and the predicted target site are conserved between humans and at least two of three additional mammalian species (rat, mouse and dog). The use of a conservation filter focuses the analysis on putative miRNA regulation of OCM genes that may be important enough to have persisted across  $\sim 80$  million years of evolution [27].

**Table 1.** Genes involved in one-carbon metabolism (OCM).

| Symbol   | Function category  | HGNC  | Reference |
|----------|--|-------|-----------|
| AHCY     | Methylation cycle – regeneration of homocysteine           | 343   | [38]      |
| ALDH1L1  | Distribution of one –carbon units                          | 3978  | [39]      |
| AMT      | Distribution of one –carbon units -glycine cleavage        | 473   | [40]      |
| ATIC     | Purine biosynthesis - DNA synthesis                        | 794   | [41]      |
| BHMT     | Homocysteine remethylation                                 | 1047  | [42]      |
| CD320    | Vitamin B12 cellular receptor                              | 16692 | [43]      |
| CUBN     | Vitamin B12 intestinal receptor                            | 2548  | [44]      |
| DHFR     | Folate homeostasis   | 2861  | [45]      |
| DMGDH    | Distribution of one –carbon units                          | 24475 | [46]      |
| DNMT1    | DNA methylation  | 2976  | [47]      |
| DNMT3A   | DNA methylation  | 2978  | [47]      |
| DNMT3B   | DNA methylation  | 2979  | [47]      |
| FOLH1    | Folate hydrolysis (intestinal)                             | 3788  | [48]      |
| FOLR1    | Cellular folate uptake                                     | 3791  | [49]      |
| FPGS     | Folate homeostasis   | 3824  | [50]      |
| FTCD     | Provision of one-carbon units – histidine catabolism       | 3974  | [51]      |
| GART     | Purine biosynthesis - DNA synthesis                        | 4163  | [52]      |
| GGH      | Folate hydrolysis (lysosomal)                              | 4248  | [53]      |
| GIF      | B12 absorption   | 4268  | [44]      |
| GNMT     | Glycine methylation – folate homeostasis                   | 4415  | [54]      |
| MAT1A    | Synthesis of SAM - methylation                             | 6903  | [33]      |
| MAT2A    | Synthesis of SAM - methylation                             | 6904  | [33]      |
| MAT2B    | Synthesis of SAM - methylation,                            | 6905  | [33]      |
| MMAB     | B12 metabolism   | 19331 | [55]      |
| MTFMT    | Mitochondrial protein synthesis –formyl-methionyl transfer | 29666 | [56]      |
| MTHFD1   | Distribution of one –carbon units                          | 7432  | [34]      |
| MTHFD1L  | Distribution of one –carbon units                          | 21055 | [35]      |
| MTHFD2   | Distribution of one –carbon units                          | 7434  | [57]      |
| MTHFD2L  | Distribution of one –carbon units                          | 31865 | [57]      |
| MTHFR    | Distribution of one –carbon units                          | 7436  | [1]       |
| MTHFS    | Distribution of one –carbon units                          | 7437  | [58]      |
| MTR      | Homocysteine remethylation                                 | 7468  | [59]      |
| MTRR     | Homocysteine remethylation                                 | 7473  | [60]      |
| SARDH    | Distribution of one –carbon units                          | 10536 | [46]      |
| SHMT1    | Provision of one –carbon units -cytosol                    | 10850 | [61]      |
| SHMT2    | Provision of one –carbon units -mitochondrion              | 10852 | [61]      |
| SLC19A1  | Reduced folate carrier                                     | 10937 | [49]      |
| SLC25A32 | Mitochondrial folate transporter                           | 29683 | [62]      |
| SLC46A1  | Proton coupled folate transporter                          | 30521 | [49]      |
| TCN1     | B12 transport protein                                      | 11652 | [28]      |
| TCN2     | B12 transport protein                                      | 11653 | [28]      |
| TYMS     | Thymidylate biosynthesis – DNA synthesis and repair        | 12441 | [63]      |

For each gene, the gene symbol, known function, HGNC identification, and literature citations indicating involvement in OCM are provided.  
doi:10.1371/journal.pone.0021851.t001

### miR-22 is the top candidate master miRNA regulator of OCM genes

Using these target predictions, we implemented a Monte-Carlo algorithm to identify miRNAs that are predicted to target OCM genes significantly more than expected by chance (Methods). The analysis revealed that miR-22 is the top candidate master regulator

