Dihydropyrimidine Accumulation Is Required for the Epithelial-Mesenchymal Transition

Yoav D. Shaul,[1,2] Elizaveta Freinkman,[1] William C. Comb,[1] Jason R. Cantor,[1] Wai Leong Tam,[1,9] Prathapan Thiru,[1] Dohoon Kim,[1] Naama Kanarek,[1] Michael E. Pacold,[1,7] Walter W. Chen,[1] Brian Bierie,[1] Richard Possemato,[1] Ferenc Reinhardt,[1] Robert A. Weinberg,[1,4,6] Michael B. Yaffe,[2,4,5] and David M. Sabatini[1,2,3,4,8,*]

[1]Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142, USA
[2]Koch Institute for Integrative Cancer Research at MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[3]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[6]MIT Ludwig Center for Molecular Oncology, Cambridge, MA 02139, USA
[7]Department of Radiation Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA
[8]Broad Institute, Cambridge, MA 02142, USA
[9]Genome Institute of Singapore, Singapore 138672, Singapore
*Correspondence: sabatini@wi.mit.edu
http://dx.doi.org/10.1016/j.cell.2014.07.032

SUMMARY

It is increasingly appreciated that oncogenic transformation alters cellular metabolism to facilitate cell proliferation, but less is known about the metabolic changes that promote cancer cell aggressiveness. Here, we analyzed metabolic gene expression in cancer cell lines and found that a set of high-grade carcinoma lines expressing mesenchymal markers share a unique 44 gene signature, designated the "mesenchymal metabolic signature" (MMS). A FACS-based shRNA screen identified several MMS genes as essential for the epithelial-mesenchymal transition (EMT), but not for cell proliferation. Dihydropyrimidine dehydrogenase (DPYD), a pyrimidine-degrading enzyme, was highly expressed upon EMT induction and was necessary for cells to acquire mesenchymal characteristics in vitro and for tumorigenic cells to extravasate into the mouse lung. This role of DPYD was mediated through its catalytic activity and enzymatic products, the dihydropyrimidines. Thus, we identify metabolic processes essential for the EMT, a program associated with the acquisition of metastatic and aggressive cancer cell traits.

INTRODUCTION

Alterations in cellular metabolism are now recognized as an emerging hallmark of cancer (Hanahan and Weinberg, 2011). Almost a century ago, Otto Warburg observed that, under aerobic conditions, tumor cells display increased glucose uptake and glycolytic rates compared to resting cells (reviewed in Hsu and Sabatini, 2008; Ward and Thompson, 2012). Subsequently, many studies have revealed how this and other metabolic changes allow cancer cells to accumulate building blocks for the biosynthesis of macromolecules, while simultaneously maintaining energetic and redox balance (reviewed in Cantor and Sabatini, 2012). Whereas many of these mechanisms are shared with normal rapidly proliferating cells, in recent years cancer genomic data have revealed metabolic alterations that appear to occur only in specific tumor types. These changes include the loss of succinate dehydrogenase (SDH) or fumarate hydratase (FH) in certain renal cell carcinomas and other familial cancer syndromes (reviewed in Gottlieb and Tomlinson, 2005), mutation of isocitrate dehydrogenase (IDH) 1 or 2 in glioma, acute myeloid leukemias, chondrosarcomas (Dang et al., 2009; Schulze and Harris, 2012), and amplification of phosphoglycerate dehydrogenase (PHGDH) in estrogen receptor (ER)-negative breast cancer and melanoma (Locasale et al., 2011; Possemato et al., 2011). These examples suggest that, in addition to fueling increased proliferation, cancer-associated alterations in metabolism can also satisfy tumor-specific demands.

Relatively few studies have examined the metabolic underpinnings of the cellular programs that increase cancer cell aggressiveness (Nomura et al., 2010; Ulanovskaya et al., 2013; Zhang et al., 2012). One such program is the epithelial-mesenchymal transition (EMT) (reviewed in Nieto and Cano, 2012) that operates in carcinoma cells and is thought to confer stem-like properties, such as enhanced survival, self-renewal, and anchorage-independent growth, all of which contribute to increased aggressiveness in vivo (Scheel and Weinberg, 2011). Indeed, EMT markers are predictive for increased invasion, loss of differentiated characteristics, metastasis, and poor prognosis in a number of human tumor types (Nieto and Cano, 2012).

To understand how cellular metabolism contributes to these and other proliferation-independent features of cancer, we created a framework for the systematic identification of metabolic alterations specific to particular tumor types, as well as

those that may characterize high-grade malignancies. By analyzing metabolic gene expression patterns in a large number of cancer cell lines, we identified a metabolic gene signature that is present in high-grade tumors bearing mesenchymal markers. Among the enzymes encoded by these genes is dihydropyrimidine dehydrogenase (DPYD), which catalyzes the rate-limiting step in pyrimidine degradation and whose physiological role in cancer was previously unknown. We find that EMT-promoting transcription factors induce the expression of DPYD and that its products, the dihydropyrimidines, must accumulate for cells to undergo an EMT. These findings reveal that the EMT induces a particular metabolic state and suggest that DPYD may have value as a diagnostic marker or therapeutic target in high-grade carcinomas.

RESULTS

A Mesenchymal-like Metabolic Gene Expression Signature in High-Grade Carcinoma Cells

In order to study metabolic gene expression patterns in cancer, we used publicly available data to generate a database of mRNA expression profiles for 1,704 metabolic genes in 978 human cancer cell lines (see the Experimental Procedures) (Possemato et al., 2011). Aided by unsupervised hierarchical clustering, we organized the profiles into five distinct groups (Figure 1A; Table S1 available online); for four of these groups, the basis for clustering was readily apparent (Figure 1B). One group consisted of melanoma cell lines, which uniquely express skin pigment biosynthesis genes. The cell lines in a second group were derived from hematopoietic system cancers (e.g., leukemia, lymphoma, and multiple myeloma) and in a third, from neuroendocrine or neuroectodermal cancers (e.g., small cell lung cancer, medulloblastoma, and neuroblastoma; Onganer et al., 2005; Parham, 2001). A fourth group consisted mostly of epithelial cancer cell lines, in which cell lines originating from breast, liver, colon, kidney, etc., clustered together. These results indicate that patterns of metabolic gene expression are sufficient to organize most cancer cell lines by tissue of origin, suggesting that many cancers retain significant portions of the metabolic programs of their normal tissue counterparts.

The cell lines in the fifth group proved more difficult to classify and thus were initially named the "mixed-lineage group" (Figure S1A). While this group contained almost all the cell lines derived from mesenchymal tumors (soft-tissue sarcoma, osteosarcoma; 20% of the cell lines in this group) and glioblastomas, it also included a large number of carcinoma lines (e.g., non-small-cell lung, hepatocellular, and breast; 43% of the cell lines in this group). Notably, all the breast cancer lines in the mixed-lineage group were of the Basal B subtype, which are derived from high-grade carcinomas (Carey et al., 2010) (Figure 1C). Likewise, all the hepatocellular carcinoma (HCC) cell lines in this group were also derived from high-grade tumors (Park et al., 1995) and retained fewer of the metabolic gene expression features of normal liver than did the HCC lines that were in the epithelial group (Figure S1B). Such loss-of-epithelial and gain-of-mesenchymal characteristics have been associated with high-grade malignancy in a variety of carcinoma types (Brabletz, 2012). Moreover, several of the glioblastoma and the majority of Basal B breast cancer cell lines are known to bear mesenchymal characteristics (Kao et al., 2009; Verhaak et al., 2010). Thus, we thought it likely that the cell lines in the mixed-lineage group shared a common mesenchymal-like phenotype.

Indeed, gene set enrichment analysis (GSEA) (Subramanian et al., 2005) of the entire genome (~17,000 genes) showed that expression of the mesenchymal gene set (EMT_UP) was significantly elevated in the mixed-lineage group relative to the other groups (FDR q-value < 0.0001; Figure S1C). Furthermore, the mixed-lineage group had elevated expression of key mesenchymal markers (Mani et al., 2008), including vimentin (VIM), Snail family zinc finger 1 and 2 (SNAI1/2), N-cadherin (CDH2), Twist basic helix-loop-helix transcription factor 1 (TWIST1), and the zinc-finger E-box binding homeobox 1 (ZEB1) transcription factor (Figure 1D). Lastly, the epithelial markers claudin 1 (CLDN1) and E-cadherin (CDH1) were expressed at low levels in this group (Figure 1D). Collectively, these data suggest that the cell lines in the mixed-lineage group, regardless of tissue of origin, display a mesenchymal-like gene expression profile. Accordingly, hereafter, we refer to the mixed-lineage group as the mesenchymal group of cell lines.

Identification of a Mesenchymal Metabolic Gene Expression Signature

We identified a mesenchymal metabolic signature (MMS), composed of 44 metabolic genes associated with diverse metabolic pathways, as highly and differentially expressed in the mesenchymal group of cell lines relative to the other groups (see the Experimental Procedures) (Table 1; Figure 2A). The MMS is particularly enriched for glycan biosynthesis genes (36% of the genes in the set), including glutamine-fructose-6-phosphate aminotransferase 2 (GFPT2) and UDP-N-acetylglucosamine pyrophosphorylase 1 (UAP1), which encode the rate-limiting and endpoint enzymes of the hexosamine biosynthetic pathway (HBP), respectively (Elbein et al., 2004; Zhang et al., 2004). The HBP end product, UDP N-acetylglucosamine (UDP-GlcNAc), is used by the enzyme O-GlcNac transferase (OGT) as a donor substrate to modify proteins via covalent attachment of GlcNAc to serine and/or threonine residues (Ma and Vosseller, 2013). Of special interest, this modification plays an important role in mesenchymal cells by stabilizing the EMT-inducing transcription factor SNAI1, which in turn downregulates the key epithelial marker CDH1 (Park et al., 2010). The MMS list includes other genes with known connections to cancer aggressiveness, such as ecto-5′-nucleotidase (NT5E, also known as CD73), a mesenchymal stem cell marker (Stagg et al., 2010), ectonucleotide pyrophosphatase/phosphodiesterase 2 (ENPP2, autotaxin), which promotes cell migration and metastasis (Benesch et al., 2014), and monoacylglycerol lipase (MGLL), which enhances cancer cell aggressiveness through the production of signaling lipids (Nomura et al., 2010). These examples suggest that the remaining MMS genes may also play an important role in the mesenchymal phenotype and/or aggressiveness of certain cancer cells.

