Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer

Roland Nilsson1,2,*, Mohit Jain3,4,5,6,*, Nikhil Madhusudhan3,4,5, Nina Gustafsson Sheppard1,2, Laura Strittmatter3,4,5, Caroline Kampf7, Jenny Huang8, Anna Asplund7 & Vamsi K. Mootha3,4,5

Metabolic remodeling is now widely regarded as a hallmark of cancer, but it is not clear whether individual metabolic strategies are frequently exploited by many tumours. Here we compare messenger RNA profiles of 1,454 metabolic enzymes across 1,981 tumours spanning 19 cancer types to identify enzymes that are consistently differentially expressed. Our meta-analysis recovers established targets of some of the most widely used chemotherapeutics, including dihydrofolate reductase, thymidylate synthase and ribonucleotide reductase, while also spotlighting new enzymes, such as the mitochondrial proline biosynthetic enzyme PYCR1. The highest scoring pathway is mitochondrial one-carbon metabolism and is centred on MTHFD2. MTHFD2 RNA and protein are markedly elevated in many cancers and correlated with poor survival in breast cancer. MTHFD2 is expressed in the developing embryo, but is absent in most healthy adult tissues, even those that are proliferating. Our study highlights the importance of mitochondrial compartmentalization of one-carbon metabolism in cancer and raises important therapeutic hypotheses.
Divergent metabolism in tumours was first recognized nearly a century ago, is consistently observed across a number of tumour types, and has been exploited for diagnostic as well as therapeutic purposes. For example, rapid glucose consumption in tumours may be imaged by positron emission tomography and used to diagnose malignancy and to monitor the response to therapy. Furthermore, the dependence of cancer cells on nucleotide metabolism forms the basis for the use of several common chemotherapeutics, including agents targeting dihydrofolate reductase (DHF), thymidylate synthase (TYMS) and ribonucleotide reductase (RRM2). Moreover, recent data have suggested that many of the growth factor signalling pathways commonly perturbed in cancer impinge on metabolic enzymes, as well as that metabolic enzymes may act as bona fide oncogenes and even transform cells. Collectively, these observations underscore the need for a deeper understanding of metabolic reprogramming in cancer. While classic biochemical studies have identified a number of enzymes whose activities are increased in cancers, the complex, coordinated changes in metabolism that occur during cancer transformation have only begun to be understood.

Over the past decade, a wealth of data on tumours, normal tissues and cell models have been generated using microarrays and analysed to identify genes differentially expressed in cancer. These data provide a unique opportunity to study expression patterns of metabolic enzymes in cancer and thereby define the metabolic programme of cancer on a genome-wide scale. Yet, as most studies have compared tumour tissue with a quiescent, postmitotic normal control tissue, these analyses do not indicate whether enzymes overexpressed in tumours are also active in proliferative normal tissues. Identification of cancer-specific metabolic activities is essential, as current chemotherapeutic agents target metabolic enzymes found both in transformed cells as well as normal proliferating cells, notably immune cells, hair follicles and intestinal epithelium, resulting in the on-target side effects in proliferative tissues that limit these agents’ therapeutic index.

Here we report a systematic re-analysis of previously published microarray data sets, focusing on genes known or predicted to encode metabolic enzymes, and identify several enzymes and pathways consistently overexpressed or underexpressed across a large number of different cancer types. In particular, we find that enzymes of the mitochondrial folate metabolic pathway, which are ordinarily low or absent in normal adult tissues, are highly upregulated in cancer. Finally, we show that the MTHFD2 enzyme in this pathway is highly expressed on the protein level in a variety of human tumours and negatively correlates with survival in breast cancer patients.

Results

Meta-analysis of enzyme mRNA expression in human tumours.

To systematically investigate expression of metabolic pathways across multiple tumour types, we first searched the GeneChip Oncology Database (GCOD) for studies containing primary tumour tissue samples and suitable normal tissue controls. We found 51 independent data sets satisfying this criterion, covering a total of 1,981 tumours of 19 different types versus 931 matched normal tissue controls (Supplementary Data 1), and interrogating a total of 20,103 genes. To avoid artifacts from comparison across different array platforms and laboratories, we calculated the differential expression (quantified by Z-scores) within each data set and estimated the statistical significance for each gene by permutation tests. While the extent of differential expression varied widely across studies (Fig. 1a)—likely reflecting variability in tumour types, nature of the control tissue and experimental design—we reasoned that genes consistently differentially expressed across these cancers would represent processes of fundamental importance to transformed cells. We therefore scored each gene by counting the number of data sets where differential expression was detected at a 5% false discovery rate (FDR) (Fig. 1b, full genome-wide analysis available as Supplementary Data 2). Most high-scoring overexpressed genes have previously been found to be expressed in the committed S, G2 and M phases of the cell cycle, where cells are thought to be most vulnerable to pharmacologic intervention. This likely reflects the high proportion of proliferating cells in tumours compared with control tissues.