( $p = 0.0126$ , Figure 1 and Table 2). Though not significant after stringent Bonferroni correction for multiple testing, miR-22 is predicted to have conserved targets in 4 important OCM genes: SLC19A1, MAT2A, MTHFD2 and MTHFR. Notably, the last three genes have at least one miR-22 target site that is conserved among >5 vertebrates, including a predicted target site within

MTHFR that is ranked in the 96<sup>th</sup> percentile according to the TargetScanS scoring criteria. MTHFR also harbors four additional non-conserved miR-22 target sites, indicating a higher likelihood of miR-22 targeting functionality.

We repeated the analysis with a less stringent conservation requirement (among primates; human, rhesus, chimp) in order to investigate whether any potential master regulators have arisen recently in order to adapt to evolving physiology. This approach predicts miR-22 target sites within two additional OCM genes, TCblR and TCN2, which are involved in vitamin B12 uptake [28]. However, because the liberal criteria leads to the prediction of many more miR-22 target sites genome-wide, a large proportion of which are likely false positives, the significance of miR-22 targeting of OCM genes is diluted ( $p = 0.1235$ ). Nonetheless, this less stringent analysis does indeed identify a different candidate master regulator of OCM, miR-125/351, that may have functional relevance in relatively recent evolutionary history ( $p = 0.0033$ ; Table S1).

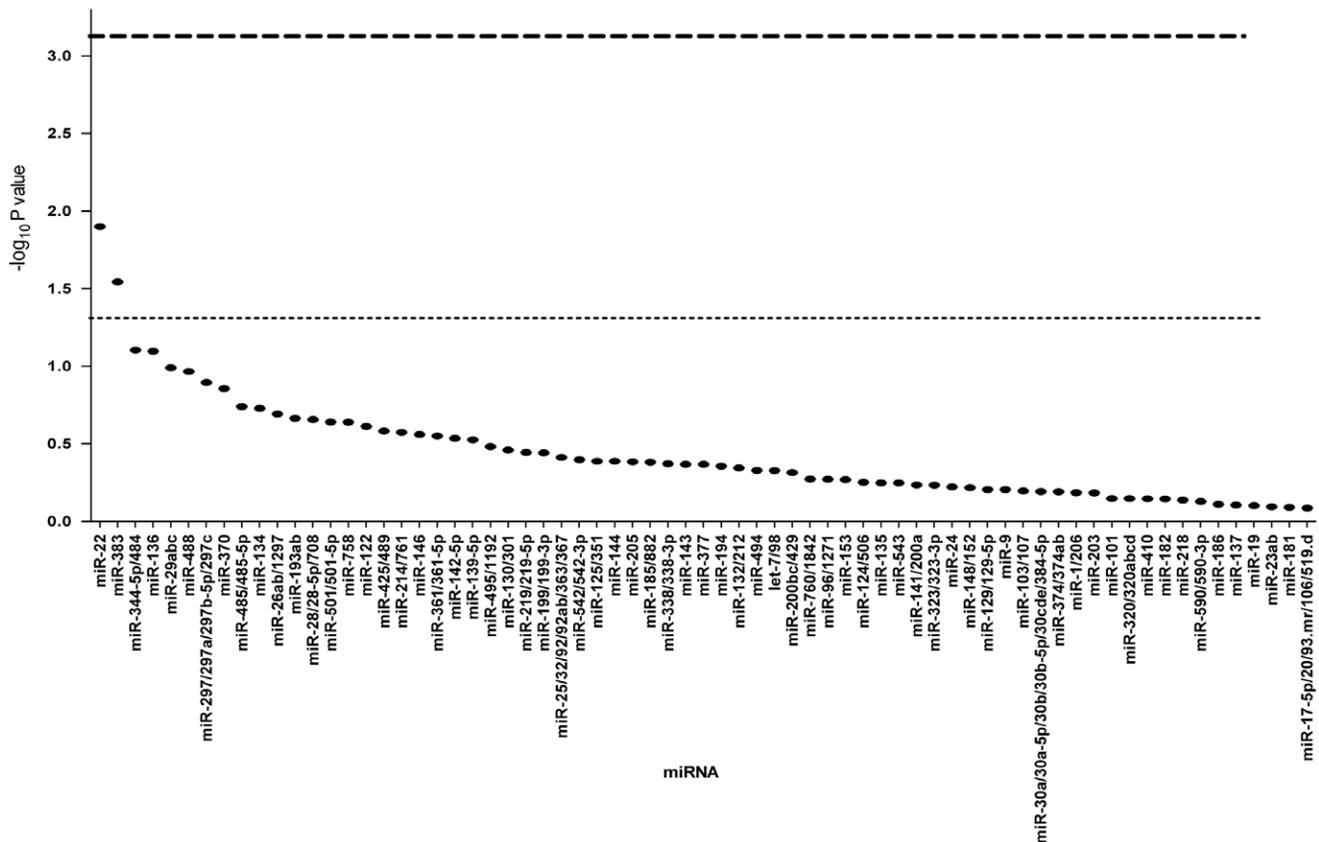
### miR-344-5p/484 and miR-488 are the top candidate master co-regulators of OCM genes