We found that the MMS genes were significantly upregulated in cell lines that express known mesenchymal markers (Figure 2B, left). For example, this gene set is upregulated in cell
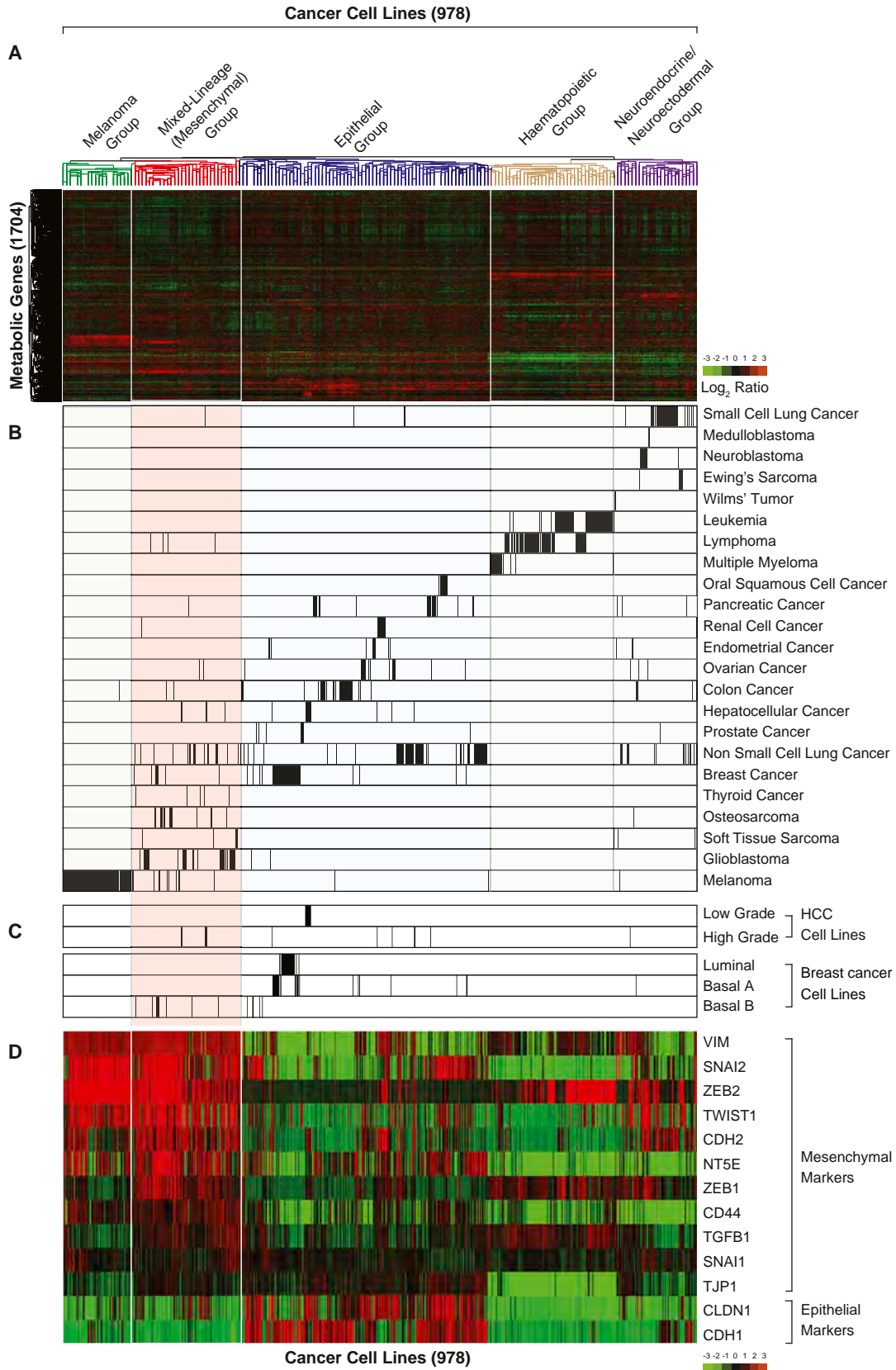
(legend on next page)

Table 1. The Mesenchymal Metabolic Signature Genes, Classified by Metabolic Pathway

	Metabolic Pathways	Metabolic Enzymes
Nucleotide	pyrimidine degradation	DPYD[a,b]
	Other	AK5,[a] ENPP1,[a] NT5E[a]
Lipid	triacylglycerol degradation	MGLL[a]
	cholesterol	TM7SF2, AKR1B1[a]
	sphingosine	SPHK1,[a] UGCG[a]
	signaling	ENPP2,[a] PPAP2B,[a] PPAPDC1A,[a] PDE1C,[a] PLCB4,[a] PTGR1,[a] PIK3C2B, PLCG2, ALDH1A1, PIP5K1B
Amino Acid	branched amino acid degradation	BCAT1[a]
	amino acid degradation	CYP1B1[a]
	tetrahydrobiopterin biosynthesis	GCH1
Carbon	TCA cycle	CYBRD1,[a] COX7A1,[a] CYBA
Redox	glutathione	MGST1,[a] GPX8,[a] GPX2
Glycan	beta-galactose	ST6GAL1, GLB1L2
	dermatan sulfate	DSE,[a] DSEL[a]
	heparin sulfate	HS3ST3A1,[a] EXT1[a]
	sulfatases	ARSJ,[a] SULF1,[a] PAPSS2[a]
	GlcNAc	GFPT2,[a,b] GALNT10,[a] UAP1[a]
	glycan	GXYLT2,[a] GBE1,[a] GLT8D2,[a] GALNT3
	hyaluronan	HAS2[a]
	glycoprotein	PAM,[a] CHI3L1[a]
	other	B3GNT9,[a] MFNG, HPDL
Cofactor	nicotinamide	AOX1,[a] NNMT,[a] QPRT
Other		MICAL2,[a] MME,[a] DDAH1,[a] MSRB3,[a] CA12,[a] PTER, CA2

TCA, tricarboxylic acid; GlcNAc, N-acetylglucosamine.
[a]These genes are upregulated in mesenchymal cells.
[b]These genes are rate-limiting enzymes.

lines derived from Basal B breast cancer and high-grade HCC relative to their luminal and low-grade counterparts, respectively (Figure 2C). Quantitative reverse-transcriptase PCR (qRT-PCR) and immunoblotting confirmed the overexpression of several individual MMS genes, including nicotinamide N-methyltransferase (*NNMT*) and *DPYD*, in high-grade breast cancer and HCC

cell lines (Figures 2D and 2E), which also expressed mesenchymal markers, such as ZEB1 and TWIST1, and low levels of CDH1 (Figure S2B).

Next, we asked if MMS gene expression correlates with that of known mesenchymal markers in primary human tumors and in cancer cell lines. From a database of expression profiles for 1,460 human primary tumors, including many of mesenchymal origin, we identified tumors with high expression of known mesenchymal markers (see the Experimental Procedures) (Figure S2A). In such tumors, the MMS genes were significantly more highly expressed than in tumors not expressing these markers (Figures 2B and S2A). Thus, MMS gene expression correlates with that of known mesenchymal markers in both cancer cell lines and tumors, suggesting that a particular metabolic program characterizes the mesenchymal cell state.

EMT-Dependent Induction of Mesenchymal Metabolic Signature Genes

Given the high expression of MMS genes in mesenchymal-like relative to epithelial cancer cell lines, we hypothesized that the EMT program may directly affect the expression of these genes. To investigate this possibility, we examined engineered human mammary epithelial (HMLE) cells that undergo an EMT upon the activation of Twist (HMLE-Twist-ER) following treatment with hydroxytamoxifen (OHT) (Mani et al., 2008). Over a 15-day treatment with OHT, the HMLE-Twist-ER cells shifted their cell-surface markers from an epithelial (CD24[high], CD44[low]) to a mesenchymal (CD24[low], CD44[high]) profile (Al-Hajj et al., 2003; Figure S2C), induced *ZEB1* and *TWIST1* expression, and suppressed *CDH1* (Figures 2F and 2G). Like the mesenchymal markers, MMS genes, such as DPYD and NNMT, also displayed a progressive increase in mRNA and protein levels over the course of OHT treatment (Figures 2F and 2G). Moreover, NAMEC cells, an HMLE-derived cell line that spontaneously acquired the mesenchymal state (Tam et al., 2013) (Figure S2D), also expressed high levels of several MMS genes (Figure 2F). Lastly, reanalysis of a previous expression profiling study comparing HMLE cells expressing an empty vector or Twist (Taube et al., 2010) showed that, unlike the majority of metabolic genes, MMS genes were upregulated upon EMT induction in culture (Figure 2H). Collectively, these results suggest that the EMT program and MMS gene induction are coupled processes.

A Fluorescence-Activated Cell Sorting-Based Pooled Small Hairpin RNA Screen for MMS Genes Required for the EMT

To identify which, if any, of the MMS genes play a critical role in the EMT, we developed a fluorescence-activated cell sorting (FACS)-based RNAi screen using a pool of 514 lentivirus

Figure 1. Based on Metabolic Gene Expression Patterns, High-Grade Carcinoma Cell Lines Cocluster with Mesenchymal Cells
(A) Metabolic gene expression patterns are sufficient to cluster most, but not all, cancer cell lines based on their tissue of origin. Two-way hierarchical clustering of the expression levels of 1,704 metabolic genes in 978 different cell lines is presented as a heatmap.
(B) Cell lines derived from related cancer types cocluster based on metabolic gene expression patterns. Each row shows all the cell lines in the data set derived from the indicated cancer type. Within each row, each black line represents the position of a cell line in the cluster.
(C) Most high-grade hepatocellular carcinoma (HCC) and Basal B breast cancer cell lines cluster within the mesenchymal group.
(D) Known mesenchymal markers are highly expressed in the mesenchymal group. Cancer cell lines were ordered as in (A).
See also Figure S1 and Table S1.
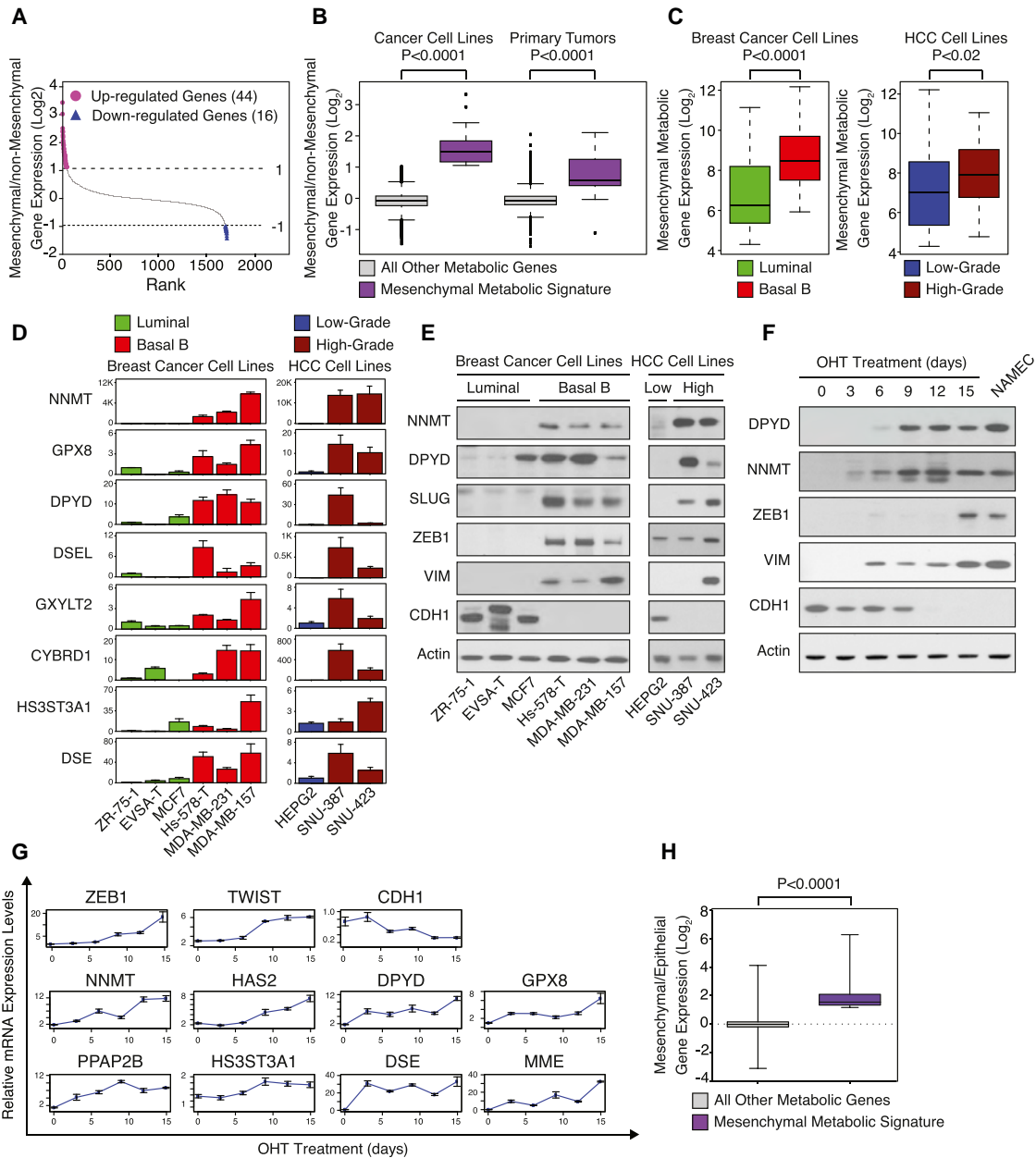
Figure 2. High Expression of MMS Genes in Mesenchymal Cell Lines