We focused our analysis on 1,454 metabolic enzymes annotated in a previously established model of the human metabolic network. Among the top 50 consistently overexpressed enzymes (Table 1), we recovered several metabolic pathways previously associated with cancer, including multiple enzymes involved in glycolysis, de novo synthesis and salvage of nucleotides and in particular deoxynucleotides, as well as prolyl hydroxylases responsive to hypoxia and glycosylation enzymes, in agreement with a recent independent study. Among these metabolic enzymes, we recovered a number of genes targeted by existing cancer therapeutics, including DHFR (target of methotrexate), TYMS (target of 5-fluorouracil) and RRM2 (target of gencitabine) (Fig. 1b; Table 1). Beyond metabolic enzymes, our analysis also identified a number of other enzymes targeted by chemotherapeutics, including topoisomerase 2A, aurora A kinase and cyclin-dependent kinase 4 (Fig. 1b; Supplementary Data 2), as consistently overexpressed in cancers. Hence, our systematic analysis of metabolic reprogramming identifies a number of established cancer-related metabolic enzymes and pathways across varied tumour types.

Our analysis additionally reveals several metabolic enzymes consistently underexpressed in cancer and whose roles in cancer have not been previously appreciated. Among the metabolic enzymes consistently underexpressed in cancer (Table 2), we noted a number of enzymes related to fatty acid metabolism, including catalysis of branched fatty acids (ACOX2, ETFDH), synthesis of ketones (HMGC2S) and fatty acid synthesis (ACAB). A number of antioxidant enzymes were also consistently downregulated in cancer (Fig. 1c; Table 2). Taken together, these findings suggest consistent downregulated metabolic pathways and enzymes across multiple tumour types, which remain to be explored.

Cancer cells have previously been proposed to exhibit stem cell-like properties, expressing genes characteristically found in embryonic cells. Of the 35 enzymes previously established to be expressed by embryonic stem cells, a large fraction were overexpressed in tumours (Fig. 1e), suggesting a tendency to revert to an embryonic-like metabolic programme. Among the top 50 consistently overexpressed enzymes (Table 1) identified in our study, 26 enzymes have previously been evaluated using large-scale RNA interference (RNAi) screening to determine those essential for tumorigenesis in vitro and in vivo in breast cancer cells. Among these 26 metabolic genes, six genes (CTPS, GAPDH, PYCR1, MTHFD2, TP11 and TSTA3) were found to be required for tumorigenesis in vitro and five genes (CTPS, GAPDH, GMPS, PYCR1 and TP11) in vivo. Whereas several of these genes are components of glycolytic or nucleotide metabolism pathways previously associated with tumorigenesis; TSTA3, PYCR1 and MTHFD2 are members of pathways less appreciated in the context of cancer. TSTA3 catalyses the production of GDP-1-fucose (Fig. 1c), which serves as the substrate for fucosyltransferases. Increased levels of GDP-1-fucose and fucosylated glycoproteins have been noted in cancer, their role in tumorigenesis remains unclear. Similarly, our analysis also...
Identified the mitochondrial enzyme PYCR1 (Fig. 1c), an established target of the oncogene myc\(^2\). PYCR1 catalyses the final step of proline synthesis from glutamate (Fig. 1c), including in cancer cells\(^2\). In conjunction with its cytosolic isoform, PYCR1 may serve as a cycle for the transfer of reducing equivalents from the cytosol into the mitochondrion\(^2\) modulating both intracellular redox status\(^2\) and sensitivity to oxidant injury\(^2\).

