

ARTICLE

Received 1 Nov 2013 | Accepted 17 Dec 2013 | Published 23 Jan 2014

DOI: 10.1038/ncomms4128

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

Roland Nilsson^{1,2,*}, Mohit Jain^{3,4,5,6,*},†, Nikhil Madhusudhan^{3,4,5}, Nina Gustafsson Sheppard^{1,2}, Laura Strittmatter^{3,4,5}, Caroline Kampf⁷, Jenny Huang⁸, Anna Asplund⁷ & Vamsi K. Mootha^{3,4,5}

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.

¹Unit of Computational Medicine, Department of Medicine, Karolinska Institutet, 17176 Stockholm, Sweden. ²Center for Molecular Medicine, Karolinska Institutet, 17176 Stockholm, Sweden. ³Broad Institute, Cambridge, Massachusetts 02142, USA. ⁴Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁵Department of Molecular Biology and the Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. ⁶Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. ⁷Department of Immunology, Genetics & Pathology, Science for Life Laboratory, Uppsala University, 75185 Uppsala, Sweden. ⁸La Jolla Institute for Allergy and Immunology, San Diego, California 92037, USA. * These authors contributed equally to this work. † Present address: Department of Medicine and Pharmacology, University of California, San Diego, California, USA. Correspondence and requests for materials should be addressed to R.N. (email: roland.nilsson@ki.se) or to V.K.M. (email: vamsi@hms.harvard.edu).

Divergent metabolism in tumours was first recognized nearly a century ago¹, 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 therapy². Furthermore, the dependence of cancer cells on nucleotide metabolism forms the basis for the use of several common chemotherapeutics, including agents targeting dihydrofolate reductase (DHFR), thymidylate synthase (TYMS) and ribonucleotide reductase (RRM2)³. Moreover, recent data have suggested that many of the growth factor signalling pathways commonly perturbed in cancer impinge on metabolic enzymes⁴, as well as that metabolic enzymes may act as *bona fide* oncogenes⁵ and even transform cells⁶. 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^{7,8}, the complex, coordinated changes in metabolism that occur during cancer transformation have only begun to be understood⁹.

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^{10,11}. 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 scale⁹. 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 index³.

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)¹⁰ 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^{12,13}, 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 network¹⁴. Among the top 50 consistently overexpressed enzymes (Table 1), we recovered several metabolic pathways previously associated with cancer, including multiple enzymes involved in glycolysis¹⁵, *de novo* synthesis and salvage of nucleotides⁷ and in particular deoxynucleotides, as well as prolyl hydroxylases responsive to hypoxia¹⁶ and glycosylation enzymes¹⁷ (Fig. 1c,d), in agreement with a recent independent study⁹. 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 gemcitabine) (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 catabolism of branched fatty acids (*ACOX2*, *ETFDH*), synthesis of ketones (*HMGCS2*) and fatty acid synthesis (*ACACB*). 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 cells¹⁸. Of the 35 enzymes previously established to be expressed by embryonic stem cells¹⁸, 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 cells¹⁹. Among these 26 metabolic genes, six genes (*CTPS*, *GAPDH*, *PYCR1*, *MTHFD2*, *TPI1* and *TSTA3*) were found to be required for tumorigenesis *in vitro* and five genes (*CTPS*, *GAPDH*, *GMPS*, *PYCR1* and *TPI1*) *in vivo*¹⁹. 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-*L*-fucose (Fig. 1c), which serves as the substrate for fucosyltransferases. Increased levels of GDP-*L*-fucose and fucosylated glycoproteins have been noted in cancer²⁰; their role in tumorigenesis remains unclear. Similarly, our analysis also