(A) Identification of the MMS. For each metabolic gene, the ratio between the mean expression level in the mesenchymal group of cell lines and in all other groups (see Figure 1) was determined and used to rank the genes. Genes that are upregulated (purple, 44 genes) or downregulated (blue, 16 genes) by at least 2-fold in mesenchymal relative to nonmesenchymal cells are highlighted.

(B) Elevated MMS gene expression in mesenchymal cancer cell lines and primary tumors. Cancer cell lines and primary tumors were divided into mesenchymal and nonmesenchymal groups based on the expression of known mesenchymal markers (Figures 1D and S2A). For each metabolic gene, the ratio of the mean expression level among the groups was determined. The box plots represent the $\log_2$ ratio distribution of MMS genes (purple) and all other metabolic genes (gray).

(C) MMS gene expression is elevated in Basal B breast and high-grade HCC cancer cell lines. Box plots represent the expression levels of the MMS genes in the indicated breast cancer (green, luminal; red, Basal B) and HCC (blue, low-grade; brown, high-grade) subtypes.

(D) Individual validation of MMS mRNA levels in breast cancer (green, luminal; red, Basal B) and HCC (blue, low-grade; brown, high-grade) cell lines by qRT-PCR. Each value represents the mean ± SEM for n = 3.

(E) Individual validation of MMS protein levels in the indicated breast cancer and HCC cell lines by immunoblotting.

(F) MMS protein upregulation an HMLE-Twist-ER-inducible EMT system. HMLE-Twist-ER cells were treated with hydroxytamoxifen (OHT) to induce an EMT for 15 days. Every 3 days, cellular proteins were isolated and subjected to immunoblotting using the indicated antibodies. NAMEC cells are mesenchymal cells derived from HMLE cells (see the Experimental Procedures).

(legend continued on next page)

vector-expressed small hairpin RNAs (shRNAs) targeting 42 of the MMS genes, 12 known mesenchymal genes, as well as non-expressed control genes (*GFP*, *RFP*, *Luciferase*, and *LacZ*) and 16 non-MMS metabolic genes (Figure S3A). We then induced the EMT in HMLE-Twist-ER cells expressing the shRNA hairpin library, and after 15 days, we compared the abundance of each hairpin in FACS-sorted epithelial versus mesenchymal cell populations isolated using the CD44 and CD24-surface antigens (Figure 3A). We reasoned that knockdown (KD) of an EMT-essential gene would cause cells to remain in the epithelial state (CD24^{high}/CD44^{low}) even upon OHT treatment. Indeed, hairpins targeting the EMT-promoting transcription factors *ZEB1* and *SNAI1* were enriched in the epithelial population (Figure 3B). We also found that hairpins against 16 MMS genes were similarly enriched, suggesting that knockdown of these genes blocks activation of the EMT program (Figure 3C). Among the MMS genes, *DPYD* was a top hit, with 5 out of 12 hairpins scoring in the screen (Figure 3B). DPYD is the rate-limiting enzyme of the pyrimidine degradation pathway (Amstutz et al., 2011) and is also capable of degrading the chemotherapeutic agent 5-fluorouracil (5-FU), but the physiological role of this enzyme in cancer cells is unclear (Offer et al., 2013).

We wished to rule out the possibility that knockdown of *DPYD* and the other MMS hit genes may block the EMT by affecting the proliferation or viability of epithelial cells. Thus, in a parallel experiment, we determined the abundance of each hairpin in HMLE-Twist-ER cells before and after a 15-day period of proliferation in the absence of EMT induction (Figure 3A; uninduced day 0 and day 15). As expected, the control hairpins as a group had a neutral effect on proliferation (median log2 hairpin abundance ratio = −0.28). Importantly, the abundance distributions of the *ZEB1*, *SNAI1*, and *DPYD* hairpins did not differ significantly from the control group (Figure 3D), indicating that these hairpins did not affect cellular viability or proliferation; by contrast, hairpins targeting ribonucleotide reductase M1 (*RRM1*) and thymidylate synthetase (*TYMS*), which are critical for cell division (Tennant et al., 2010), caused a significant antiproliferative effect (median log$_2$ hairpin abundance ratio = −3.23 and −2.4, respectively) (Figure 3D). Therefore, knockdown of DPYD suppressed the EMT program without inhibiting the viability or proliferation of epithelial cells, suggesting that this enzyme plays a specific role in inducing the mesenchymal cell state.

To validate the identification of DPYD as essential for the EMT, we individually infected HMLE-Twist-ER cultures with eight distinct shRNAs targeting DPYD. We found that DPYD knockdown, in a dose-dependent manner, decreased the percentage of cells with a mesenchymal-like profile (CD24^{low}/CD44^{high}) after 15 days of Twist induction by OHT treatment (Figure S3C). DPYD knockdown with the hairpins that most strongly reduced DPYD expression (shDPYD_1 and shDPYD_4) (Figures S3D and S3G)

did not affect the viability of untreated HMLE-Twist-ER cells (Figure S3E) but decreased the percentage of OHT-treated cells with a mesenchymal-like profile (Figure S3F). In order to further demonstrate the reliability of the screen, we validated another hit from the screen, glutathione peroxidase 8 (*GPX8*) (Figure 3C). As for DPYD, GPX8 hairpins demonstrated a similar expression-phenotype relationship (Figure S3H), thus confirming that several MMS genes are essential for the EMT program.

DPYD Expression Promotes the EMT

Next, we further characterized the role of DPYD in the EMT program. In addition to inhibiting the EMT-associated changes in cell-surface marker expression (CD24^{high}/CD44^{low} to CD24^{low}/CD44^{high}) (Figure 4A), DPYD knockdown suppressed the Twist-induced expression of known mesenchymal markers, as judged by qRT-PCR (Figure 4B), immunoblotting (Figure 4C), and immunofluorescence (Figure 4E). Moreover, DPYD knockdown suppressed EMT-induced changes in cell morphology (Figures 4D and 4E).

To rule out the possibility that the effects of the DPYD shRNAs are due to off-target effects, we restored DPYD levels in shDPYD-expressing HMLE-Twist-ER cells by ectopically expressing the mouse isoform of DPYD (mDpyd), which is 86% identical at the amino acid level to the human isoform (Figure S4A) but is unaffected by the shRNAs targeting the human DPYD mRNA (Figure S4B). Importantly, mDpyd rescued the effect of shDPYD on the expression of cell-surface markers (Figure 4A) and known EMT genes (Figures 4B, 4C, and 4E) as well as changes in cell morphology (Figures 4D and 4E). DPYD knockdown had a similar effect on mesenchymal marker expression in PANC1 pancreatic cancer cells engineered to express Twist-ER (Figure S4C), showing that the role of DPYD in EMT induction is not restricted to breast tissue-derived cell lines.

As an in vitro functional readout of the EMT, we measured the capacity of HMLE-Twist-ER cells to form mammospheres, migrate, and invade, which are unique properties of the mesenchymal-like (CD24^{low}/CD44^{high}), but not of epithelial (CD24^{high}/CD44^{low}) cells (Mani et al., 2008). DPYD knockdown decreased the efficiency of mammosphere formation (Figures 4F and S4D), as well as the capacity of the cells to migrate (Figures 4G and S4E) and invade matrigel basement membrane (Figure 4H). Therefore, blockade of EMT induction by DPYD suppression affects multiple in vitro phenotypes of mesenchymal cells.

We then turned our attention to the role of DPYD expression in the EMT program during the progression of human cancer cells through specific aspects of the metastatic cascade in mice. Extravasation is a key step in metastasis, permitting cancer cells to migrate out of capillaries into target tissues (Nieto and Cano, 2012), and recent studies demonstrate that Twist-induced EMT promotes tumor cell extravasation (Tsai et al., 2012). To test