Among all 1,454 enzymes examined, the metabolic enzyme most consistently overexpressed in tumours was the mitochondrial folate-coupled dehydrogenase MTHFD2 (Fig. 1a,b), ranked within the top three of all 20,103 genes interrogated (Supplementary Data 2). MTHFD2 is integral to mitochondrial one-carbon metabolism (Fig. 1d), a metabolic system recently implicated in rapid cancer cell proliferation\(^2\). MTHFD2 is a

**Figure 1 | Transcriptional regulation of metabolic pathways in human tumours.** (a) Differential expression (Z-score) distributions for 51 tumour-versus-normal data sets representing 19 tumour types are shown as violin plots (grey). Dots indicate Z-scores for MTHFD2 in each study; red colour denotes significance at \(< 5\% \) FDR. (b) Distribution of meta-analysis scores for all 20,103 genes interrogated across the 51 data sets. Gene symbols indicated, see text for further description. (c) Metabolic pathways detected as strongly overexpressed (among top 50 metabolic enzymes; red gene symbols) or underexpressed (among bottom 50 metabolic enzymes; blue gene symbols) in tumours. (d) Schematic of one-carbon metabolism with overexpressed (red) and underexpressed (blue) genes indicated. Grey symbols, not measured. (e) Gene set enrichment analysis for the set of 35 embryonic metabolic enzymes compared with mRNAs for all enzymes.
A bifunctional enzyme, catalysing the NAD$^+$-dependent CH$_2$-THF dehydrogenase and CH$^+$-THF cyclohydrolase reactions within the mitochondria. Within the mitochondrial one-carbon metabolism, we also noted frequent overexpression of the preceding enzyme, SHMT2, which catalyses the production of glycine and one-carbon groups in the form of CH$_2$-THF from serine (Fig. 1d). The subsequent enzyme, MTHFD1L, catalysing the synthesis of formate and regeneration of the THF cofactor (Fig. 1d), was only measured on 11/51 data sets, and our meta-analysis was not well-powered in this case; however, it appeared