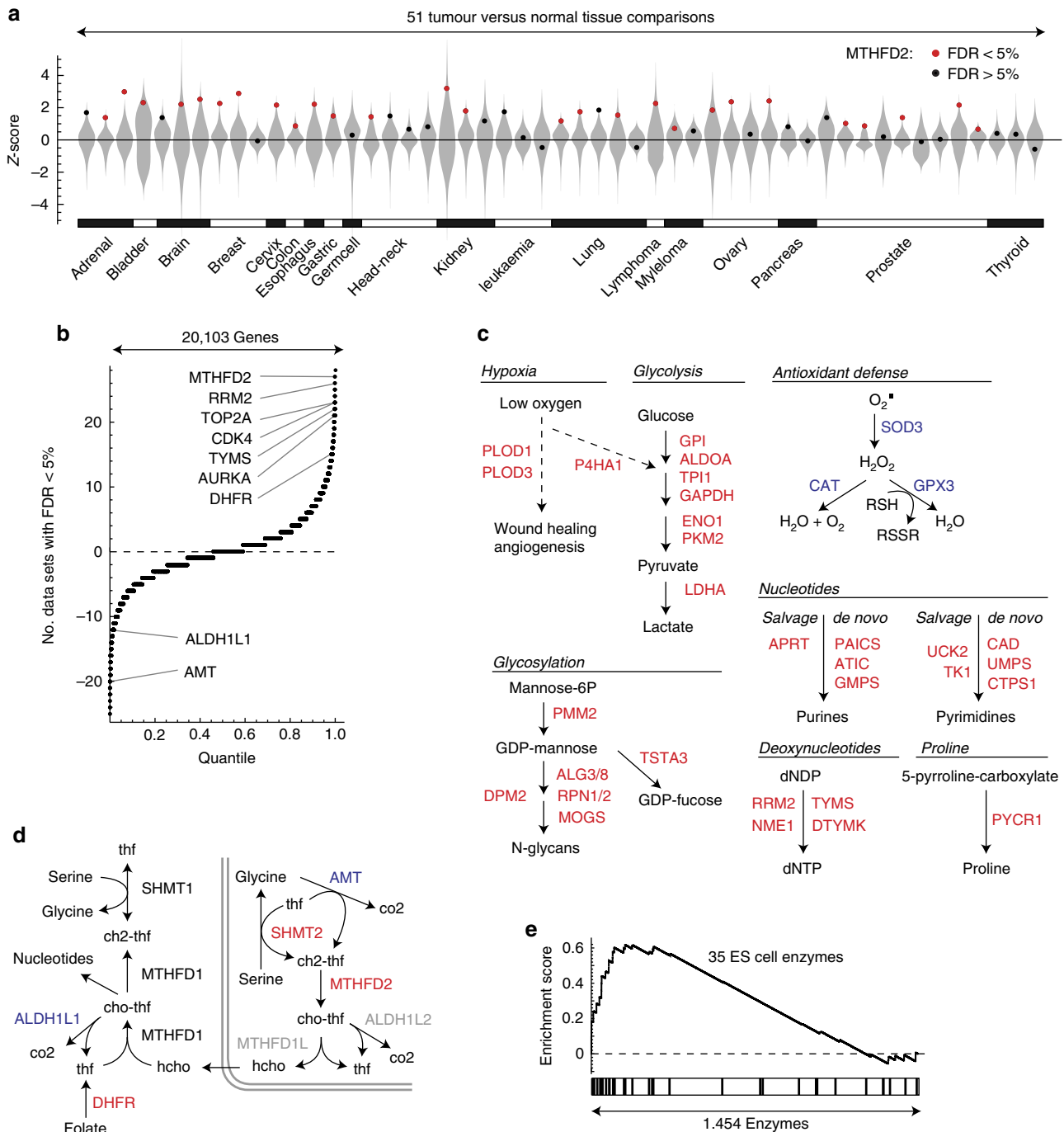


Figure 1 | Transcriptional regulation of metabolic pathways in human tumours. (a) Differential expression (Z-score) distributions for 51 tumour-versus-normal data sets representing 19 tumour types are shown as violin plots (grey). Dots indicate Z-scores for *MTHFD2* in each study; red colour denotes 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.

identified the mitochondrial enzyme *PYCR1* (Fig. 1c), an established target of the oncogene *myc*²¹. *PYCR1* catalyses the final step of proline synthesis from glutamate (Fig. 1c), including in cancer cells²². In conjunction with its cytosolic isoform, *PYCR1* may serve as a cycle for the transfer of reducing equivalents from the cytosol into the mitochondrion²³, modulating both intracellular redox status²⁴ and sensitivity to oxidant injury²⁵.

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 proliferation²⁶. *MTHFD2* is a

Table 1 | Top 50 metabolic enzymes frequently overexpressed in human tumours.

Score	Symbol	Description	Drug target	One-carbon	Nucleotides	Glycolysis	Hypoxia	Glycosylation
27	MTHFD2	Methylene-THF dehydrogenase/cyclohydrolase (mitochondrial)		X				
26	RRM2	Ribonucleotide reductase M2	X		X			
25	NME1	Protein expressed in non-metastatic cells 1			X			
23	GMPS	Guanine monophosphate synthetase			X			
23	SHMT2	Serine hydroxymethyltransferase 2 (mitochondrial)		X				
22	GGCT	Gamma-glutamylcyclotransferase						
22	UCK2	Uridine-cytidine kinase 2			X			
22	TYMS	Thymidylate synthetase	X		X			
22	TK1	Thymidine kinase 1, soluble			X			
21	PLOD3	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3					X	
21	ENO1	Enolase 1 (alpha)				X		
20	PYCR1	Pyrroline-5-carboxylate reductase 1						
20	DNMT1	DNA (cytosine-5-)-methyltransferase 1						
19	FLAD1	Flavin adenine dinucleotide synthetase homologue (<i>S. cerevisiae</i>)						
19	ALG3	Asparagine-linked glycosylation 3 homologue (<i>S. cerevisiae</i>)						X
19	TSTA3	Tissue-specific transplantation antigen P35B						X
19	TPI1	Triosephosphate isomerase 1				X		
19	RPN2	Ribophorin II						X
19	GPI	Glucose-6-phosphate isomerase				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	PAFAH1B3	Platelet-activating factor acetylhydrolase 1b, catalytic subunit 3						
17	P4HA1	Prolyl 4-hydroxylase, alpha polypeptide I					X	
16	RPN1	Ribophorin I						X
16	PKM2	Pyruvate kinase, muscle				X		
16	LDHA	Lactate dehydrogenase A				X		
16	HMBS	Hydroxymethylbilane synthase						
16	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						
14	ALG8	Asparagine-linked glycosylation 8 (<i>S. cerevisiae</i>)						X
14	CHPF2	Chondroitin polymerizing factor 2						
14	UMPS	Uridine monophosphate synthetase			X			
14	SRD5A1	Steroid-5-alpha-reductase, alpha polypeptide 1						
14	SLC7A1	Solute carrier family 7 (y+ system), member 1						
14	PMM2	Phosphomannomutase 2						X
14	INPPL1	Inositol polyphosphate phosphatase-like 1						
14	APRT	Adenine phosphoribosyltransferase			X			
13	CHPF	Chondroitin polymerizing factor						
13	B3GAT3	Beta-1,3-glucuronyltransferase 3 (glucuronosyltransferase I)						
13	EHMT2	Euchromatic histone-lysine N-methyltransferase 2						
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.

bifunctional enzyme, catalysing the NAD^+ -dependent CH_2 -THF dehydrogenase and CH^+ -THF cyclohydrolase reactions within the mitochondria²⁷. 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, *MTHFDIL*, 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

Table 2 | Top 50 metabolic enzymes frequently underexpressed in human tumours.