miRNAs often co-regulate their target genes in order to induce effective repression [13]. Accordingly, we implemented a Monte-Carlo algorithm to identify *pairs* of miRNAs that are predicted to co-target OCM genes significantly more than expected by chance. miR-22 does not appear in the top 30 pairs of master co-regulator candidates, suggesting that its influence upon OCM genes may

have evolved independently from that of other miRNAs. However, we did identify miR-344-5p/484 and miR-488 as a candidate pair for master co-regulators ( $p = 0.0004$ ). Interestingly, though both of these miRNAs are among the top 6 candidates for individual candidate master regulators, neither is statistically significant by itself (Table 2, Figure 1). That their targeting of OCM genes is significant only when considered together as a pair suggests that their influence upon OCM may have co-evolved.

### Genotype-phenotype correlation analysis for polymorphisms within predicted miRNA target sites in OCM genes

In order to investigate the potential phenotypic relevance of the predicted miRNA targeting of OCM genes, we first identified all common HapMap Phase III and dbSNP v131 single nucleotide polymorphisms (SNPs with minor allele frequency > 5%,  $n = 21$ ) that map to within the predicted miRNA target sites. DNA samples obtained from healthy Irish college students ( $n = 2,506$ ) were genotyped at each SNP locus. Of these 21 SNPs, 14 were successfully genotyped, and proxy SNPs ( $r^2 > 0.95$ ) were genotyped for 3 others. 15 of these 17 SNPs were polymorphic in our sample set (Table S2). Blood, plasma and serum samples from these same students were previously phenotyped for serum metabolites related to OCM. We found significant associations between 2 SNPs and several related metabolites: TCblR rs9426 with serum B12 (sB12) ( $p < 0.0001$ ), holo transcobalamin (holoTC) ( $p < 0.0001$ ), total transcobalamin (total TC) ( $p <$



**Figure 1. Statistical analysis of predicted microRNA targeting in the one-carbon metabolism (OCM) pathway.** For each microRNA, the empirical p-value for the level of enrichment of predicted target sites in OCM genes (Methods) is shown. Analysis is restricted to predicted target sites that, together with their cognate microRNAs, are conserved between humans and at least two other mammalian species among mouse, rat and dog. Lower dashed line indicates  $p = 0.05$ ; upper dashed line indicates equivalent significance level after correction for multiple hypothesis testing. doi:10.1371/journal.pone.0021851.g001

**Table 2.** Top 6 candidate microRNA master regulators of one-carbon metabolism genes in mammals.

| microRNA       | P-value for enrichment of predicted target sites in OCM genes | Predicted OCM target genes    |
|----------------|---|-------------------------------|
| miR-22         | 0.0126  | MTHFD2, MTHFR, SLC19A1, MAT2A |
| miR-383        | 0.0286  | SHMT2, DNMT3A                 |
| miR-344-5p/484 | 0.0787  | DNMT3A, MAT2A                 |
| miR-136        | 0.08  | MAT2A, MTHFS                  |
| miR-29abc      | 0.1024  | DNMT3A, MAT2A, DNMT3B, MAT1A  |
| miR-488        | 0.1081  | DNMT3A, MAT2A                 |

For each microRNA, the empirical p-value for the level of enrichment of predicted target sites in OCM genes (Methods) and the gene symbol for each predicted target gene are provided. Analysis is restricted to predicted target sites that, together with their cognate microRNAs, are conserved between humans and at least two other mammalian species among mouse, rat and dog. doi:10.1371/journal.pone.0021851.t002