| Score | Symbol | Description | Drug target | One-carbon | Nucleotides | Glycolysis | Hypoxia | Glycosylation |
|-------|--------|-------------|-------------|------------|-------------|------------|---------|--------------|
| 27    | MTHFD2 | Methylene-THF dehydrogenase/cyclohydrolase | (mitochondrial) |            |             |            |         |              |
| 26    | RRM2   | Ribonucleotide reductase M2 |            | X          | X           |            |         |              |
| 25    | NME1   | Protein expressed in non-metastatic cells 1 |            | X          |             |            |         |              |
| 23    | GMP5   | Guanine monophosphate synthetase |            |            |             |            |         |              |
| 23    | SHMT2  | Serine hydroxymethyltransferase 2 (mitochondrial) |            |            |             | X          |         |              |
| 22    | GGCT   | Gamma-glutamylcyclotransferase |            |            |             |            |         |              |
| 22    | UCK2   | Uridine-cytidine kinase 2 |            | X          |             |            |         |              |
| 22    | TYMS   | Thymidylate synthetase |            | X          | X           |            |         |              |
| 21    | TK1    | Thymidine kinase 1, soluble |            | X          |             |            |         |              |
| 21    | PLOD3  | Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 |            |            |             | X          |         |              |
| 20    | ENO1   | Enolase 1 (alpha) |            |             |             |            |         | X            |
| 20    | PYCR1  | Pyrroline-5-carboxylate reductase 1 |            |            |             |            |         |              |
| 19    | DNMT1  | DNA (cytosine-5-)-methyltransferase 1 |            |            |             |            |         |              |
| 19    | FLAD1  | Flavin adenine dinucleotide synthetase homologue (S. cerevisiae) |            |            |             |            |         | X            |
| 19    | ALG3   | Asparagine-linked glycosylation 3 homologue (S. cerevisiae) |            |            |             | X          |         |              |
| 18    | PAICS  | AIR carboxylase/SAICAR synthetase |            |            |             |            |         | X            |
| 18    | DTYMK  | Deoxythymidylate kinase (thymidylate kinase) |            | X          |             |            |         |              |
| 18    | ATIC   | AICAR formyltransferase/IMP cyclohydrolase |            |            |             | X          |         |              |
| 17    | MOGS   | Mannosyl-oligosaccharide glucosidase |            |            |             | X          |         |              |
| 17    | PLOD1  | Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1 |            |            |             | X          |         |              |
| 17    | PFAH1B3| Platelet-activating factor acetylhydrolase 1b, catalytic subunit 3 |            |            |             |            |         |              |
| 17    | P4HA1  | Prolin 4-hydroxylase, alpha polypeptide 1 |            |            |             | X          |         |              |
| 16    | RPN1   | Ribophorin I |            |            |             | X          |         |              |
| 16    | PKM2   | Pyruvate kinase, muscle |            |            |             | X          |         |              |
| 16    | LDHA   | Lactate dehydrogenase A |            |            |             | X          |         |              |
| 16    | HMB5   | Hydroxymethylbilane synthase |            |            |             |            |         |              |
| 15    | ALDOA  | Aldolase A, fructose-bisphosphate |            |            |             | X          |         |              |
| 15    | SRM    | Spermidine synthase |            |            |             |            |         |              |
| 15    | GAPDH  | Glyceraldehyde-3-phosphate dehydrogenase |            |            |             | X          |         |              |
| 15    | DHFR   | Dihydrofolate reductase |            |             |             | X          |         | X            |
| 15    | CTPS1  | CTP synthase 1 |            |            |             | X          |         |              |
| 15    | CAD    | Carbamoyl-P synthetase 2/aspartate transcarbamylase/dihydroorotase |            |            |             | X          |         |              |
| 15    | ACLY   | ATP citrate lyase |            |            |             |            |         | X            |
| 14    | ALG8   | Asparagine-linked glycosylation 8 (S. cerevisiae) |            |            |             | X          |         |              |
| 14    | CHPF2  | Chondroitin polymerizing factor 2 |            |            |             |            |         |              |
| 14    | UMP5   | Uridine monophosphate synthetase |            |            |             | X          |         |              |
| 14    | SRD5A1 | Steroid-5-alpha-reductase, alpha polypeptide 1 |            |            |             |            |         | X            |
| 14    | SLC7A1 | Solute carrier family 7 (y + system), member 1 |            |            |             |            |         |              |
| 14    | PMM2   | Phosphomannomutase 2 |            |            |             | X          |         |              |
| 14    | INPPL1 | Inositol polyphosphate phosphatase-like 1 |            |            |             | X          |         |              |
| 14    | APRT   | Adenine phosphoribosyltransferase |            |            |             | X          |         |              |
| 13    | CHPF   | Chondroitin polymerizing factor |            |            |             |            |         |              |
| 13    | B3GAT3 | Beta-1,3-glucuronosyltransferase 3 (glucuronosyltransferase 1) |            |            |             | X          |         |              |
| 13    | EHMT2  | Euchromatic histone-lysine N-methyltransferase 2 |            |            |             |            |         | X            |
| 13    | SLC25A13| Solute carrier family 25 (aspartate/glutamate carrier), member 13 |            |            |             |            |         |              |
| 13    | DPM2   | Dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory |            |            |             | X          |         |              |

Score denotes number of data sets (of 51 total), where each gene was significantly overexpressed at a FDR of 5%, as in Fig. 1. Rightmost columns indicate membership in pathways, as in Fig. 1.
consistently overexpressed in those 11 data sets (data not shown).

In the paralogous cytosolic one-carbon pathway, the NADP$^+$-linked, trifunctional MTHFD1 enzyme also exhibited overexpression in multiple cancers, although it did not reach the top 50 threshold (Fig. 1d). This is consistent with the notion that the mitochondrial one-carbon system serves to produce formate that is then coupled to THF in the cytosol by the MTHFD1 enzyme.27 Interestingly, the cytosolic enzyme ALDH1L1, which breaks down cytosolic formyl-THF into CO$_2$ and THF and thus opposes the synthesis of one-carbon units, was consistently underexpressed in tumours (Fig. 1d; Table 2). The paralogous mitochondrial enzyme, ALDH1L2 was not reliably measured on the microarray platforms used (Fig. 1d). The aminomethyltransferase AMT, a component of the enzyme system that catabolizes glycine, was also underexpressed (Fig. 1d; Table 2). These findings suggest that the metabolic pathway channelling one-carbon units through the mitochondria via MTHFD2 has an important role in cancer biology.