Score	Symbol	Description
-24	ASPA	Aspartoacylase
-23	AOX1	Aldehyde oxidase 1
-22	GPX3	Glutathione peroxidase 3 (plasma)
-21	ADH1B	Alcohol dehydrogenase 1B (class I), beta polypeptide
-20	PTGDS	Prostaglandin D2 synthase 21 kDa (brain)
-20	MAOB	Monoamine oxidase B
-20	AMT	Aminomethyltransferase
-19	COX7A1	Cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)
-17	UST	Uronyl-2-sulfotransferase
-17	SOD3	Superoxide dismutase 3, extracellular
-17	PTGIS	Prostaglandin I2 (prostacyclin) synthase
-17	ADH1C	Alcohol dehydrogenase 1C (class I), gamma polypeptide
-16	PIP5K1B	Phosphatidylinositol-4-phosphate 5-kinase, type I, beta
-16	PDE2A	Phosphodiesterase 2A, cGMP-stimulated
-16	CYP3A5	Cytochrome P450, family 3, subfamily A, polypeptide 5
-16	ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)
-16	ALDH1A1	Aldehyde dehydrogenase 1 family, member A1
-15	ECHDC2	Enoyl CoA hydratase domain containing two
-15	DUOX1	Dual oxidase 1
-15	ST6GALNAC2	ST6 -N-acetylgalactosaminide alpha-2,6-sialyltransferase 2
-15	PPAP2B	Phosphatidic acid phosphatase type 2B
-15	PLCL1	Phospholipase C-like 1
-15	IDI1	Isopentenyl-diphosphate delta isomerase 1
-15	ALDH9A1	Aldehyde dehydrogenase 9 family, member A1
-15	ACACB	Acetyl-CoA carboxylase beta
-14	MAN1C1	Mannosidase, alpha, class 1C, member 1
-14	CHST3	Carbohydrate (chondroitin 6) sulfotransferase 3
-14	MAOA	Monoamine oxidase A
-14	CYP11A1	Cytochrome P450, family 11, subfamily A, polypeptide 1
-14	ATP1A2	ATPase, Na ⁺ /K ⁺ transporting, alpha 2 polypeptide
-14	ALDH3A2	Aldehyde dehydrogenase 3 family, member A2
-14	ADH1A	Alcohol dehydrogenase 1A (class I), alpha polypeptide
-13	NMRK1	Nicotinamide riboside kinase 1
-13	INPP5A	Inositol polyphosphate-5-phosphatase, 40 kDa
-13	HSD11B1	Hydroxysteroid (11-beta) dehydrogenase 1
-13	CYP27A1	Cytochrome P450, family 27, subfamily A, polypeptide 1
-13	CHKB	Choline kinase beta
-13	CAT	Catalase
-12	ACOX2	Acyl-CoA oxidase 2, branched chain
-12	SRD5A2	Steroid-5-alpha-reductase, alpha polypeptide 2
-12	SLC6A4	Solute carrier family 6 (neurotransmitter transporter, serotonin), member 4
-12	PIK3R1	Phosphoinositide-3-kinase, regulatory subunit 1 (alpha)
-12	HMGCS2	3-Hydroxy-3-methylglutaryl-CoA synthase 2 (mitochondrial)
-12	GSTM1	Glutathione S-transferase mu 1
-12	GATM	Glycine amidinotransferase (L-arginine:glycine amidinotransferase)
-12	ETFDH	Electron-transferring-flavoprotein dehydrogenase
-12	CDO1	Cysteine dioxygenase, type I
-12	ATP6V1E1	ATPase, H ⁺ transporting, lysosomal 31 kDa, V1 subunit E1
-12	ATP1B2	ATPase, Na ⁺ /K ⁺ transporting, beta 2 polypeptide
-12	SLC25A4	Solute carrier family 25 (mitochondrial carrier; ANT), member 4

Negative scores denote number of data sets (of 51 total), where each gene was significantly underexpressed at a FDR of 5%, 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* enzyme²⁷. Interestingly, the cytosolic enzyme *ALDH1L1*, which breaks down cytosolic formyl-THF into CO₂ 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 amino-