(G) MMS gene upregulation in HMLE-Twist-ER-inducible EMT system. Every 3 days, cells were collected and mRNA isolated and subjected to qRT-PCR using the indicated probes. Each value represents the mean ± SEM for n = 3.
(H) MMS genes are upregulated during the EMT. Microarray analysis for gene expression changes during EMT was described previously (GSE24202; Taube et al., 2010). Here, the same data set was reanalyzed for the log$_2$ expression ratio of MMS and all other metabolic genes between HMLE-Twist-ER cells forced to express Twist and Snai1 (mesenchymal) to HMLE-Twist-ER expressing empty vector (epithelial). The box plots represent the log$_2$ ratio expression distributions of MMS genes (purple) and all other metabolic genes (gray). The p value for the comparison between the two groups was determined using the Student's t test. See also Table S2.
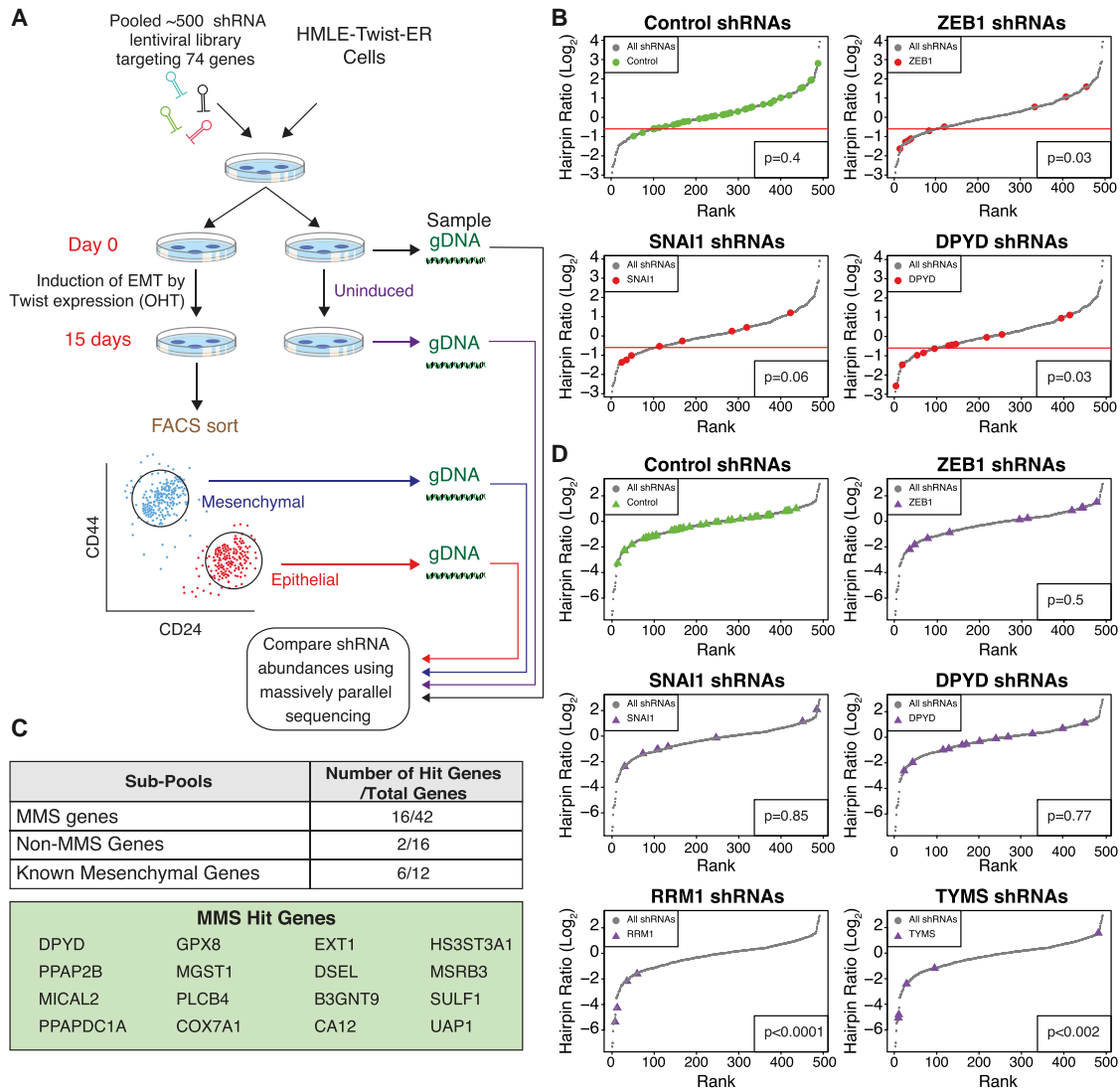
Figure 3. A FACS-Based Pooled shRNA Screen Identifies DPYD as Required for EMT

(A) Schematic presentation of the FACS-based pooled shRNA screen. OHT, hydroxytamoxifen; gDNA, genomic DNA.

(B) DPYD knockdown (KD) inhibits the EMT. All hairpins from the screen were ranked based on the $\log_2$ ratio of their abundance in the mesenchymal relative to the epithelial population of OHT-induced HMLE-Twist-ER cells after FACS sorting (see Figure 3A). Hairpin subpools pictured include those targeting control genes (39 hairpins targeting RFP, GFP, luciferase, and LacZ), ZEB1 (nine hairpins), SNAI1 (eight hairpins), and DPYD (12 hairpins). One SD below the mean of the distribution of the control hairpins was set as a cutoff (red line). Every hairpin with a $\log_2$ ratio below the cutoff was considered a hit. The significance of the differences in distribution between the selected genes and the other genes in the screen was quantified using the Student's t test.

(C) Several of the MMS genes are critical for the EMT. Genes with at least two hairpins scoring below the cutoff (see B) were classified as hit genes. The numbers in the table represent the hit genes as a fraction of the total genes in a given subpool.

(D) DPYD KD does not affect cell viability. All hairpins were ranked based on the $\log_2$ ratio of their abundance in uninduced HMLE-Twist-ER cells on day 15 relative to day 0. The same hairpin subpools as in (B), with the addition of shRNAs targeting the essential genes RRM1 (four hairpins) and TYMS (five hairpins), are shown. The significance of the differences in distribution between the selected genes and the control genes was quantified using Student's t test.

See also Figure S3 and Tables S3 and S4.

whether DPYD plays a role in the ability of cells to extravasate, we used HMLE-Twist-ER transformed with an H-Ras^{G12V} oncogene expression vector that also expressed GFP (HMLER-Twist-ER cells). These cells were then infected with hairpins against DPYD, together with a vector control or mDpyd, induced by OHT to undergo EMT in vitro, and injected into the mouse tail vein. As measured by GFP staining, DPYD depletion reduced the ability of HMLER-Twist-ER cells to enter the mouse lung by ∼95% (Figures 4I and 4J), a process that was significantly rescued by the expression of mDpyd (Figures 4I and 4J). Thus, DPYD plays an essential role during the lung extravasation of cancer cells that are undergoing the EMT.

DPYD Activity Is Essential for Its Function in the EMT

We asked whether the role of DPYD in the EMT requires its enzymatic activity. Accordingly, we tested the ability of a cata-lytically attenuated mouse DPYD mutant (mDpyd-I560S, also known as DPYD*13, which has a 75% reduction in enzymatic activity relative to wild-type [WT]; Offer et al., 2013), to rescue the inhibitory effect of shDPYD_1 on EMT induction. Whereas expression of wild-type mDpyd in the presence of shDPYD_1 restored the EMT induction following 15 days of OHT treat-ment, mDpyd-I560S had a greatly reduced capacity to rescue CD44/CD24 expression and mammosphere formation and completely failed to restore expression of the EMT-inducing transcription factor ZEB1 (Figured 5A–5C). In addition, whereas control cells (expressing empty vector) treated with OHT for only 10 days displayed an intermediate CD44/CD24 marker expression profile, cell lines ectopically expressing either mouse or human DPYD (DPYD-FLAG) displayed higher mesen-chymal marker expression at this earlier time point (Figure 5D), resembling the profile of control cells after a full 15 days of OHT treatment (Figure 5A). Compared to wild-type DPYD, expres-sion of the mutant DPYD-I560S (human DPYD-I560S-FLAG) had a greatly attenuated effect on cell-surface marker expres-sion, mammosphere formation, and ZEB1 expression (Figures 5D–5F). Thus, the pro-EMT role of DPYD requires DPYD residues needed for its full enzymatic activity. Moreover, the accelerated progression of the EMT in DPYD-overexpressing cells suggests that the DPYD products may be rate-limiting in EMT.

Cellular Dihydropyrimidine Levels Are Elevated during the EMT

Having demonstrated that DPYD enzymatic activity plays a crit-ical role in the EMT program, we asked whether its metabolic products increase in abundance during this process. To do so, we used liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Büchel et al., 2013) to determine the cellular con-centration of DPYD substrates (uracil and thymine) and immedi-ate products (dihydrouracil [DHU] and dihydrothymine [DHT]) (Figure 5G) (Lohkamp et al., 2010). In HMLE-Twist-ER cells, overexpression or knockdown of DPYD resulted in a corresponding ∼10-fold increase or decrease, respectively, in the intracellular DHU/uracil molar ratio (Figure 5H). Moreover, NAMEC cells exhibited higher DHU/uracil and DHT/thymine ra-tios than HMLE-Twist-ER cells (by 10- and 6-fold, respectively; Figures 5H and S5A), consistent with the higher endogenous DPYD expression level in the former cells (Figure 2F). In addi-tion, OHT treatment of HMLE-Twist-ER cells, which progres-sively upregulates DPYD expression (Figures 2F and 2G), gradually increased the cellular DHU/uracil molar ratio by 5-fold after 15 days of Twist induction (Figure 5I). DPYD expres-sion and DHU/uracil ratios were also correlated in breast cancer and HCC cell lines (Figures 5J and 5K). Notably, the higher DHU/uracil molar ratio in MCF7 breast cancer cells compared to the other luminal breast cancer cell lines (Figure 5J) corre-lated with the relatively high expression of DPYD in this partic-ular cell line (Figure 2D). Hence, DHU/uracil and DHT/ ratios correlate closely with DPYD expression levels and mesen-chymal character in a number of cellular settings, suggesting that DPYD is enzymatically active in the cancer cell lines that we examined.

DPYD is normally expressed in the liver, where it is the rate-limiting enzyme of a three-step pyrimidine degradation pathway that converts uracil and thymine to β-alanine and 2-methyl-β-alanine, respectively (Figure 5G) (Lohkamp et al., 2010). In the liver, the immediate products of DPYD are further catabo-lized by dihydropyrimidinase (DPYS) and beta-ureidopropionase (UPB1) (van Kuilenburg et al., 2004) (Figure 5G). In contrast, we found that HMLE-Twist-ER and NAMEC cells express only DPYD, but not the other components of this catabolic pathway (Figure S5B). In addition, unlike that of *DPYD*, *DPYS* and *UPB1* expression was not elevated in breast Basal B and high-grade HCC cell lines (Figure S5C). These observations explain why the products of DPYD activity can accumulate in mesen-chymal-like cancer cells but may not do so in normal liver.

## Intracellular Dihydropyrimidines Are Essential for the EMT

To establish the requirement for the immediate products of DPYD in the EMT program, we ectopically expressed DPYS-FLAG, which to our knowledge is the only known enzyme that uses dihydropyrimidines (DHPs) as substrates (van Kuilenburg et al., 2010). We found DPYS-FLAG to be an active enzyme, as it reduced the intracellular concentration of its substrates (DHPs; Figure 6A), while elevating the intracellular abundance of its immediate product, 3-ureidopropionate (Figure S6A). After 15 days of Twist induction by OHT treatment, DPYS-FLAG expression decreased the percentage of cells with a mesen-chymal profile (CD24^{low}/CD44^{high}) (Figure 6B) and inhibited the upregulation of mesenchymal markers (Figure 6C) and the ability of the cells to migrate (Figure 6D) and form mammo-spheres (Figure 6E). These results provide further evidence that the accumulation of DHPs plays a key role in the EMT program.

Next, we asked whether addition of DHU or DHT to culture media could substitute for DPYD loss. Indeed, treatment of shDPYD_1 cells with these metabolites at 10 or 100 μM resulted in a dose-dependent rescue of mammosphere formation (Fig-ure 6F), whereas the DPYD substrate uracil had a significantly smaller effect (Figure S6B), despite the fact that uracil and DHU accumulated to comparable intracellular concentrations (Figure S6C). Therefore, the effect of DPYD knockdown on mam-mosphere formation can be reversed either by ectopic expres-sion of active DPYD (Figures 4C and 5C) or by supplementation of the cell-culture media with its products. Together, these re-sults confirm that the MMS gene product DPYD plays a critical role in the EMT via its capacity to generate dihydropyrimidines.