0.0001), total homocysteine (tHcy) ( $p = 0.033$ ) and methylmalonic acid (MMA) ( $p = 0.045$ ); and MTHFR rs1537514 with red blood cell folate (RCF) ( $p < 0.0001$ ) (Table 3; Table S2). Upon re-analysis with only those participants who do not take vitamin supplements (which can affect metabolite levels and may make a genetic effect more difficult to detect), all associations remained significant. After Bonferroni correction for multiple testing, the TCbIR rs9426 association with sB12, holoTC and total TC and the MTHFR rs1537514 association with RCF remained significant (Table 3).

TCbIR rs9426 is in high  $D'$  with another variant within the TCbIR gene, E88del ( $D' = 0.94$ ;  $r^2 = 0.37$ ), which was previously associated with increased levels of MMA in newborns of European ancestry [25] and increased risk of neural tube defects (NTD) in a large Irish population [29]. In order to compare the effects of the two variants, we additionally genotyped E88del in our cohort. Similar to rs9426, E88del displayed significant associations with several metabolites: sB12 ( $p < 0.0001$ ), holoTC

( $p < 0.0001$ ), total TC ( $p < 0.0001$ ), tHcy ( $p < 0.011$ ), and MMA ( $p = 0.0016$ ). However, the differences between alleles in the average level of each of the 5 metabolites are greater for the E88del locus than for rs9426. This observation supports the hypothesis that the deletion of a glutamic acid residue at position 88 (E88del), rather than the removal of possible miRNA regulation (rs9426), is the likely source of the metabolic variation observed between genotypes.

MTHFR rs1537514 is in high  $D'$  ( $D' = 0.988$ ;  $r^2 = 0.059$ ) with the well-studied MTHFR 677C->T (rs1801133; MAF = 0.24) [1]. Both of these SNPs have minor allele frequencies in our cohort of at least 0.11. We compared phenotypic variation of homozygotes at each SNP locus independently. Specifically, as all MTHFR C667T (SNP1) T (i.e. minor allele) alleles occur on a rs1537514 (SNP2) G (i.e. major allele) background, we compared the following three genotype combinations by a two-tailed t-test: (#1) SNP1 TT, SNP2 GG ( $n = 267$ ); (#2) SNP1 CC, SNP2 GG ( $n = 728$ ); and (#3) SNP1 CC, SNP2 CC ( $n = 49$ ). The results of this analysis reveal that the major (#2) and minor allele (#1) homozygotes of 677C->T differ substantially more from one another in RCF (204 nM difference between means,  $p < 0.0001$ ) than the major (#2) and minor (#3) allele homozygotes of rs1537514 (152 nM difference between means,  $p = 0.0352$ ). This suggests that 677C->T is more likely to be the causative variant.

## Discussion

In this study we describe a computational strategy for the identification of candidate master miRNA regulators of a group of 42 OCM-related genes. Based on our target prediction analyses with the most stringent conservation criteria, we discover a novel role for miR-22 as a candidate master regulator of OCM. miR-22 is widely expressed and has been previously linked to cancer [30,31]. It is predicted to target OCM genes (MTHFR, TCbIR,TCN2, SLC19A1, MAT2A and MTHFD2) that are critical for folate and vitamin B12 transport, folate cofactor distribution and methylation (Table 1). The enzyme encoded by the *MTHFR* gene plays a key role in the regulation of OCM by channeling one-carbon units away from DNA synthesis and into the production of methionine by MTR [32]. Methionine is then converted to the methyl donor, S-Adenosyl methionine (SAM), by

**Table 3.** Summary metabolite data for individuals with different genotypes at significantly associated single nucleotide polymorphic (SNP) loci that are within predicted microRNA target sites.