methyltransferase *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

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

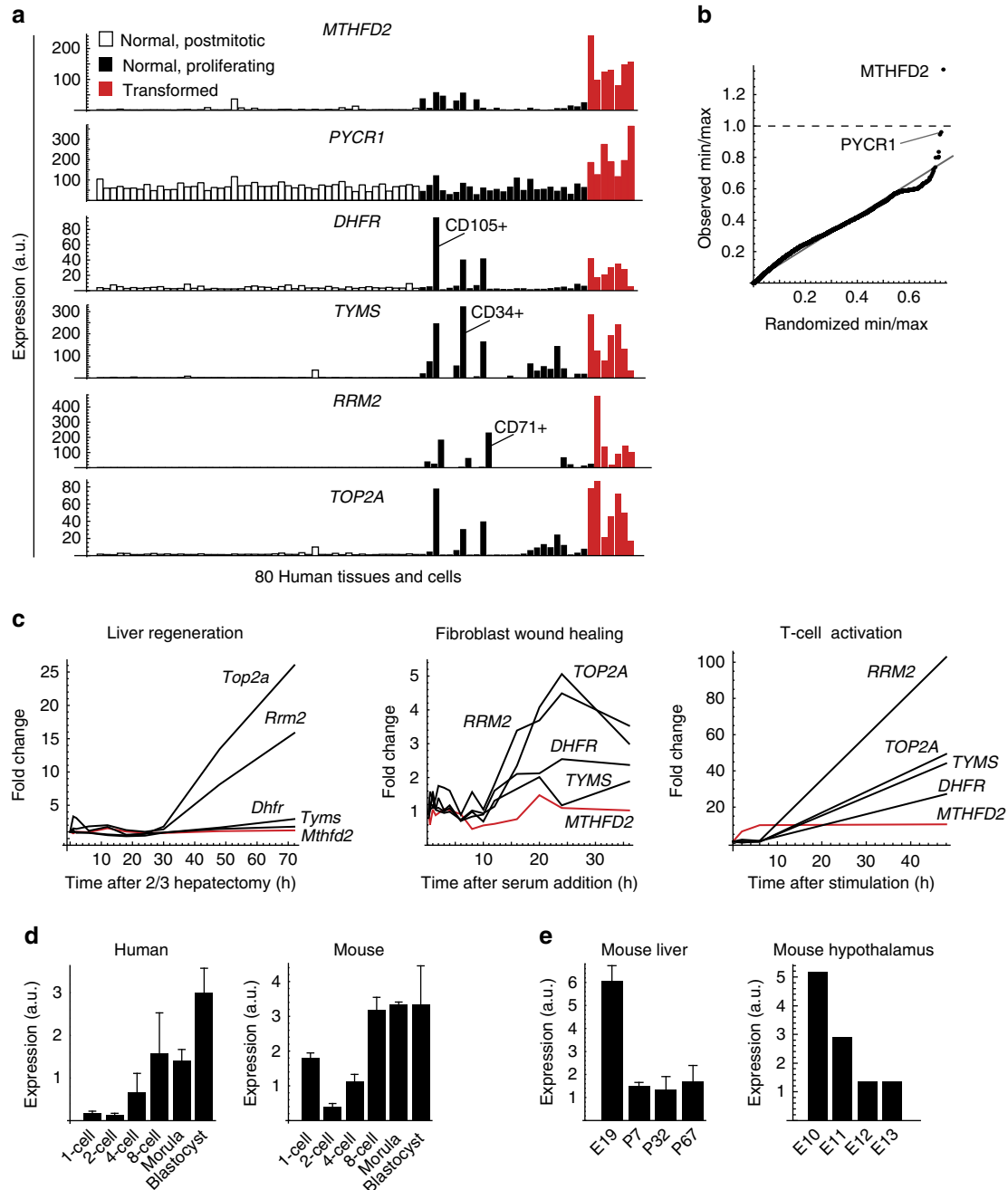


Figure 2 | *MTHFD2* expression in transformed proliferating cells compared with normal proliferating cells. (a) mRNA expression levels in normal post-mitotic (open bars), normal proliferating (black bars) and transformed (red bars) cells or tissues for *MTHFD2* and four established cancer drug targets. 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 fibroblasts²⁸ as well as in human embryonic stem cells¹⁸, 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 observations¹⁹, 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

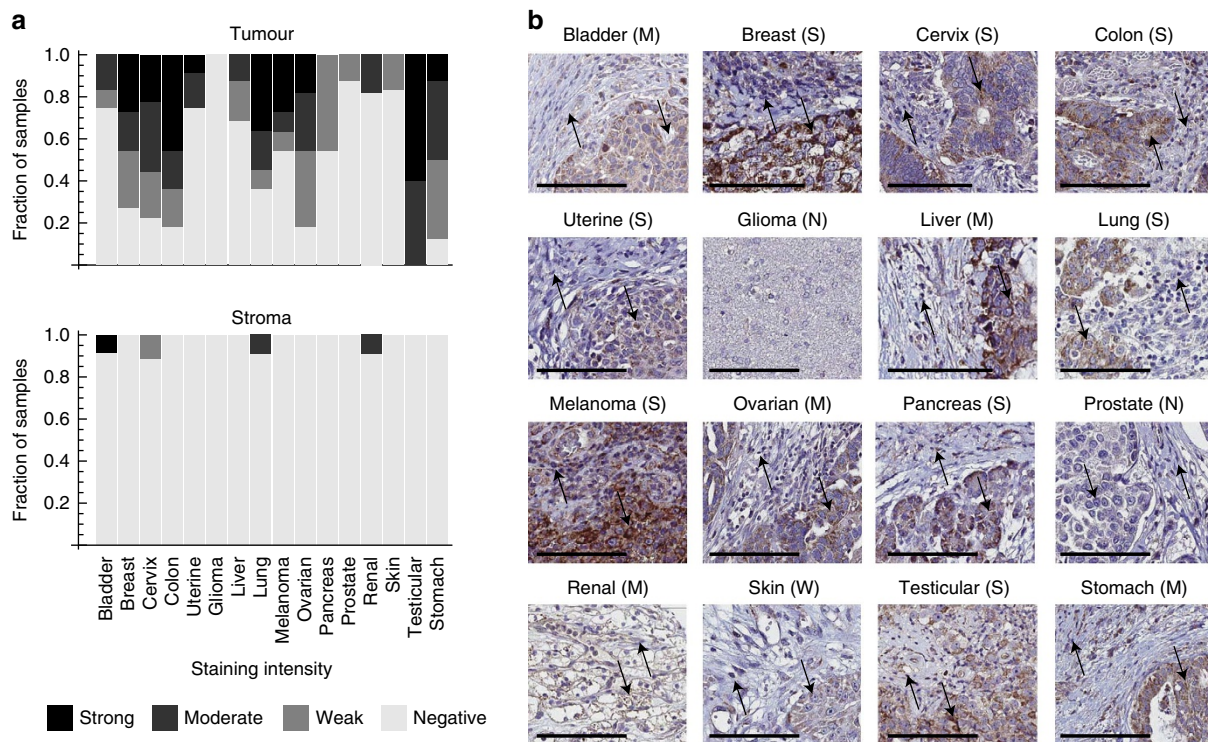


Figure 3 | *MTHFD2* protein expression in human tumours. (a) Top, fraction of samples with none, weak, moderate or strong immunohistochemistry staining for *MTHFD2* in transformed cells, across 16 solid tumour types. Bottom, same analysis as above for stromal cells. For each tumour type, tumours from 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 μ m.

was associated with increased mortality in breast cancer patients (Supplementary Fig. 5b), suggesting a potential role in tumour progression.