DISCUSSION

We identified a mesenchymal metabolic signature (MMS) con-sisting of 44 metabolic genes upregulated in cancers bearing mesenchymal markers. Several of these metabolic genes are essential for the EMT, including DPYD, the rate-limiting enzyme of the pyrimidine degradation pathway. Remarkably, the expres-sion of DPYD is not required for cell viability or proliferation, demonstrating the existence of metabolic processes that
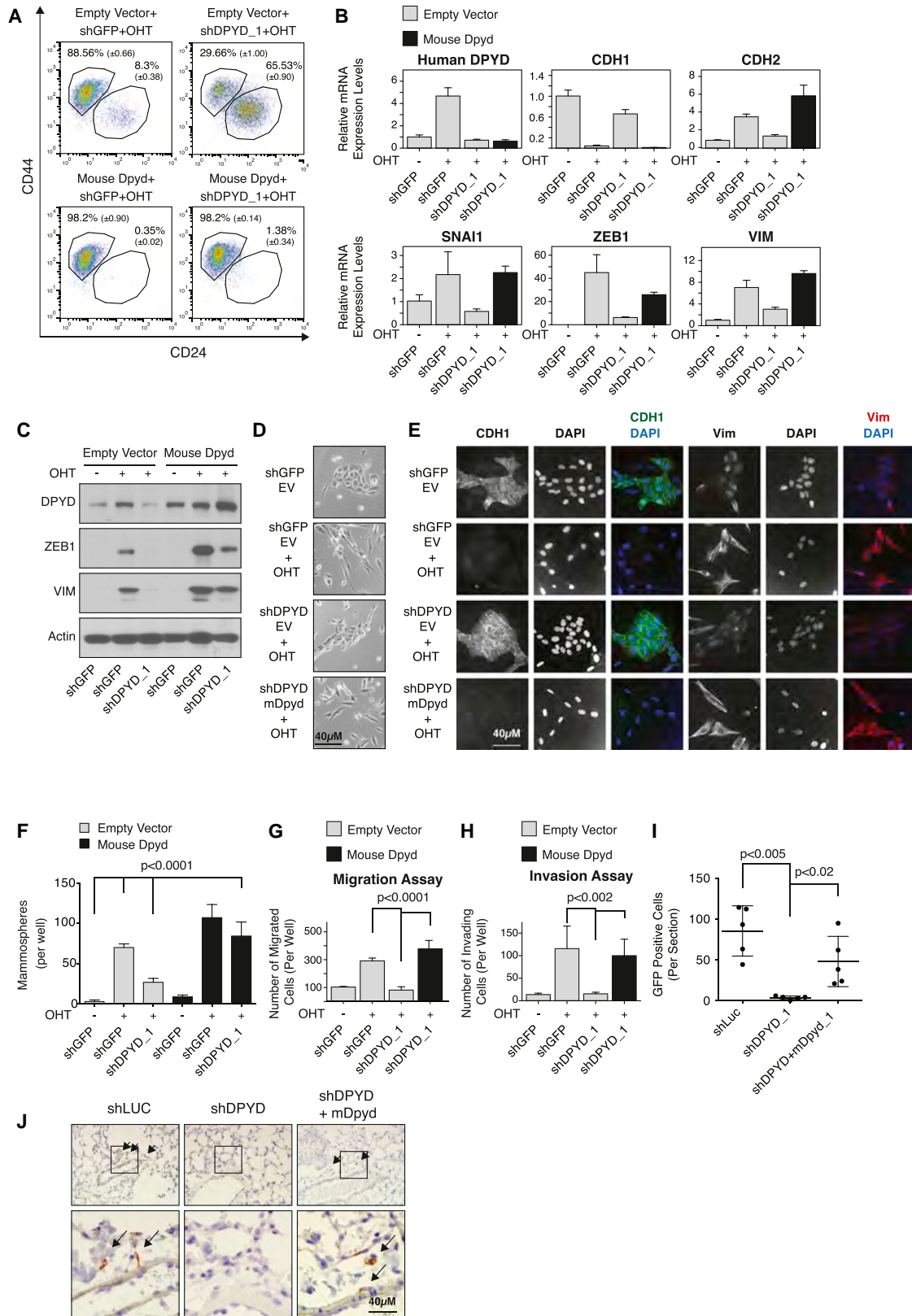
(legend on next page)

specifically enable carcinoma cells to acquire mesenchymal-like characteristics. Because these characteristics are associated with increased cancer aggressiveness, these findings suggest that DPYD activity may play a role in carcinoma progression.

There is a clear difference between the metabolic pathways that are associated with proliferation and those upregulated during the EMT. Compared to resting cells, proliferating cells up-regulate glycolysis and nucleotide biosynthesis pathways (Hu et al., 2013), whereas the mesenchymal metabolic signature (MMS) is enriched in glycan biosynthesis genes. Glycosylation is thought to be one of the most common covalent protein modifications in eukaryotic cells, with a major role in differentiation and mediating cell-cell interactions (Li et al., 2013). Because the EMT is accompanied by major changes in cell morphology and detachment from the surrounding cells, it is reasonable to assume that a major glycan remodeling may occur during the EMT program. Furthermore, glycosylation regulates the function of several key players in the EMT, including the products of the SNAI1 and CD44 genes (Jaggupilli and Elkord, 2012; Park et al., 2010). Thus, we anticipate that future studies will further demonstrate an important role for specific glycan remodeling events in both the mesenchymal phenotype and the EMT program.

After executing the EMT program, epithelial-derived cancer cells acquire traits associated with high-grade malignancy, including resistance to apoptosis and chemotherapy, dedifferentiation, and invasiveness, which can lead to metastatic dissemination from primary tumors (Tam and Weinberg, 2013). Thus, inhibiting the EMT may maintain a tumor in a lower-grade state, potentially increasing therapeutic efficacy and slowing metastasis. The feasibility of manipulating epithelial plasticity is reinforced by studies showing that depletion of ZEB1 by RNA interference in mesenchymal-like cells results in a partial mesenchymal-epithelial transition (MET), presumably through the induction of CDH1 expression (Chaffer et al., 2013). However,

the development of inhibitors targeting transcription factors, such as ZEB1, remains a challenge (Singh and Settleman, 2010). By contrast, many of the enzymes encoded by the MMS have well-defined active sites that can potentially be targeted by small molecules. Here, we demonstrate that DPYD expression and activity are essential for the induction of ZEB1 expression, suggesting that the expression of transcriptional drivers of the EMT program can be modulated through inhibition of metabolic enzymes, such as DPYD.

Many studies have linked DPYD function with acquired tumor resistance to the chemotherapeutic agent 5-FU, but the physiologic role of this enzyme in cancer cells is unknown (Amstutz et al., 2011). By demonstrating that DPYD plays an essential role in the EMT, we provide one of the first indications for its function in cancer. However, there is a clear distinction between this function and the normal role of DPYD in the liver. In the latter, DPYD functions as the first enzyme in a three-step pathway of pyrimidine degradation, whereas we show that in mesenchymal-like cells, the expression of the two downstream enzymes (DPYS and UPB1) is not detectable at the mRNA level. Therefore, the EMT program reconfigures the pyrimidine degradation pathway in order to use only DPYD, presumably because its enzymatic activity fulfills a specific metabolic demand. We suggest that this EMT-dependent metabolic rewiring, which activates only selected components of a given metabolic pathway, is not exclusive to DPYD but can potentially occur in other MMS-related metabolic processes. Thus, through such rewiring, the EMT may confer novel cellular functions to other pathways represented in the MMS as well. Further studies aimed at understanding the role of the MMS genes in cancer may reveal novel metabolic processes that promote cancer aggressiveness.

The function of DPYD in the EMT is dependent upon its products, the dihydropyrimidines (DHPs), DHU and DHT. However, understanding the role of these metabolites in the EMT program is challenging, because no biological function has been ascribed

Figure 4. DPYD Expression Is Essential for EMT Induction

(A) Mouse Dpyd expression rescues the effects of DPYD KD on the EMT. HMLE-Twist-ER cells were infected and treated as indicated, followed by FACS analysis for the cell-surface markers CD24 and CD44. The percentage of cells in each gate is presented (mean ± SD, for n = 3).

(B) Mouse Dpyd rescues the effects of DPYD KD on mesenchymal gene expression. HMLE-Twist-ER cells were infected and treated as indicated. The expression levels of the indicated genes were measured using qRT-PCR. Each value represents the mean ± SEM for n = 3.

(C) Mouse Dpyd rescues the effects of DPYD KD on ZEB1 expression. HMLE-Twist-ER cells were infected and treated as indicated. Followed by immunoblotting with the indicated antibodies.

(D) Mouse Dpyd rescues the effects of DPYD KD on cell morphology. HMLE-Twist-ER cells were infected and treated as in (B) and visualized with bright-field microscopy. EV, empty vector.

(E) Mouse Dpyd rescues the effects of DPYD KD on cell morphology and gene expression. HMLE-Twist-ER cells were infected and treated as in (B), and the indicated proteins were visualized by immunofluorescence.