| SNP       | Gene  | Predicted microRNA binding site | Genotype | N         | tHcy (uM) | Met A (uM) | SB12 (pM) | Sfol (nM) | RCF (nM)   | holoTC (pM) | total TC (pM) | MMA A (uM)    |
|-----------|-------|---------------------------------|----------|-----------|-----------|------------|-----------|-----------|------------|-------------|---------------|---------------|
| rs9426    | TCbIR | miR-136                         | CC       | 2192–2246 | *8.7±3.0  | N/A        | **328±140 | N/A       | N/A        | **57±27     | **832±163     | *0.188±0.0853 |
|           |       |                                 | CT       | 198–202   | *8.2±2.5  | N/A        | **382±190 | N/A       | N/A        | **88±52     | **975±211     | *0.176±0.0721 |
|           |       |                                 | TT       | 1         | 9.2±0     | N/A        | 538±0     | N/A       | N/A        | 173±0       | 1307±0        | 0.140±0       |
| rs1537514 | MTHFR | miR-596 and miR-518a-5p/527     | GG       | 1926–1946 | 8.7±3.1   | 29.6±8.5   | N/A       | 34±18     | **1057±422 | N/A         | N/A           | N/A           |
|           |       |                                 | GC       | 410–414   | 8.4±2.5   | 28.6±7.6   | N/A       | 35±29     | **1142±470 | N/A         | N/A           | N/A           |
|           |       |                                 | CC       | 51        | 8.1±2.3   | 28.2±7.2   | N/A       | 37±23     | **1269±499 | N/A         | N/A           | N/A           |

Numerical data represent mean metabolite levels ±SD; N indicates the range in the number of individuals with a particular genotype for which data was available for different metabolites. \* indicates a significant ( $p < 0.05$ ) result; \*\* indicates a significant result ( $p < 0.05$ ) after correction for multiple testing; N/A indicates that the SNP was not tested for association with a particular metabolite. doi:10.1371/journal.pone.0021851.t003

MAT2A [33]. Vitamin B12 transport via TCblR and TCN2 influences SAM production by regulating the activity of the B12-dependent MTR enzyme. Folate is required for this process and is transported from the circulation into cells by the reduced folate carrier, the product of the *SLC19A1* gene. Finally, MTHFD2 is important for the exchange of one-carbon units between the cytoplasm and the mitochondrion [34–35]. Coordinated regulation of these genes by miR-22 is likely to influence OCM and downstream epigenetic processes.

Our results revealed another potential master regulator of OCM, miR-125/351. The large majority of its predicted targets are conserved only in primates, indicating that its putative role in regulating OCM may be more evolutionarily recent. It also appears likely that while miR-125/351 functions cooperatively with other miRNAs to impose regulation on OCM genes, miR-22's influence in OCM may have evolved independently from other miRNAs. Notably, both miR-22 and miR-125/351 significantly increase in expression upon folate deficiency [19], lending further support to the prediction that these two miRNAs are relevant to folate-mediated OCM. Our results suggest that miR-344-5p/484 and miR-488 may act in a cooperative fashion to regulate OCM. Neither miR-344-5p/484 nor miR-488 was included in the Marsit et al. study that examined miRNA expression under folate-deficient conditions. Future investigations, including loss-of-function experiments using antagomirs [36] or “sponge” constructs [37] are required to validate the predicted role of miR-22, miR-125/351, miR-344-5p/484 and miR-488 as important regulators of OCM genes.

In a complementary, independent experiment, we genotyped 17 SNPs located within predicted miRNA target sites in OCM genes and found significant associations between two of the SNPs and several metabolites. However, further analyses suggested that these associations could be accounted for by nearby functional variants that are in strong linkage disequilibrium with the miRNA target site SNPs. Nonetheless, we believe that our approach, which combines bioinformatic and genetic experiments, provides a useful

model for exploring the role of miRNAs in basic physiological processes.

## Supporting Information

**Table S1 Top 5 candidate microRNA master regulators of one-carbon metabolism genes in primates.** For each microRNA, the empirical p-value for the level of enrichment of predicted target sites in OCM genes (Methods) and the gene symbol for each predicted target gene are provided. Analysis is restricted to predicted target sites that, together with their cognate microRNAs, are conserved among humans, rhesus monkey, and chimpanzee.

(DOC)

**Table S2 Complete metabolite association results for single nucleotide polymorphisms (SNPs) within predicted human microRNA target sites in OCM genes.**

\* indicates a significant ( $p < 0.05$ ) SNP-metabolite association; \*\* indicates a significant SNP-metabolite association ( $p < 0.05$ ) after correction for multiple testing; N/A indicates the SNP was not present within a predicted microRNA target site.