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 overexpressed 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 *PYCR1*, 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 particular compartments, such as the mitochondria. As with the overexpressed metabolic genes, further investigation of these underexpressed metabolic enzymes may uncover metabolic vulnerabilities, which may be targeted therapeutically.

While the enzymatic function of the *MTHFD2* enzyme is well established, its broader role in cancer cell metabolism is not yet clear. Interestingly, while *SHMT2* knockdown renders cancer cells auxotrophic for glycine²⁶, we find that *MTHFD2* knockdown results in death of human cancer cells in a complete medium containing glycine (Supplementary Fig. 1g,h). This observation indicates that the observed cell death on loss of *MTHFD2* is not due to blocking of glycine synthesis. Moreover, as *SHMT2* is

thought to be the main source of mitochondrial one-carbon units, this observation also suggests that depletion of formate, the other main metabolic product of the mitochondrial one-carbon pathway, is unlikely to be the cause of cell death on loss of *MTHFD2*. In support of this hypothesis, we have been unable to rescue the *MTHFD2* defect with exogenous formate. Moreover, both normal and transformed mouse embryonic fibroblasts lacking *MTHFD2* are viable but auxotrophic for glycine³¹, suggesting that the strict dependence on *MTHFD2* may be unique to human cells. A recent report indicates that RNAi targeting *MTHFD2* also affects cell migration and invasion of human cancer cells³².

The cell death observed on *MTHFD2* depletion may therefore be due to some type of toxic event that does not arise in wild-type cells and is unrelated to normal cellular metabolism. One hypothesis is that, since *MTHFD2* is considered required for recycling the THF cofactor in mitochondria (Fig. 1d), loss of the protein could result in mitochondrial folates becoming 'trapped' in the CH₂-THF form, which would be detrimental to all folate-coupled activities in the mitochondria. In addition, because CH₂-THF cannot be transported out of the mitochondria²⁷, folate trapping within the mitochondria with *MTHFD2* silencing could potentially impact a number of folate-dependent cytosolic reactions, as well. A similar 'folate trap' phenomenon has been observed on loss of cytosolic folate-coupled enzymes, including loss of methionine synthase³³. Alternatively, loss of *MTHFD2* could result in accumulation of a hitherto unknown toxic metabolite that rapidly causes cell death, perhaps from unwanted side reactions of metabolic enzymes occurring in the absence of *MTHFD2* (ref. 34). While we have consistently observed cell death in response to several independent RNAi sequences targeting *MTHFD2*, it must be emphasized that we have hitherto not been able to rescue cell death by overexpression of an RNAi-resistant *MTHFD2* cDNA, which is required to conclusively demonstrate that loss of this protein causes cancer cell death.

While additional experiments are needed to understand the role of *MTHFD2* function in embryonic cells and in cancer, the specific expression of *MTHFD2* in transformed cells relative to other adult tissues and the growth and survival defects caused by loss of the mRNA suggests that this enzyme should be evaluated as a potential target for cancer chemotherapeutics. More generally, the current analysis, combined with our previous report²⁶ and other recent studies^{6,19} points to a remodelling of mitochondrial one-carbon metabolism in cancer. Although it has long been appreciated that one-carbon folate metabolism is 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 GCOD¹⁰ 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-U133Av2, 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 described¹⁰. 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 procedure³⁵.

Meta-analysis of the FDR was done by counting the number of times a gene was overexpressed in tumours and subtracting the number of times it was underexpressed. Enrichment analysis was done using the GSEA-p method³⁶ with the parameter p set to 1.

The compendium of 80 human tissues was generated from duplicate samples for each tissue using Affymetrix U133A arrays³⁷. The gcRNA-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, GSE6998 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 a control sequence (shCtrl) not matching any human gene. Accession numbers are: shCtrl, Clone ID TRCN0000072232, Clone name lacZ_27s1c1, Target sequence 5'-CGTCGTATTACAACGTCGTGA-3'; shD2-50, Clone ID TRCN0000036550, Clone name NM_006636.2-548s1c1, Target sequence 5'-CGAA TGTGTTGGATCAGTAT-3'; shD2-53, Clone ID TRCN0000036553, Clone name NM_006636.2-772s1c1, Target sequence 5'-GCAGTTGAAGAAACATACAAT-3'. The pLKO.1 plasmids were each packaged with VSVG and pCMV-dR8.91 vectors in 293T cells to generate lentivirus, as previously described²⁶. One day before transfection, 850,000 293T cells were seeded per well/plasmid. Roche X-tremeGENE 9 was used to transfect 293T cells according to the manufacturer's protocols (100 ng VSVG, 900 ng pCMV-dR8.91 and 1 µg plasmid with 6 µl transfection reagent for each well/plasmid). After 18 h of transfection, high serum (30% Gibco 16,000 FBS in DMEM) media was added. Virus was harvested at 24 and 48 h, pooled and frozen at -80 °C in single-use aliquots²⁶.