Enzymes in transformed and normal proliferating cells. Many of the metabolic changes accompanying transformation are also active in normal proliferating cells4. To identify cancer-related metabolic enzymes more specific to the transformed state, we analysed gene expression from 80 normal human tissues, including proliferative tissues such as intestine and bone (Table 2).
marrow, compared with transformed cell lines. The well-known chemotherapeutic targets DHFR, RRM2, TYMS and topoisomerase 2A (TOP2A) were expressed in several normal proliferating cell types (Fig. 2a), in which the adverse side effects of drugs targeting these enzymes are commonly observed. In contrast, MTHFD2 exhibited little expression in these tissues, and was in fact the most cancer-specific messenger RNA (mRNA) measured on these arrays, as quantified by a transformed/normal expression ratio (Fig. 2b). In this analysis, we also noted the PYCR1 enzyme (Fig. 2b), which was also among the most commonly overexpressed enzymes in cancer (Table 1). MTHFD2 was not detected in serum-stimulated normal human fibroblasts in vitro or in hepatocytes proliferating in vivo in response to partial liver resection (Fig. 2c), whereas other established enzymes (Fig. 2b), which was also among the most commonly overexpressed enzymes in cancer (Table 1). MTHFD2 was not detected in serum-stimulated normal human fibroblasts in vitro or in hepatocytes proliferating in vivo in response to partial liver resection (Fig. 2c), whereas other established

Figure 2 | MTHFD2 expression in transformed proliferating cells compared with normal proliferating cells. (a) mRNA expression levels in normal post-mitotic (open bars), normal proliferating (black bars) and transformed (red bars) cells or tissues for MTHFD2 and four established cancer drug targets. Normal hematopoietic cell fractions with strong expression are indicated. (b) Quantile–quantile plot for the ratio of minimal expression among transformed cells to maximum expression among normal (proliferating and post-mitotic) cells, defined as in (a), for each of the 12,529 human mRNAs measured. Randomized quantiles (X-axis) were obtained by permuting samples. MTHFD2 mRNA is indicated. (c) mRNA expression of human MTHFD2 or mouse Mthfd2 and four established cancer drug targets during mouse liver regeneration following partial hepatectomy (left); human fibroblasts proliferating in response to serum stimulation (center); and human CD4+ T lymphocytes activated by CD3 and CD28 antibodies (right). Red line indicates Mthfd2/MTHFD2. (d) mRNA expression of human MTHFD2 (left) and mouse Mthfd2 (right) during early embryonic development. Error bars denote s.d. (n = 3). (e) mRNA expression of mouse Mthfd2 during embryonic development of liver (left) and hypothalamus (right). Error bars denote s.d. (n = 2).
chemotherapeutic drug targets were induced in these normal proliferating cells. However, we did observe MTHFD2 induction at an early time point in activated lymphocytes (Fig. 2c), consistent with previously observed MTHFD2 activity in bone marrow28, indicating a possible role for MTHFD2 in normal hematopoietic cells.

MTHFD2 enzymatic activity has previously been detected in mouse embryonic fibroblasts29 as well as in human embryonic stem cells30, and Mthfd2 deletion is embryonic lethal in mice29. We found that the MTHFD2 mRNA increases during the initial rounds of cellular division in both mouse and human fertilized oocytes (Fig. 2d). In addition, expression is high in mouse fetal liver and hypothalamus, but decreases markedly around birth (Fig. 2e). These observations support a role for the enzyme in embryonic development.

RNA interference targeting MTHFD2 causes cancer cell death. To further evaluate whether expression of the MTHFD2 gene product is required for cancer cell proliferation, we used two different RNAi modalities to silence MTHFD2 across 16 diverse cancer cell types. In total, we used two lentiviral-delivered short hairpin RNAs (shRNAs) and six sequence-independent non-viral-delivered short interfering RNA (siRNA) oligonucleotides (Supplementary Figs 1,2). MTHFD2 mRNA, protein levels and enzymatic activity were substantially reduced by shRNA (Supplementary Figs 1a–c and 3) or siRNA targeting MTHFD2 (Supplementary Fig. 1e). In most cell lines, proliferation was severely reduced, consistent with prior observations31, with noted variability in the degree of defect among the various cell lines (Supplementary Fig. 1d,f). Propidium iodide (PI) staining and flow cytometry revealed marked cell death at the 48 h time point, with 40% of cells nonviable (Supplementary Fig. 1g,h), while non-targeting control siRNA transfections did not impair cell viability (Supplementary Fig. 1i). While RNAi-mediated loss of the MTHFD2 gene product was associated with slowed cancer cell proliferation and marked cell death, we have hitherto been unable to rescue this phenotype by exogenous expression of the MTHFD2 cDNA, which is required to definitively prove that the effect is due to the enzymatic activity of MTHFD2 protein and not a consequence of silencing its RNA or an off-target effect.