(DOC)

## Acknowledgments

The authors acknowledge the contribution of the study participants. Tracey Claxton assisted with sample preparation. Additional assistance and analyses were provided by Sinead Gibney and Eileen Gibney.

## Author Contributions

Conceived and designed the experiments: NS PS LCB. Performed the experiments: NS FP. Analyzed the data: NS PS FP LCB AMM. Contributed reagents/materials/analysis tools: LCB PMU. Wrote the paper: NS PS LCB. Provided cohort and related study design: AMM PNK JLM JMS BS. Edited manuscript: AMM PNK JLM JMS BS PMU FP.

## References

- Frosst P, Blom HJ, Milos R, Goyette P, Sheppard CA, et al. (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet* 10: 111–113.
- Kremer JM, Galivan J, Streckfuss A, Kamen B (1986) Methotrexate metabolism analysis in blood and liver of rheumatoid arthritis patients. Association with hepatic folate deficiency and formation of polyglutamates. *Arthritis Rheum* 29: 832–835.
- Blount BC, Mack MM, Wehr CM, MacGregor JT, Hiatt RA, et al. (1997) Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci U S A* 94: 3290–3295.
- Lee JE, Willett WC, Fuchs CS, Smith-Warner SA, Wu K, et al. (2011) Folate intake and risk of colorectal cancer and adenoma: modification by time. *Am J Clin Nutr*.
- Zittoun J (1993) [Anemias due to disorder of folate, vitamin B12 and transcobalamin metabolism]. *Rev Prat* 43: 1358–1363.
- Molloy AM, Brody LC, Mills JL, Scott JM, Kirke PN (2009) The search for genetic polymorphisms in the homocysteine/folate pathway that contribute to the etiology of human neural tube defects. *Birth Defects Res A Clin Mol Teratol* 85: 285–294.
- Mills JL, McPartin JM, Kirke PN, Lee YJ, Conley MR, et al. (1995) Homocysteine metabolism in pregnancies complicated by neural-tube defects. *Lancet* 345: 149–151.
- Crott JW, Liu Z, Keyes MK, Choi SW, Jang H, et al. (2008) Moderate folate depletion modulates the expression of selected genes involved in cell cycle, intracellular signaling and folate uptake in human colonic epithelial cell lines. *J Nutr Biochem* 19: 328–335.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281–297.
- Chekulaeva M, Filipowicz W (2009) Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr Opin Cell Biol* 21: 452–460.
- Gu S, Kay MA (2010) How do miRNAs mediate translational repression? *Silence* 1: 11.
- Griffiths-Jones S (2010) miRBase: microRNA sequences and annotation. *Curr Protoc Bioinformatics* Chapter 12: Unit 12 19 11–10.
- Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215–233.
- Rosa A, Brivanlou AH (2009) microRNAs in early vertebrate development. *Cell Cycle* 8.
- Esau C, Davis S, Murray SF, Yu XX, Pandey SK, et al. (2006) miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting. *Cell Metab* 3: 87–98.
- Li G, Luma C, Qiu J, Epstein DL, Gonzalez P (2009) Alterations in microRNA expression in stress-induced cellular senescence. *Mech Ageing Dev* 130: 731–741.
- Taganov KD, Boldin MP, Chang KJ, Baltimore D (2006) NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc Natl Acad Sci U S A* 103: 12481–12486.
- Couzin J (2008) MicroRNAs make big impression in disease after disease. *Science* 319: 1782–1784.
- Marsit CJ, Eddy K, Kelsey KT (2006) MicroRNA responses to cellular stress. *Cancer Res* 66: 10843–10848.
- Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19: 92–105.
- Kelleher BP, Broin SD (1991) Microbiological assay for vitamin B12 performed in 96-well microtitre plates. *J Clin Pathol* 44: 592–595.
- Molloy AM, Scott JM (1997) Microbiological assay for serum, plasma, and red cell folate using cryopreserved, microtiter plate method. *Methods Enzymol* 281: 43–53.
- Consortium IH (2003) The International HapMap Project. *Nature* 426: 789–796.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, et al. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 29: 308–311.
- Quadros EV, Lai SC, Nakayama Y, Sequeira JM, Hannibal L, et al. (2010) Positive newborn screen for methylmalonic aciduria identifies the first mutation