(F) Mouse Dpyd rescues the effects of DPYD KD on mammosphere formation. Quantification of in vitro mammosphere formation by cells treated as in (C). The data are reported as the number of mammospheres formed per 500 seeded cells; each value represents the mean ± SD for n = 6.
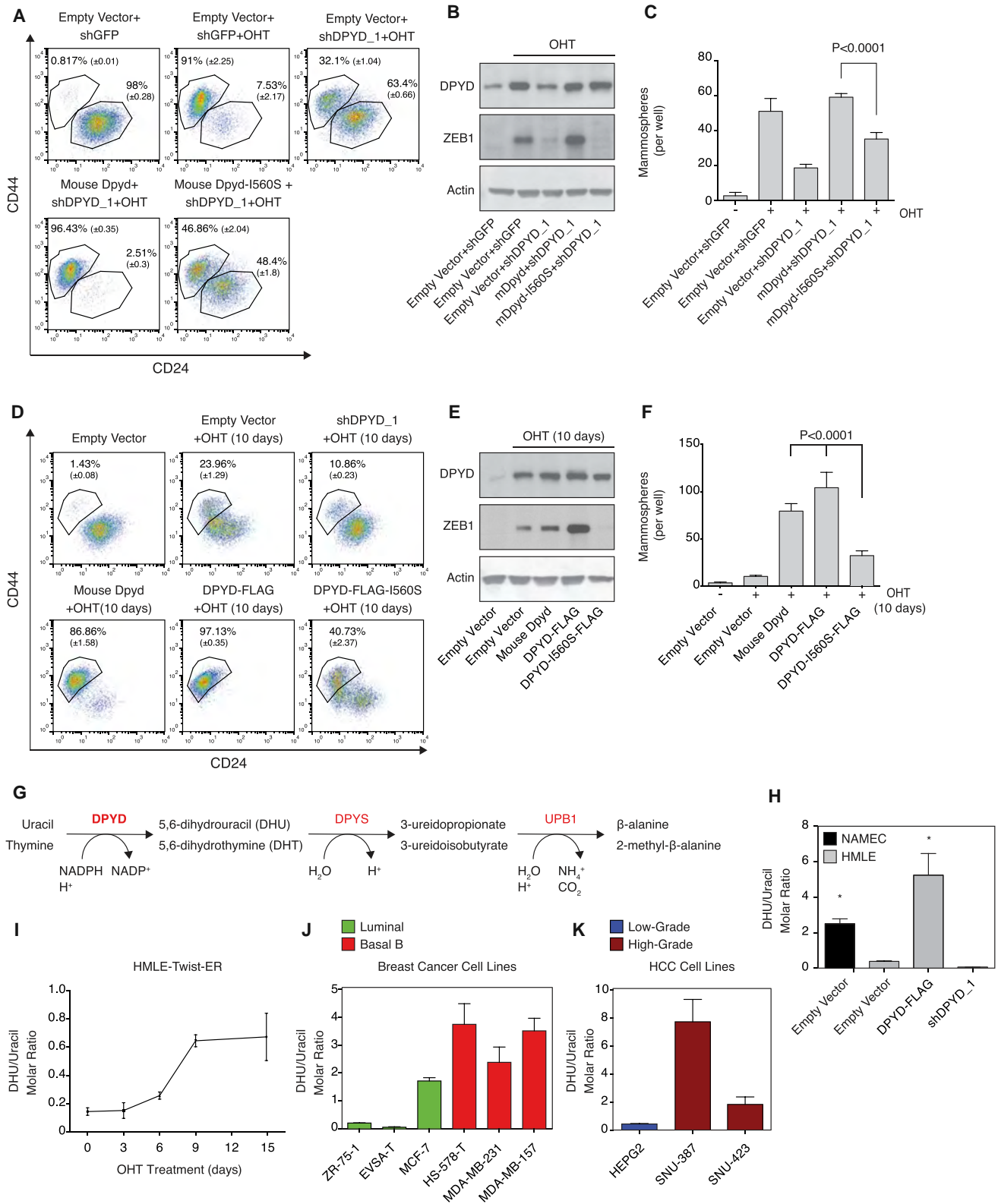
(G) Mouse Dpyd rescues the effects of DPYD KD on cell migration. HMLE-Twist-ER cells were infected and treated as in (B), and their ability to migrate was measured. The data are reported as the number of migrated cells per 50,000 seeded cells; each value represents the mean ± SD for n = 3. EV, empty vector.

(H) Mouse Dpyd rescues the effects of DPYD KD on cell invasiveness. HMLE-Twist-ER cells were infected and treated as in (B) and their ability to migrate through Matrigel was measured. The data are reported as in (G).

(I) DPYD KD inhibits lung extravasation of cancer cells in vivo. HMLER-Twist-ER cells expressing the indicated hairpins, and open reading frames were treated with OHT for 15 days and then injected into the mouse tail vein. After 3 days, the number of GFP-positive cells in each lung was determined by immunohisto-chemistry. Each value (●) represents the average of three nonadjacent lung sections from a single mouse (five mice per group).

(J) DPYD KD inhibits lung extravasation of cancer cells in vivo. Representative GFP-positive cells from the same mouse lungs as in (I) are indicated with arrows. The bottom panels show magnifications of the boxed areas in the upper panels.

See also Figure S4.

A

Empty Vector+ shGFP

Empty Vector+ shGFP+OHT

Empty Vector+ shDPYD_1+OHT

0.817% (±0.01) 98% (±0.28)

91% (±2.25) 7.53% (±2.17)

32.1% (±1.04) 63.4% (±0.66)

Mouse Dpyd+ shDPYD_1+OHT

Mouse Dpyd-I560S + shDPYD_1+OHT

96.43% (±0.35) 2.51% (±0.3)

46.86% (±2.04) 48.4% (±1.8)

CD44

CD24

B

OHT

DPYD

ZEB1

Actin

Empty Vector+shGFP
Empty Vector+shGFP
Empty Vector+shDPYD_1
mDpyd+shDPYD_1
mDpyd-I560S+shDPYD_1

C

P<0.0001

Mammospheres (per well)

80
60
40
20
0

Empty Vector+shGFP
Empty Vector+shGFP
Empty Vector+shDPYD_1
mDpyd+shDPYD_1
mDpyd-I560S+shDPYD_1

- + + + + OHT

D

Empty Vector

Empty Vector +OHT (10 days)

shDPYD_1 +OHT (10 days)

1.43% (±0.08)

23.96% (±1.29)

10.86% (±0.23)

Mouse Dpyd +OHT(10 days)

DPYD-FLAG +OHT (10 days)

DPYD-FLAG-I560S +OHT (10 days)

86.86% (±1.58)

97.13% (±0.35)

40.73% (±2.37)

CD44

CD24

E

OHT (10 days)

DPYD

ZEB1

Actin

Empty Vector
Empty Vector
Mouse Dpyd
DPYD-FLAG
DPYD-I560S-FLAG

F

P<0.0001

Mammospheres (per well)

150
100
50
0

Empty Vector
Empty Vector
Mouse Dpyd
DPYD-FLAG
DPYD-I560S-FLAG

- + + + + OHT (10 days)

G

Uracil → 5,6-dihydrouracil (DHU) → 3-ureidopropionate → β-alanine
Thymine   5,6-dihydrothymine (DHT)   3-ureidoisobutyrate   2-methyl-β-alanine

DPYD ⇵ $NADPH$ H^+ → $NADP^+$

DPYS H_2O → H^+

UPB1 H_2O H^+ → NH_4^+ CO_2

H

■ NAMEC
□ HMLE

DHU/Uracil Molar Ratio

8
6
4
2
0

Empty Vector *
Empty Vector
DPYD-FLAG *
shDPYD_1

I

HMLE-Twist-ER

DHU/Uracil Molar Ratio

1.0
0.8
0.6
0.4
0.2
0

0 3 6 9 12 15
OHT Treatment (days)

J

■ Luminal
■ Basal B

Breast Cancer Cell Lines

DHU/Uracil Molar Ratio

5
4
3
2
1
0

ZR-75-1, EVSA-T, MCF-7, HS-578-T, MDA-MB-231, MDA-MB-157

K

■ Low-Grade
■ High-Grade

HCC Cell Lines

DHU/Uracil Molar Ratio

10
8
6
4
2
0

HEPG2, SNU-387, SNU-423

Figure 5. DPYD Activity Is Essential for the EMT

(A) Mouse Dpyd-I560S fails to rescue the effects of DPYD KD on the EMT. HMLE-Twist-ER cells were infected and treated as indicated, followed by FACS analysis as in Figure 4A.

(legend continued on next page)

the DHPs other than as substrates for the enzyme DPYS. One possibility is that DHPs may act as allosteric regulators of other enzymes (similar to serine regulating the glycolytic enzyme pyruvate kinase M2 [Chaneton et al., 2012]) or as receptor ligands (like the citric acid cycle intermediates succinate and α-ketoglutarate acting as ligands to the G-protein-coupled receptors GPR99 and GPR91, respectively [He et al., 2004]). In this scenario, the DHPs themselves could act as key signaling molecules without further enzymatic processing.

Another potential function for the DHPs is that these pyrimidine bases could be converted to pyrimidine deoxynucleosides or nucleosides and thus could be possibly incorporated into DNA or RNA, respectively. Support for this latter possibility comes from previous studies showing that genotoxic agents can damage DNA precursors (dNTPs) in the nucleotide pools of bacterial cells (Dolinnaya et al., 2013). These chemically altered dNTPs, including the deoxynucleotide triphosphate form of DHT (DHdTTP), have been found to be incorporated into bacterial genomes (Dolinnaya et al., 2013) and are able to substitute for deoxythymidine triphosphate (dTTP) as substrates for the *Escherichia coli* DNA polymerase I and Klenow fragments in vitro (Ide and Wallace, 1988; Ide et al., 1987). It remains to be determined whether such modified nucleotides can be produced in human cells and if so how they affect cellular phenotypes.

EXPERIMENTAL PROCEDURES

For further details, please refer to the Extended Experimental Procedures.

Identification of the MMS Genes

For each metabolic gene, the ratio between the mean expression level in mesenchymal (mesenchymal group; Figure 1) and nonmesenchymal cell lines (all other groups) was determined. The mean and SD of all the metabolic gene expression ratios was calculated, and all genes upregulated above a 2-fold or below a −2 were classified as MMS (Table S2).

Pooled shRNA Screen

pLKO.1 lentiviral plasmids encoding shRNAs targeting 74 genes (listed in Table S3) were obtained and combined to generate a plasmid pool (Possemato et al., 2011). HMLE cells were infected with the pooled lentivirus at an MOI of 0.2–0.5 so as to ensure that most cells contained only one viral integrant. Cells were selected for 3 days with 0.5 mg/ml puromycin, after which time 10^6 cells were removed, washed, and frozen at −80°C (Figure 3A, day 0). The remaining cells were split into OHT-treated and untreated samples. After 15 days, the OHT-treated cells were trypsinized, washed with PBS +1% inactivated fetal calf (IFC) serum, and FACS-sorted using CD44/CD24 antibodies in order to separate the mesenchymal and epithelial populations.

Genomic DNA was isolated from all the cells using the QIAampDNA Mini Kit (QIAGEN). To amplify the shRNAs encoded in the genomic DNA, PCR was performed for 33 cycles at an annealing temperature of 66°C using 3.5 μg of genomic DNA, the primer pair indicated below, and DNA polymerase (TAKARA Ex taq, Clontech lab, RR001A). Forward primers containing unique 4 nt barcodes were used (see below) so that PCR products obtained from many samples could be sequenced together. After purification, the PCR products from each cell sample were quantified by ethidium bromide staining (Sigma-Aldrich, E1510) after gel electrophoresis, pooled in equal proportions, and analyzed by high-throughput sequencing (Illumina). The shRNAs from all four DNA samples (day 0, day 15 untreated, day 15 OHT-treated mesenchymal, and day 15 OHT-treated epithelial) were sequenced together. Sequencing reads were deconvoluted using GNU Octave software by segregating the sequencing data by barcode and matching the shRNA stem sequences to those expected to be present in the shRNA pool, allowing for mismatches of up to 3 nt. The $\log_2$ values reported are the average $\log_2$ of the fold change in the abundance of each shRNA in the mesenchymal-like samples compared to epithelial cells. The mean and SD of the control hairpins (*GFP*, *RFP*, *Luciferase*, and *LacZ*) were calculated and used to set a cutoff (one SD below the control mean). Every gene that had at least two hairpins with a $\log_2$ value below the cutoff was considered a hit (hairpin ratio list is in Table S4).