Lentiviral infection was performed as previously described²⁶. Cells were seeded at a density of 1,500–2,500 cells per well, depending on cell type, in a 96-well plate in RPMI-1640 medium (200 µl total) supplemented with 5% FBS and 2 mM glutamine. Fifteen hours after plating, the cell culture medium was exchanged with fresh medium containing polybrene (8 µg ml⁻¹) and viral supernatant. Plates were centrifuged for 30 min at 2,000 g and 37 °C, after which the medium was replaced. Twenty-four hours later, cells were selected for infection by adding puromycin to the cell culture medium (2 µg ml⁻¹). Uninfected control cells demonstrated 100% cell death within 48 h. Seven days following plating, cells were washed with phosphate-buffered saline and fixed using 4% paraformaldehyde for 20 min at room temperature. Cells were stained using the Hoechst 33,342 dye (Invitrogen) according to manufacturer's specifications, imaged using an ImageXpress Micro (Molecular Devices), and counted using the 'count nuclei' module of MetaXpress (Molecular Devices). Cell counts are expressed normalized to shCtrl-infected cells for each cell line. All experiments were performed using at least four independent cell cultures.

For *MTHFD2* mRNA and protein measurements following shRNA knockdown, HCT-116 cells were infected with lentiviral shRNA vectors in six-well dishes as described²⁶. Seven days following infection, mRNA was isolated from cells using an RNeasy kit (Qiagen), and quantitative reverse transcription-PCR was performed for *MTHFD2* and *HPRT* using the Taqman assays (Applied Biosystems, assay ID Hs01074341_g1 and Hs99999909_m1, respectively), according to the manufacturer's instructions. Quantitative reverse transcription-PCR data are expressed as $\Delta\Delta C_T$ values. For western blot analysis, whole-cell lysates were prepared and resolved by SDS-PAGE, followed by transfer to a PVDF membrane using the Trans-Blot Turbo Transfer System. Membranes were stained using primary antibody for ATP5A (Abcam mouse monoclonal, no. AB147480, 1:100,000) or MTHFD2 (Abcam mouse monoclonal, no. AB56772, 1:500), followed by secondary antibody (polyclonal sheep anti-IgG mouse antibody, no. NA931V, 1:10,000). The signal was detected using enhanced chemiluminescence. Full images of western blots are provided in Supplementary Fig. 3a. For assay of MTHFD2 activity, 1×10^6 cells were lysed in 24% glycerol, 1% Triton-X diluted in phosphate-buffered saline (pH 7.4) and MTHFD2 activity measured in lysate, as previously described²⁸.

Transient siRNA knockdown. siRNA oligonucleotides targeting *MTHFD2* were from Dharmacon, catalogue no. J-009349-10, J-009349-11, J-009349-12, D-009349-02, D-0093848-03 and D-009349-04, referred to in this study as siD2-10, siD2-11, siD2-12, siD2-2, siD2-3 and siD2-3, respectively. Control siRNAs were Dharmacon ON-TARGET, referred to as siCtrl-1, and Dharmacon siGenome, referred to as siCtrl-2. Cells were transfected 24 h after seeding (at 40–60% confluence) with 10 nM oligonucleotide, using the INTERFERin transfection reagent (Polyplus

Transfection). Western blot analysis was performed on cells seeded at 150,000 cells per well in a six-well plate and lysed in $2 \times$ Laemmli sample loading buffer 48 h after siRNA transfection. Lysates were separated by 10% Mini-PROTEAN TGX Precast Gel (Bio-Rad), transferred to a 0.45-µm nitrocellulose membrane (Bio-Rad) and blocked in 5% milk in tris-buffered saline and Tween-20. The following antibodies were used for immunoblotting: monoclonal mouse β -actin antibody (Sigma-Aldrich, 1:30,000), monoclonal mouse MTHFD2 (Abnova, no. H00010797-M01, 1:500) and anti-mouse IgG (H + L) HRP conjugated (Pierce, 1:100,000). Full images of western blots are provided in Supplementary Fig. 3b.

For assessment of cellular proliferation following transient siRNA transfection, cells were seeded at 4,000 cells per well in a 96-well plate 24 h before siRNA transfections and lysed at the indicated time points using the CyQuant Cell Proliferation Assay (Invitrogen) according to manufacturer's instructions. Binding of CyQuant to DNA was analysed by measuring fluorescence in a VICTOR2 plate reader (PerkinElmer).

Flow cytometric measurement of cell death. Cells were seeded at 10^6 cells per 100 mm dish 24 h before siRNA transfections. Cells were collected by trypsinization at the indicated time points, pelleted for 5 min at 1,500g, rinsed twice in ice-cold phosphate-buffered saline and then resuspended in 0.01 M HEPES/NaOH (pH 7.4), 0.14 M NaCl and 2.5 mM CaCl₂. The cell suspension was diluted to a concentration of 5×10^5 cells ml⁻¹, 2.5×10^5 cells were transferred to a 5-ml culture tube and PI(BD Pharmingen) was added at a concentration of $0.5 \mu\text{g ml}^{-1}$; cells were gently vortexed and incubated for 15 min at 25 °C in the dark. Samples were analysed within 1 h in a CyAn ADP Analyzer (Beckman Coulter), with a flow rate ~ 300 events s⁻¹ and a total of 25,000 events were gated. Debris and doublets were excluded using forward scatter and pulse width before gating of PI-positive populations.