- in TCbIR/CD320, the gene for cellular uptake of transcobalamin-bound vitamin B(12). *Hum Mutat* 31: 924–929.
26. Sethupathy P, Megraw M, Hatzigeorgiou AG (2006) A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat Methods* 3: 881–886.
  27. Foote M, Hunter JP, Janis CM, Sepkoski JJ, Jr. (1999) Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. *Science* 283: 1310–1314.
  28. Moestrup SK, Birn H, Fischer PB, Petersen CM, Verroust PJ, et al. (1996) Megalin-mediated endocytosis of transcobalamin-vitamin-B12 complexes suggests a role of the receptor in vitamin-B12 homeostasis. *Proc Natl Acad Sci U S A* 93: 8612–8617.
  29. Pangilinan F, Mitchell A, Vandermeer J, Molloy AM, Troendle J, et al. (2010) Transcobalamin II receptor polymorphisms are associated with increased risk for neural tube defects. *J Med Genet*.
  30. Xiong J, Du Q, Liang Z (2010) Tumor-suppressive microRNA-22 inhibits the transcription of E-box-containing c-Myc target genes by silencing c-Myc binding protein. *Oncogene*.
  31. Neeley LA, Patel S, Garver J, Gallo M, Hackett M, et al. (2006) A single-molecule method for the quantitation of microRNA gene expression. *Nat Methods* 3: 41–46.
  32. Yamada K, Strahler JR, Andrews PC, Matthews RG (2005) Regulation of human methylenetetrahydrofolate reductase by phosphorylation. *Proc Natl Acad Sci U S A* 102: 10454–10459.
  33. Ramani K, Yang H, Kuhlenkamp J, Tomasi L, Tsukamoto H, et al. (2010) Changes in the expression of methionine adenosyltransferase genes and S-adenosylmethionine homeostasis during hepatic stellate cell activation. *Hepatology* 51: 986–995.
  34. MacFarlane AJ, Perry CA, Ginary HH, Gao D, Allen RH, et al. (2009) Mthfd1 is an essential gene in mice and alters biomarkers of impaired one-carbon metabolism. *J Biol Chem* 284: 15333–15339.
  35. Pike ST, Rajendra R, Artzt K, Appling DR (2010) Mitochondrial C1-tetrahydrofolate synthase (MTHFD1L) supports the flow of mitochondrial one-carbon units into the methyl cycle in embryos. *J Biol Chem* 285: 4612–4620.
  36. Czech MP (2006) MicroRNAs as therapeutic targets. *N Engl J Med* 354: 1194–1195.
  37. Ebert MS, Sharp PA (2010) MicroRNA sponges: progress and possibilities. *RNA* 16: 2043–2050.
  38. Mudd SH (2011) Hypermethioninemias of genetic and non-genetic origin: A review. *Am J Med Genet C Semin Med Genet* 157: 3–32.
  39. Krupenko NI, Dubard ME, Strickland KC, Moxley KM, Oleinik NV, et al. (2010) ALDH1L2 is the mitochondrial homolog of 10-formyltetrahydrofolate dehydrogenase. *J Biol Chem* 285: 23056–23063.
  40. Nanao K, Takada G, Takahashi E, Seki N, Komatsu Y, et al. (1994) Structure and chromosomal localization of the aminomethyltransferase gene (AMT). *Genomics* 19: 27–30.
  41. Bulock KG, Beardsley GP, Anderson KS (2002) The kinetic mechanism of the human bifunctional enzyme ATIC (5-amino-4-imidazolecarboxamide ribonucleotide transformylase/inosine 5'-monophosphate cyclohydrolase). A surprising lack of substrate channeling. *J Biol Chem* 277: 22168–22174.
  42. Garrow TA (1996) Purification, kinetic properties, and cDNA cloning of mammalian betaine-homocysteine methyltransferase. *J Biol Chem* 271: 22831–22838.
  43. Quadros EV, Nakayama Y, Sequeira JM (2009) The protein and the gene encoding the receptor for the cellular uptake of transcobalamin-bound cobalamin. *Blood* 113: 186–192.
  44. Kozyraki R, Kristiansen M, Silahatoglu A, Hansen C, Jacobsen C, et al. (1998) The human intrinsic factor-vitamin B12 receptor, cubilin: molecular characterization and chromosomal mapping of the gene to 10p within the autosomal recessive megaloblastic anemia (MGA1) region. *Blood* 91: 3593–3600.