Metabolite Extraction

Solvents were obtained from Fisher Scientific and were Optima LC-MS/MS grade, except where otherwise specified. Cells grown in standard tissue culture plates (~500,000 cells per sample) were washed twice in an ice-cold solution of 0.9% NaCl in deionized water, followed by extraction on dry ice in 1 ml 80% methanol containing 10 ng/ml phenylalanine-d_8 and valine-d_8

(B) Mouse Dpyd-I560S fails to rescue the effects of DPYD KD on ZEB1 expression. HMLE-Twist-ER cells were infected and treated as indicated, followed by immunoblotting with the indicated antibodies.

(C) The I560S mutation prevents the capacity of mDpyd to promote mammosphere formation in DPYD KD cells. Cells treated as in (B) were subjected to the in vitro mammosphere formation assay as in Figure 4F. Each value represents the mean ± SD for n = 6.

(D) DPYD activity accelerates the EMT. HMLE-Twist-ER cells infected with the indicated constructs were either left untreated or treated with OHT for 10 days, followed by FACS analysis as in Figure 4A. The percentage of cells in each gate is presented.

(E) Unlike wild-type human DPYD, overexpression of the catalytically attenuated DPYD (DPYD-I560S) does not potentiate the Twist-induced ZEB1 upregulation after 10 days of OHT treatment. Cells infected with the indicated constructs were either left untreated or treated with OHT for 10 days, followed by immunoblotting with the indicated antibodies.

(F) DPYD activity enhances mammosphere formation. Cells treated as in (D) were subjected to the in vitro mammosphere formation assay as in Figure 4F. Each value represents the mean ± SD for n = 6.

(G) Schematic presentation of the pyrimidine degradation pathway. Gene names are marked in red: DPYD, dihydropyrimidine dehydrogenase (rate-limiting step); DPYS, dihydropyrimidinase; UPB1, beta-ureidopropionase.

(H) Modulation of DPYD expression affects the cellular DHU/uracil molar ratio. DHU and uracil levels were measured by LC-MS/MS in NAMEC or HMLE-Twist-ER cell lines expressing empty vector, DPYD-FLAG, or shDPYD_1 hairpin. Each value represents the mean ± SD for n = 3.

(I) The cellular DHU/uracil ratio increases during EMT. HMLE-Twist-ER cells were treated with OHT for 15 days. At the indicated time points, samples were collected and subjected to LC-MS/MS analysis to determine DHU and uracil levels. The molar concentration ratio between the two metabolites in each sample is presented. Each value represents the mean ± SD for n = 3.

(J) The cellular DHU/uracil ratio is elevated in Basal B relative to luminal breast cancer cell lines. The concentrations of DHU and uracil were measured in the indicated breast cancer cell lines (green, luminal; red, Basal B) using LC-MS/MS. Each value represents the mean ± SD for n = 3.

(K) The cellular DHU/uracil ratio is elevated in high-grade relative to low-grade HCC cell lines. The concentrations of DHU and uracil were measured in HCC cell lines (blue, low-grade; brown, high-grade). Each value represents the mean ± SD for n = 3.

See also Figure S5.
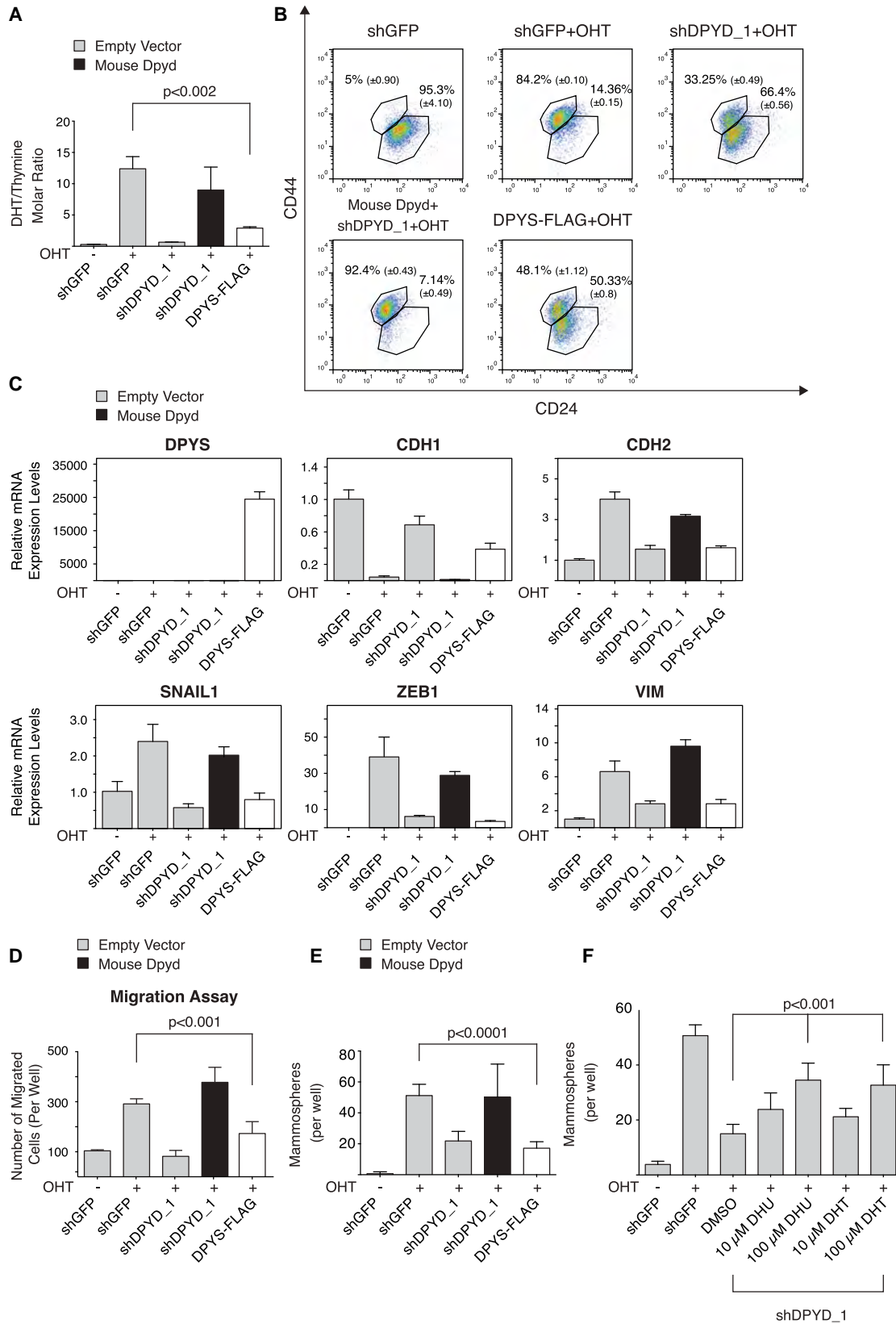
(legend on next page)

(Sigma-Aldrich, 492485 and 486612, respectively) as internal standards. The cell mixtures were shaken vigorously on a Vortex mixer for 10 min at 4°C, vacuum dried, and resuspended in 100 μl LC-MS/MS-grade water (Fisher). These extracts were then centrifuged at 15,000 × *g* at 4°C for 10 min, and the supernatants were passed through a cellulose acetate particulate filter (National Scientific).

Migration and Invasion Assays

For the migration assay, 50,000 cells were seeded into 24-well cell culture inserts with 8 μm pores (BD falcon, 353097). For invasion assay 50,000 cells were seeded on BD Matrigel invasion chambers with 8 μm pores (Corning, 62405-744). For both assays, after 20 hr, the cells on the upper surface of the filters were removed using cotton swabs. For visualization, the cells were fixed with 4% PFA in PBS, followed by methanol treatment and crystal violet. Three fields per filter were counted. Data are presented as migrated cells per filter.

Animal Studies, Histology, and Immunohistochemistry

For extravasation studies, HMLE-Twist-ER were transformed with a GFP-tagged-H-Ras[G12V] oncogene (MSCV-RAS-IRES-GFP, HMLER) and sorted for GFP-expressing cells. HMLER cells were then infected with OHT-inducible Twist (HMLER-Twist-ER) and the indicated hairpins and vectors. After 15 days of OHT treatment (in culture), 1,000,000 cells per mouse were injected through the tail vein (five mice per sample). After 3 days, the lungs were extracted, fixed in 10% formalin, paraffin embedded, and sectioned. The sections were subjected to deparaffinization and antigen retrieval with 10 mM sodium citrate buffer (pH 6.0) in a pressurized decloaking chamber (Biocare Medical). Antibodies were diluted in 5% horse serum and 0.1% Tween in TBS, which was also used for blocking. The Vectastain ABC Immunoperoxidase Detection Kit (Vector Labs, PK6101) and DAB+ Substrate Kit (Dako, K3467) were used for chromogenic labeling. Images were acquired using an Olympus BX41 microscope and CellSens[R] software. For all image-based data, acquisition and processing steps were carried out using the same parameters across the entire set. For all image-based data, acquisition and processing steps were carried out using the same parameters across the entire set. For each mouse, three nonadjacent lung sections were stained and examined to obtain the total number of GFP-positive cells. All research involving animals was complied with protocols approved by the MIT Committee on Animal Care.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2014.07.032.

AUTHOR CONTRIBUTIONS

Y.D.S. and D.M.S. designed the research, and Y.D.S. performed cell-based experiments. E.F. and M.E.P. performed mass spectrometry experiments. W.C.C. and Y.D.S. performed qRT-PCR analysis. W.W.C. and Y.D.S. performed FACS analysis. J.R.C. and D.K. performed cloning. W.L.T., B.B., N.K., R.P., F.R., and Y.D.S performed mouse experiments. P.T. and Y.D.S. wrote scripts. Y.D.S., R.A.W., M.B.Y., and D.M.S. interpreted the results. Y.D.S., E.F., and D.M.S. wrote the paper.

ACKNOWLEDGMENTS

REFERENCES

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc. Natl. Acad. Sci. USA *100*, 3983–3988.

Amstutz, U., Froehlich, T.K., and Largiadèr, C.R. (2011). Dihydropyrimidine dehydrogenase gene as a major predictor of severe 5-fluorouracil toxicity. Pharmacogenomics *12*, 1321–1336.

Benesch, M.G.K., Tang, X., Maeda, T., Ohhata, A., Zhao, Y.Y., Kok, B.P.C., Dewald, J., Hitt, M., Curtis, J.M., McMullen, T.P.W., and Brindley, D.N. (2014). Inhibition of autotaxin delays breast tumor growth and lung metastasis in mice. FASEB J. *28*, 2655–2666.

Brabletz, T. (2012). To differentiate or not—routes towards metastasis. Nat. Rev. Cancer *12*, 425–436.

Büchel, B., Rhyn, P., Schürch, S., Bühr, C., Amstutz, U., and Largiadèr, C.R. (2013). LC-MS/MS method for simultaneous analysis of uracil, 5,6-dihydrouracil, 5-fluorouracil and 5-fluoro-5,6-dihydrouracil in human plasma for therapeutic drug monitoring and toxicity prediction in cancer patients. Biomed. Chromatogr. *27*, 7–16.

Cantor, J.R., and Sabatini, D.M. (2012). Cancer cell metabolism: one hallmark, many faces. Cancer Discov. *2*, 881–898.