Immunohistochemistry. Specimens containing normal and/or cancer tissues were collected in accordance with approval from the local ethics committee (ref. no. Ups 02-577) and according to Swedish rules and legislation. Tissue microarrays holding samples from the different cancer types as well as from normal tissues were produced, as previously described³⁸. Briefly, tissue samples were fixed with formalin and embedded in paraffin; 2–3 tissue cylinders from representative areas were punched out and assembled in a recipient tissue microarray block with prepunched holes. For immunohistochemical staining 4 µm sections were then cut and placed on Superfrost Plus microscope slides (Thermo Fisher Scientific, Fremont, CA, USA). Slides were deparaffinized in xylene, hydrated in graded alcohols and blocked for endogenous peroxidase for 5 min in 0.3% H₂O₂ diluted in 95% ethanol. Heat-induced epitope retrieval was done in a decloaking chamber (Biocare Medical) with citrate buffer at pH 6.0 for 4 min at 125 °C. Before staining, slides were immersed in wash buffer containing 0.2% Tween-20 for 15 min to avoid surface tension. Staining was performed in an Autostainer 480S instrument (Thermo Fisher Scientific) at room temperature with the following steps: Ultra V block (Thermo Fisher Scientific, no. TA-125-UB) 5 min, primary antibody 30 min, primary antibody enhancer (Thermo Fisher Scientific, no. TL-125-PB) 20 min, UltraVision LP HRP polymer (Thermo Fisher Scientific, no. TL-125-PH) 30 min and diaminobenzidine (Thermo Fisher Scientific, no. TA-125-HDX) 5 min two times. MTHFD2 primary antibody was from Abnova (mouse monoclonal, no. H00010797-M01). Between incubations, slides were rinsed in wash buffer. Slides were counter-stained with Mayer's hematoxylin (Histolab, Gothenburg, Sweden, ref. 01820), dehydrated and mounted on cover slips using pterex (Histolab, ref. 00871.0500). Two independent observers evaluated staining for all samples, blinded to the identity of the antibodies used. Staining intensity was classified (negative, weak, moderate and strong) and fraction of positive cells (0, <25%, 25–75%, >75%) were estimated separately for tumour cells and tumour stroma.

Survival analysis. Six independent large cohorts of cancer patients for which survival data for at least a decade were available were examined. Microarray data from Chin *et al.*³⁹ and van de Vijver *et al.*⁴⁰, were downloaded from the Lawrence Berkeley National Laboratory (http://cancer.lbl.gov/breastcancer/list_data.php?id=9) and Rosetta Inpharmatics (<http://www.rii.com/publications/2002/nejm.html>), respectively. Microarray data from the Desmedt *et al.*⁴¹, Pawitan *et al.*⁴², Miller *et al.*⁴³ and Kao *et al.*⁴⁴ studies are available from the NCBI Gene Expression Omnibus, accessions GSE7390, GSE1456, GSE3494 and GSE20685, respectively. Survival data and clinical parameters were obtained from the original reports. Patients were split into 'positive' or 'negative' based on above- or below-median expression of the *MTHFD2* mRNA, and Kaplan–Meier curves were derived for these groups. Hazard ratios were estimated using Cox's proportional hazard model⁴⁵. Groups were tested for significant differences using the logrank test, and meta-analysis was performed using the method by DerSimonian and Laird⁴⁶.

References

- Warburg, O. Über den Stoffwechsel der Carcinomzelle. *Naturwissenschaften* **12**, 1131–1137 (1924).
- Ben-Haim, S. & Ell, P. 18F-FDG PET and PET/CT in the evaluation of cancer treatment response. *J. Nucl. Med.* **50**, 88–99 (2009).