  45. Schweitzer BI, Dicker AP, Bertino JR (1990) Dihydrofolate reductase as a therapeutic target. *FASEB J* 4: 2441–2452.
  46. Porter DH, Cook RJ, Wagner C (1985) Enzymatic properties of dimethylglycine dehydrogenase and sarcosine dehydrogenase from rat liver. *Arch Biochem Biophys* 243: 396–407.
  47. Kulis M, Esteller M (2010) DNA methylation and cancer. *Adv Genet* 70: 27–56.
  48. Halsted CH, Ling E, Luthi-Carter R, Villanueva JA, Gardner JM, et al. (2000) Folylpoly-gamma -glutamate carboxypeptidase from pig jejunum. Molecular characterization and relation to glutamate carboxypeptidase II. *J Biol Chem* 275: 30746.
  49. Zhao R, Diop-Bove N, Visentin M, Goldman ID (2010) Mechanisms of Membrane Transport of Folates into Cells and Across Epithelia. *Annu Rev Nutr*.
  50. Garrow TA, Shane B (1993) Purification and general properties of human folylpolyglutamate synthetase. *Adv Exp Med Biol* 338: 659–662.
  51. Kohls D, Croteau N, Mejia N, MacKenzie RE, Vrieling A (1999) Crystalization and preliminary X-ray analysis of the formiminotransferase domain from the bifunctional enzyme formiminotransferase-cyclodeaminase. *Acta Crystallogr D Biol Crystallogr* 55: 1206–1208.
  52. Welin M, Grossmann JG, Flodin S, Nyman T, Stenmark P, et al. (2010) Structural studies of tri-functional human GART. *Nucleic Acids Res* 38: 7308–7319.
  53. Galivan J, Ryan TJ, Chave K, Rhee M, Yao R, et al. (2000) Glutamyl hydrolase. pharmacological role and enzymatic characterization. *Pharmacol Ther* 85: 207–215.
  54. Luka Z, Cerone R, Phillips JA, 3rd, Mudd HS, Wagner C (2002) Mutations in human glycine N-methyltransferase give insights into its role in methionine metabolism. *Hum Genet* 110: 68–74.
  55. Shevell MI, Matiaszuk N, Ledley FD, Rosenblatt DS (1993) Varying neurological phenotypes among muto and mut- patients with methylmalonyl-CoA mutase deficiency. *Am J Med Genet* 45: 619–624.
  56. Takeuchi N, Kawakami M, Omori A, Ueda T, Spremulli LL, et al. (1998) Mammalian mitochondrial methionyl-tRNA transformylase from bovine liver. Purification, characterization, and gene structure. *J Biol Chem* 273: 15085–15090.
  57. Bolusani S, Young BA, Cole NA, Tibbetts AS, Momb J, et al. (2011) Mammalian MTHFD2L encodes a mitochondrial methylenetetrahydrofolate dehydrogenase isozyme expressed in adult tissues. *J Biol Chem* 286: 5166–5174.
  58. Field MS, Szebenyi DM, Stover PJ (2006) Regulation of de novo purine biosynthesis by methylenetetrahydrofolate synthetase in neuroblastoma. *J Biol Chem* 281: 4215–4221.
  59. Brody LC, Baker PJ, Chines PS, Musick A, Molloy AM, et al. (1999) Methionine synthase: high-resolution mapping of the human gene and evaluation as a candidate locus for neural tube defects. *Mol Genet Metab* 67: 324–333.
  60. Yamada K, Gravel RA, Toraya T, Matthews RG (2006) Human methionine synthase reductase is a molecular chaperone for human methionine synthase. *Proc Natl Acad Sci U S A* 103: 9476–9481.
  61. Garrow TA, Brenner AA, Whitehead VM, Chen XN, Duncan RG, et al. (1993) Cloning of human cDNAs encoding mitochondrial and cytosolic serine hydroxymethyltransferases and chromosomal localization. *J Biol Chem* 268: 11910–11916.
  62. McCarthy EA, Titus SA, Taylor SM, Jackson-Cook C, Moran RG (2004) A mutation inactivating the mitochondrial inner membrane folate transporter creates a glycine requirement for survival of chinese hamster cells. *J Biol Chem* 279: 33829–33836.
  63. Woeller CF, Anderson DD, Szebenyi DM, Stover PJ (2007) Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *J Biol Chem* 282: 17623–17631.