Figure 6. DPYS Expression Inhibits the EMT Program

(A) DPYS-FLAG expression reduces the cellular DHT/thymine ratio. HMLE-Twist-ER cells were infected and treated as indicated, followed by LC-MS/MS to measure the intracellular DHT and thymine concentration. Each value represents the mean ± SD for n = 3.

(B) DPYS-FLAG expression inhibits the EMT program. HMLE-Twist-ER cells were infected and treated as in (A), followed by FACS analysis for the cell-surface markers CD24 and CD44. The percentage of cells in each gate is presented.

(C) DPYS-FLAG expression reduces the expression level of mesenchymal genes. HMLE-Twist-ER cells were infected and treated as in (A), and the indicated gene expression levels were measured using qRT-PCR. Each value represents the mean ± SEM for n = 3.

(D) DPYS-FLAG expression inhibits cell migration. HMLE-Twist-ER cells were infected and treated as in (A), and their ability to migrate was determined. The data are reported as in Figure 4G. Each value represents the mean ± SD for n = 3.

(E) DPYS-FLAG expression inhibits mammosphere formation. Quantification of in vitro mammosphere formation by cells infected and treated as in (A). The data are reported as in Figure 4F. Each value represents the mean ± SD for n = 6.

(F) DPYD products rescue the effect of DPYD KD on mammosphere formation. HMLE-Twist-ER cells expressing shDPYD_1 were treated with the indicated concentrations of DHU or DHT and subjected to the in vitro mammosphere formation assay as in (E). The data are reported as in Figure 4F. Each value represents the mean ± SD for n = 6.

See also Figure S6.

Carey, L., Winer, E., Viale, G., Cameron, D., and Gianni, L. (2010). Triple-negative breast cancer: disease entity or title of convenience? Nat. Rev. Clin. Oncol. *7*, 683–692.

Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell *154*, 61–74.

Chaneton, B., Hillmann, P., Zheng, L., Martin, A.C.L., Maddocks, O.D.K., Chokkathukalam, A., Coyle, J.E., Jankevics, A., Holding, F.P., Vousden, K.H., et al. (2012). Serine is a natural ligand and allosteric activator of pyruvate kinase M2. Nature *491*, 458–462.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *462*, 739–744.

Dolinnaya, N.G., Kubareva, E.A., Romanova, E.A., Trikin, R.M., and Oretskaya, T.S. (2013). Thymidine glycol: the effect on DNA molecular structure and enzymatic processing. Biochimie *95*, 134–147.

Elbein, S.C., Zheng, H., Jia, Y., Chu, W., Cooper, J.J., Hale, T., and Zhang, Z. (2004). Molecular screening of the human glutamine-fructose-6-phosphate amidotransferase 1 (GFPT1) gene and association studies with diabetes and diabetic nephropathy. Mol. Genet. Metab. *82*, 321–328.

Gottlieb, E., and Tomlinson, I.P.M. (2005). Mitochondrial tumour suppressors: a genetic and biochemical update. Nat. Rev. Cancer *5*, 857–866.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

He, W., Miao, F.J.P., Lin, D.C.H., Schwandner, R.T., Wang, Z., Gao, J., Chen, J.-L., Tian, H., and Ling, L. (2004). Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. Nature *429*, 188–193.

Hsu, P.P., and Sabatini, D.M. (2008). Cancer cell metabolism: Warburg and beyond. Cell *134*, 703–707.

Hu, J., Locasale, J.W., Bielas, J.H., O'Sullivan, J., Sheahan, K., Cantley, L.C., Vander Heiden, M.G., and Vitkup, D. (2013). Heterogeneity of tumor-induced gene expression changes in the human metabolic network. Nat. Biotechnol. *31*, 522–529.

Ide, H., and Wallace, S.S. (1988). Dihydrothymidine and thymidine glycol triphosphates as substrates for DNA polymerases: differential recognition of thymine C5-C6 bond saturation and sequence specificity of incorporation. Nucleic Acids Res. *16*, 11339–11354.

Ide, H., Melamede, R.J., and Wallace, S.S. (1987). Synthesis of dihydrothymidine and thymidine glycol 5′-triphosphates and their ability to serve as substrates for Escherichia coli DNA polymerase I. Biochemistry *26*, 964–969.

Jaggupilli, A., and Elkord, E. (2012). Significance of CD44 and CD24 as cancer stem cell markers: an enduring ambiguity. Clin. Dev. Immunol. *2012*, 708036.

Kao, J., Salari, K., Bocanegra, M., Choi, Y.-L., Girard, L., Gandhi, J., Kwei, K.A., Hernandez-Boussard, T., Wang, P., Gazdar, A.F., et al. (2009). Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS ONE *4*, e6146.

Li, S., Mo, C., Peng, Q., Kang, X., Sun, C., Jiang, K., Huang, L., Lu, Y., Sui, J., Qin, X., and Liu, Y. (2013). Cell surface glycan alterations in epithelial mesenchymal transition process of Huh7 hepatocellular carcinoma cell. PLoS ONE *8*, e71273.

Locasale, J.W., Grassian, A.R., Melman, T., Lyssiotis, C.A., Mattaini, K.R., Bass, A.J., Heffron, G., Metallo, C.M., Muranen, T., Sharfi, H., et al. (2011). Phosphoglycerate dehydrogenase diverts glycolytic flux and contributes to oncogenesis. Nat. Genet. *43*, 869–874.

Lohkamp, B., Voevodskaya, N., Lindqvist, Y., and Dobritzsch, D. (2010). Insights into the mechanism of dihydropyrimidine dehydrogenase from site-directed mutagenesis targeting the active site loop and redox cofactor coordination. Biochim. Biophys. Acta *1804*, 2198–2206.

Ma, Z., and Vosseller, K. (2013). O-GlcNAc in cancer biology. Amino Acids *45*, 719–733.

Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesen-

chymal transition generates cells with properties of stem cells. Cell *133*, 704–715.

Nieto, M.A., and Cano, A. (2012). The epithelial-mesenchymal transition under control: global programs to regulate epithelial plasticity. Semin. Cancer Biol. *22*, 361–368.

Nomura, D.K., Long, J.Z., Niessen, S., Hoover, H.S., Ng, S.-W., and Cravatt, B.F. (2010). Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. Cell *140*, 49–61.

Offer, S.M., Wegner, N.J., Fossum, C., Wang, K., and Diasio, R.B. (2013). Phenotypic profiling of DPYD variations relevant to 5-fluorouracil sensitivity using real-time cellular analysis and in vitro measurement of enzyme activity. Cancer Res. *73*, 1958–1968.

Onganer, P.U., Seckl, M.J., and Djamgoz, M.B.A. (2005). Neuronal characteristics of small-cell lung cancer. Br. J. Cancer *93*, 1197–1201.

Parham, D.M. (2001). Neuroectodermal and neuroendocrine tumors principally seen in children. Am. J. Clin. Pathol. Suppl. *115*, S113–S128.

Park, J.G., Lee, J.H., Kang, M.S., Park, K.J., Jeon, Y.M., Lee, H.J., Kwon, H.S., Park, H.S., Yeo, K.S., Lee, K.U., et al. (1995). Characterization of cell lines established from human hepatocellular carcinoma. Int. J. Cancer *62*, 276–282.

Park, S.Y., Kim, H.S., Kim, N.H., Ji, S., Cha, S.Y., Kang, J.G., Ota, I., Shimada, K., Konishi, N., Nam, H.W., et al. (2010). Snail1 is stabilized by O-GlcNAc modification in hyperglycaemic condition. EMBO J. *29*, 3787–3796.

Possemato, R., Marks, K.M., Shaul, Y.D., Pacold, M.E., Kim, D., Birsoy, K., Sethumadhavan, S., Woo, H.-K., Jang, H.G., Jha, A.K., et al. (2011). Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature *476*, 346–350.

Scheel, C., and Weinberg, R.A. (2011). Phenotypic plasticity and epithelial-mesenchymal transitions in cancer and normal stem cells? Int. J. Cancer *129*, 2310–2314.

Schulze, A., and Harris, A.L. (2012). How cancer metabolism is tuned for proliferation and vulnerable to disruption. Nature *491*, 364–373.

Singh, A., and Settleman, J. (2010). EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene *29*, 4741–4751.

Stagg, J., Divisekera, U., McLaughlin, N., Sharkey, J., Pommey, S., Denoyer, D., Dwyer, K.M., and Smyth, M.J. (2010). Anti-CD73 antibody therapy inhibits breast tumor growth and metastasis. Proc. Natl. Acad. Sci. USA *107*, 1547–1552.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat. Med. *19*, 1438–1449.

Tam, W.L., Lu, H., Buikhuisen, J., Soh, B.S., Lim, E., Reinhardt, F., Wu, Z.J., Krall, J.A., Bierie, B., Guo, W., et al. (2013). Protein kinase C α is a central signaling node and therapeutic target for breast cancer stem cells. Cancer Cell *24*, 347–364.

Taube, J.H., Herschkowitz, J.I., Komurov, K., Zhou, A.Y., Gupta, S., Yang, J., Hartwell, K., Onder, T.T., Gupta, P.B., Evans, K.W., et al. (2010). Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc. Natl. Acad. Sci. USA *107*, 15449–15454.

Tennant, D.A., Durán, R.V., and Gottlieb, E. (2010). Targeting metabolic transformation for cancer therapy. Nat. Rev. Cancer *10*, 267–277.

Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell *22*, 725–736.

Ulanovskaya, O.A., Zuhl, A.M., and Cravatt, B.F. (2013). NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. Nat. Chem. Biol. *9*, 300–306.

van Kuilenburg, A.B.P., Meinsma, R., Beke, E., Assmann, B., Ribes, A., Lorente, I., Busch, R., Mayatepek, E., Abeling, N.G.G.M., van Cruchten, A., et al. (2004). beta-Ureidopropionase deficiency: an inborn error of pyrimidine degradation associated with neurological abnormalities. Hum. Mol. Genet. *13*, 2793–2801.

van Kuilenburg, A.B.P., Dobritzsch, D., Meijer, J., Meinsma, R., Benoist, J.-F., Assmann, B., Schubert, S., Hoffmann, G.F., Duran, M., de Vries, M.C., et al. (2010). Dihydropyrimidinase deficiency: Phenotype, genotype and structural consequences in 17 patients. Biochim. Biophys. Acta *1802*, 639–648.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al.; Cancer Genome Atlas Research Network (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98–110.

Ward, P.S., and Thompson, C.B. (2012). Metabolic reprogramming: a cancer hallmark even Warburg did not anticipate. Cancer Cell *21*, 297–308.

Zhang, H., Jia, Y., Cooper, J.J., Hale, T., Zhang, Z., and Elbein, S.C. (2004). Common variants in glutamine:fructose-6-phosphate amidotransferase 2 (GFPT2) gene are associated with type 2 diabetes, diabetic nephropathy, and increased GFPT2 mRNA levels. J. Clin. Endocrinol. Metab. *89*, 748–755.

Zhang, W.C., Shyh-Chang, N., Yang, H., Rai, A., Umashankar, S., Ma, S., Soh, B.S., Sun, L.L., Tai, B.C., Nga, M.E., et al. (2012). Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. Cell *148*, 259–272.