MTHFD2 protein expression in tumours and normal tissues. To evaluate whether the MTHFD2 protein is overexpressed in human tumours, we performed immunohistochemistry in 176 tumour samples, collected from 16 tumour types. Strong or moderate staining for MTHFD2 was observed in 12 of these tumour types, and the protein was detected in all cancer types with the exception of gliomas (Fig. 3a). While staining intensity was variable across tumour samples, the protein consistently appeared specific to transformed cells within the tumours, with little or no staining seen in adjacent stroma (Fig. 3b). These data indicate that the MTHFD2 protein is indeed present in multiple cancer types. Activity of the enzyme has also previously been observed in transformed cell lines30. We did observe staining in normal epithelium of the gut tissues, in particular in intestinal crypts containing stem cells, in tonsil lymphoid tissue and in exocrine pancreas (Supplementary Fig. 4). Finally, we evaluated the expression of MTHFD2 in six independent cohorts of patients with breast cancer followed for survival. Whereas there was heterogeneity among individual data sets (Supplementary Fig. 5a), likely reflecting differences in the number of patients and clinical parameters including, entry patient criteria and treatment strategies, a meta-analysis of all six independent cohorts indicated that high expression of the MTHFD2 mRNA

Figure 3 | MTHFD2 protein expression in human tumours. (a) Top, fraction of samples with none, weak, moderate or strong immunohistochemistry staining for MTHFD2 in transformed cells, across 16 solid tumour types. Bottom, same analysis as above for stromal cells. For each tumour type, tumours from nine to 12 individuals were examined. (b) Representative images from each of the 16 tumour types summarized in (a), exemplifying negative (N), weak (W), moderate (M) or strong (S) staining intensities. Up arrow, stromal cells; down arrow, cancer cells. Scale bar represent 100 μm.
Discussion

In this study, we have identified metabolic enzymes and pathways that are frequently overexpressed in tumours, and demonstrated that one of these enzymes, MTHFD2, is broadly required for cancer cell proliferation and viability. It should be emphasized that our meta-analysis only addresses changes in enzyme expression at the mRNA level, and our results do not exclude that other enzymes may be dysregulated in cancer by post-transcriptional mechanisms such as translational control or allosteric regulation. Moreover, while our meta-analysis was designed to reliably detect genes that are frequently overexpressed in tumours, it will likely miss genes absent or not well measured on common microarray platforms, and does not consider tissue-specific phenomena. Future studies investigating differences in metabolic enzyme expression between tumour types, as well as co-occurrence of enzymes and pathways in tumours, would be a valuable extension to this work.

Within our systematic analysis of metabolic reprogramming, we identified a number of metabolic genes whose expression was altered in a variety of tumour types relative to normal controls. Among the metabolic enzymes that were consistently over-expressed in tumours (Table 1), we find a high proportion of established cancer-related metabolic enzymes and known chemotherapeutic drug targets, suggesting that the additional enzymes identified here may be of interest as cancer targets as well. These additional enzymes include MTHFD2 and PYCRL1, both of which are low or absent in a large panel of normal tissues, including the proliferative tissues of the gut and a number of immune cell types; this could theoretically limit on-target side effects typically associated with these cell types. We further demonstrate the necessity of the MTHFD2 gene in cancer cell proliferation and survival across a number of cancer cell lines. While the proliferation defect and early cell death phenotypes with MTHFD2 knockdown were observed with multiple independent RNAi sequences, we cannot definitively rule out off-target RNAi effects. Moreover, while our results indicate a requirement for the MTHFD2 mRNA, further studies are needed to prove that the MTHFD2 protein and its enzymatic activity are indeed mediating the observed effects on cancer cell proliferation.

Genes consistently underexpressed in tumours include metabolic enzymes involved in fatty acid metabolism, including catabolism of branched fatty acids and ketogenesis as well as a number of antioxidant enzymes (Table 2). The consistent repression of these metabolic enzymes suggests a larger metabolic reprogramming that occurs with transformation in varied tumour types, and may reflect an effort on the part of a cancer cell to shunt metabolites to particular biosynthetic pathways rather than catabolism, to limit the production of toxic intermediates and/or to preferentially shuttle reducing equivalents to compartmentalized within human cells, these recent studies indicate that compartment-specific alterations may be critical in promoting cancer growth and survival. The recognition that compartmentalization of one-carbon folate metabolism is altered in cancer raises new opportunities for identifying novel therapies for cancer, as well as for targeting existing antifolate therapies with greater precision.