3. Tennant, D. A., Durán, R. V. & Gottlieb, E. Targeting metabolic transformation for cancer therapy. *Nat. Rev. Cancer* **10**, 267–277 (2010).
4. DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G. & Thompson, C. B. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab.* **7**, 11–20 (2008).
5. Yan, H. *et al.* IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* **360**, 765–773 (2009).
6. Zhang, W. C. *et al.* Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell* **148**, 259–272 (2012).
7. Weber, G. Biochemical strategy of cancer cells and the design of chemotherapy: G.H.A. Clowes Memorial Lecture. *Cancer Res.* **43**, 3466–3492 (1983).
8. Weinhouse, S. Glycolysis, respiration, and anomalous gene expression in experimental hepatomas: G.H.A. Clowes Memorial Lecture. *Cancer Res.* **32**, 2007–2016 (1972).
9. Hu, J. *et al.* Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat. Biotechnol.* **31**, 522–529 (2013).
10. Liu, F., White, J. A., Antonescu, C. & Quackenbush, J. GCOD-genechip oncology database. *BMC bioinformatics* **12**, 46 (2011).
11. Rhodes, D. R. *et al.* Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl Acad. Sci. USA* **101**, 9309–9314 (2004).
12. Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
13. Cho, R. J. *et al.* Transcriptional regulation and function during the human cell cycle. *Nat. Genet.* **27**, 48–54 (2001).
14. Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl Acad. Sci. USA* **104**, 1777–1782 (2007).
15. Altenberg, B. & Greulich, K. O. Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. *Genomics* **84**, 1014–1020 (2004).
16. Hofbauer, K.-H. *et al.* Oxygen tension regulates the expression of a group of procollagen hydroxylases. *Eur. J. Biochem.* **270**, 4515–4522 (2003).
17. Dennis, J. W., Granovsky, M. & Warren, C. E. Glycoprotein glycosylation and cancer progression. *Biochim. Biophys. Acta.* **1473**, 21–34 (1999).
18. Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499–507 (2008).
19. Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**, 346–350 (2011).
20. Noda, K. *et al.* Elevated FX expression and increased production of GDP-L-fucose, a common donor substrate for fucosylation in human hepatocellular carcinoma and hepatoma cell lines. *Cancer Res.* **63**, 6282–6289 (2003).
21. Liu, W. *et al.* Reprogramming of proline and glutamine metabolism contributes to the proliferative and metabolic responses regulated by oncogenic transcription factor c-MYC. *Proc. Natl Acad. Sci. USA* **109**, 8983–8988 (2012).
22. De Ingeniis, J. *et al.* Functional specialization in proline biosynthesis of melanoma. *PLoS One* **7**, e45190 (2012).
23. Hagedorn, C. H. & Phang, J. M. Transfer of reducing equivalents into mitochondria by the interconversions of proline and delta1-pyrroline-5-carboxylate. *Arch. Biochem. Biophys.* **225**, 95–101 (1983).
24. Krishnan, N., Dickman, M. B. & Becker, D. F. Proline modulates the intracellular redox environment and protects mammalian cells against oxidative stress. *Free Radic. Biol. Med.* **44**, 671–681 (2008).
25. Reversade, B. *et al.* Mutations in PYCR1 cause cutis laxa with progeroid features. *Nat. Genet.* **41**, 1016–1021 (2009).
26. Jain, M. *et al.* Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **336**, 1040–1044 (2012).
27. Tibbetts, A. S. & Appling, D. R. Compartmentalization of mammalian folate-mediated one-carbon metabolism. *Annu. Rev. Nutr.* **30**, 57–81 (2010).
28. Mejia, N. R. & MacKenzie, R. E. NAD-dependent methylenetetrahydrofolate dehydrogenase is expressed by immortal cells. *J. Biol. Chem.* **260**, 14616–14620 (1985).
29. Di Pietro, E., Sirois, J., Tremblay, M. & MacKenzie, R. Mitochondrial NAD-dependent methylenetetrahydrofolate dehydrogenase-methylenetetrahydrofolate cyclohydrolase is essential for embryonic development. *Mol. Cell Biol.* **22**, 4158 (2002).
30. Smith, G. K. *et al.* Activity of an NAD-dependent 5,10-methylenetetrahydrofolate dehydrogenase in normal tissue, neoplastic cells, and oncogene-transformed cells. *Arch. Biochem. Biophys.* **283**, 367–371 (1990).
31. Patel, H., Pietro, E. D. & MacKenzie, R. E. Mammalian fibroblasts lacking mitochondrial NAD⁺-dependent methylenetetrahydrofolate dehydrogenase-cyclohydrolase are glycine auxotrophs. *J. Biol. Chem.* **278**, 19436–19441 (2003).
32. Lehtinen, L., Ketola, K., Mäkelä, R. & Mpindi, J. High-throughput RNAi screening for novel modulators of vimentin expression identifies MTHFD2 as a regulator of breast cancer cell migration and invasion. *Oncotarget* **4**, 48–63 (2013).
33. Weir, D. G. & Scott, J. M. The methyl folate trap. *Lancet* **318**, 337–340 (1981).
34. Linster, C. L., Van Schaftingen, E. & Hanson, A. D. Metabolite damage and its repair or pre-emption. *Nat. Chem. Biol.* **9**, 72–80 (2013).
35. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 289–300 (1995).
36. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
37. Su, A. I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA* **101**, 6062–6067 (2004).
38. Kampf, C., Olsson, L., Ryberg, U., Sjöstedt, E. & Pontén, F. Production of tissue microarrays, immunohistochemistry staining and digitalization within the human protein atlas. *J. Vis. Exp.* **63**, e3620 (2012).
39. Chin, K. *et al.* Genomic and transcriptional aberrations linked to breast cancer pathophysiology. *Cancer cell* **10**, 529–541 (2006).
40. Van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
41. Desmedt, C. *et al.* Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin. Cancer Res.* **13**, 3207–3214 (2007).
42. Pawitan, Y. *et al.* Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res.* **7**, R953–R964 (2005).
43. Miller, L. D. *et al.* An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc. Natl Acad. Sci. USA* **102**, 13550–13555 (2005).
44. Kao, K.-J., Chang, K.-M., Hsu, H.-C. & Huang, A. T. Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization. *BMC Cancer* **11**, 143 (2011).
45. Cox, D. R. Regression models and life-tables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **34**, 187–220 (1972).
46. DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).

Acknowledgements

We acknowledge the National Cancer Institute for cellular reagents. We acknowledge Joseph White (Dana-Farber Cancer Institute) for assistance with the GCOD database. R.N. and N.G.S. were supported by grants from the Swedish Research Council (K2012-99X-22007-01-3) and Foundation for Strategic Research (ICA10-0023). M.J. was supported by NIH K08HL107451. This work was funded by grants from the NIH (R01DK081457) V.K.M. is an Investigator of the Howard Hughes Medical Institute.

Author contributions

R.N. and M.J. conceived and performed all experiments, collected and analysed experimental data and prepared the manuscript. N.M., L.S. and J.H. performed shRNA-mediated knockdown experiments, cell proliferation and biochemical assays. N.G.S. performed siRNA-mediated knockdown experiments, cell proliferation and cell death assays. C.K. and A.A. performed immunohistochemistry experiments. V.K.M. conceived all experiments, reviewed all experimental data and prepared the manuscript. All authors have discussed the results and reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The Massachusetts General Hospital has filed a provisional patent application listing R.N., M.J. and V.K.M. entitled 'Glycine, mitochondrial one-carbon metabolism, and cancer', patent number 61/791,082. All remaining authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Nilsson, R. *et al.* Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nat. Commun.* **5**:3128 doi: 10.1038/ncomms4128 (2014).