Methods

Microarray data analysis. Microarray data sets were selected from the GCOID according to the following criteria: we considered only data sets of human tumours with normal tissue samples as controls, excluding cell lines and cultured normal cells; we required at least three independent samples in both tumour and control groups; and we considered only data from the Affymetrix HG-Focus, HG-U133A, HG-U133A2, HG-U133 + v2, HG_U95A and HG_U95Av2 microarray platforms. A list of data sets included is provided in Supplementary Data 1. All data sets were uniformly processed from the Affymetrix CEL using the Robust multi-array average algorithm, as described. For each microarray platform, in cases of multiple probe sets per gene, we selected the probe set with a maximal mean signal rank across all arrays. Differential expression between tumour and normal groups was quantified for each data set using Z-scores, and nominal P-values of differential expression were calculated for each gene using a permutation test (1,000 permutations). FDRs were computed using the Benjamini–Hochberg procedure.
underexpressed. Enrichment analysis was done using the GSEA-p method with the parameter set p = 0.01.

The compendium of 80 human tissues was generated from duplicate samples for each tissue using Affymetrix U133A arrays. The gRNA-normalized data was downloaded from www.biogps.org, and duplicates were averaged for the analysis presented in Fig. 2. Our classification of tissues and cells as normal/postmitotic, normal/proliferating or transformed is described in Supplementary Table 1.

Expression data from proliferating human fibroblasts, human T cells, mouse regenerating liver and mouse/human embryonic development was obtained from the NCBI Gene Expression Omnibus, series accession GSE3945, GSE2770, GSE6999, and GSE18290, respectively. No additional normalization was performed.

Cell culture. Human cancer cell lines (passage number 5–18 in all cases except HCT-116) were obtained from the National Cancer Institute. All cell lines were cultured in RPMI-1640 medium (Invitrogen) with 2 mM L-glutamine and 5% fetal bovine serum (FBS) (HyClone Laboratories). Cells were cultured at 37°C in 5% CO2.

shRNA knockdown of MTHFD2. Cells were cultured according to standard techniques as described above. Lentiviral vectors (pLKO.1) expressing shRNA clones were generated by the Broad RNAi Consortium, as previously described26. Sequence-independent shRNAs were generated against human MTHFD2 (sh50, sh53) or to a scrambled vector control (shCtrl) not matching any human gene sequences are: shCtrl, Clone ID TRCN0000072232, Clone name lacZ_27s1c1, Target sequence 5'-CCTGCGTATTACAAAGTCGTAAG-3'; sh52-50, Clone ID TRCN0000036500, Clone name NM_006636.2-5481s1c1, Target sequence 5'-CGCGTGTGTTTGTACATGAT-3'; sh52-53, Clone ID TRCN0000036533, Clone name NM_006636.2-5481s1c1, Target sequence 5'-CGAGTTGAAGAAACATACAAT-3', sh50, Clone ID TRCN0000072232, Clone name lacZ_27s1c1, Target sequence 5'-CCTG CGTATTACAAAGTCGTAAG-3'. Cells were infected with shRNA lentiviral vectors in six-well dishes as cell cultures. siCtrl-2. Cells were transfected 24 h after seeding (at 40–60% confluence) with X-tremeGENE 9 was used to transfect 293T cells according to the manufacturer's instructions. Transfections were performed in 6-well plates with 100 ng VSVG, 900 ng pCMV-dR8.91 and 1 μg plasmid with 6 μl transfection reagent pooled and frozen at 80°C in single-use aliquots26.

Lentiviral infection was performed as previously described26. Cells were seeded at a density of 20,000 cells per well for all cell types, except for control cells, where a density of 40,000 cells per well was used. Transfection. Western blot analysis was performed on cells seeded at 150,000 cells per well in a six-well plate and lysed in 2 x Lysis buffer (8% sodium dodecyl sulfate, 200 mM Tris, pH 6.8, 10% glycerol, 1% Triton X-100,000, and 5 μl of 20 mM Reduced GSH). Protein was measured using the Pierce BCA Protein Assay Kit (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierce) with a detergent based assay (Pierc...
