Glycine Decarboxylase Activity Drives Non-Small Cell Lung Cancer Tumor-Initiating Cells and Tumorigenesis

Wen Cai Zhang, Ng Shyh-Chang, He Yang, Shivshankar Umashankar, Siming Ma, Boon Seng Soh, Li Li Sun, Bee Choo Tai, Min En Nga, Kishore Kumar Bhakoo, Senthil Raja Jayapal, Massimo Nichane, Qiang Yu, Dokeu A. Ahmed, Christie Tan, Wong Poo Sing, John Tam, Agasthian Thirugananam, Monireh Soroush Noghabi, Yin Huei Pang, Haw Siang Ang, Wayne Mitchell, Paul Robson, Philipp Kaldis, Ross Andrew Soo, Sanjay Swarup, Elaine Hsuen Lim, and Bing Lim

SUMMARY

Identification of the factors critical to the tumor-initiating cell (TIC) state may open new avenues in cancer therapy. Here we show that the metabolic enzyme glycine decarboxylase (GLDC) is critical for TICs in non-small cell lung cancer (NSCLC). TICs from primary NSCLC tumors express high levels of the oncogenic stem cell factor LIN28B and GLDC, which are both required for TIC growth and tumorigenesis. Overexpression of GLDC and other glycine/serine enzymes, but not catalytically inactive GLDC, promotes cellular transformation and tumorigenesis. We found that GLDC induces dramatic changes in glycolysis and glycine/serine metabolism, leading to changes in pyrimidine metabolism to regulate cancer cell proliferation. In the clinic, aberrant activation of GLDC correlates with poorer survival in lung cancer patients, and aberrant GLDC expression is observed in multiple cancer types. This link between glycine metabolism and tumorigenesis may provide novel targets for advancing anticancer therapy.

INTRODUCTION

Despite numerous advances in our knowledge of cancer (Vogelstein and Kinzler, 2004; Hanahan and Weinberg, 2011), our ability to develop clinically effective therapies based on this understanding has met with limited success. Current therapies can control tumor growth initially, but most patients ultimately relapse. One prominent example is lung cancer, the leading cause of cancer-related mortality with over 1 million deaths each year (Jemal et al., 2011). Non-small cell lung cancer (NSCLC) accounts for approximately 85% of all lung cancers. Although NSCLC patients with epidermal growth factor receptor (EGFR) mutations respond to EGFR inhibitors initially, most patients experience a relapse within 1 year (Sequist et al.,...
These findings underscore the urgent need for both combination therapies and also new approaches to treat cancerous tumors. One such approach may be to target tumor-initiating cells (TICs).

Data from leukemias, germ cell tumors, and a number of solid tumors support the notion that cancers are maintained by a subpopulation of self-renewing and evolving TICs. This is also popularly known as the cancer stem cell (CSC) model (Reya et al., 2001; Rosen and Jordan, 2009). Although the validity of the CSC model is an issue of controversy in melanoma (Quintana et al., 2008, 2010; Boiko et al., 2010; Civenni et al., 2011), many other solid tumors appear to follow the CSC model (Ishizawa et al., 2010). Recently it was proposed that at earlier stages of tumorigenesis, rare TIC clones differentiate into nonmalignant progeny to form the bulk of the tumor, whereas at advanced stages, TIC clones constitute the bulk of the tumor (Boiko et al., 2010). Studies with mouse models of lung cancer have also begun to reconcile the connection between the evolving genotype of TIC clones and the surface phenotype of TICs (Curtis et al., 2010). Thus accumulated findings suggest that targeting TICs may be a promising approach for eradicating tumors early. However, progress in the targeting of TICs to improve cancer therapy has been hindered by a lack of understanding of the molecular pathways that are critical to TICs.

Recent studies have led to an emerging appreciation of the importance of metabolic reprogramming in cancer (Hsu and Sabatini, 2008; Vander Heiden et al., 2009; Hanahan and Weinberg, 2011). Most recently, the embryonic isoform of pyruvate kinase PKM2, in collaboration with phosphoglycerate mutase, was found to regulate the shift from oxidative phosphorylation to glycolysis in cancer cells (Christofk et al., 2008; Vander Heiden et al., 2009; Hanahan and Weinberg, 2011). Most recently, the embryonic isoform of pyruvate kinase PKM2, in collaboration with phosphoglycerate mutase, was found to regulate the shift from oxidative phosphorylation to glycolysis in cancer cells (Christofk et al., 2008; Vander Heiden et al., 2009; Hanahan and Weinberg, 2011). These findings have led to a resurgence of interest in the Warburg effect—the phenomenon whereby cancer cells, like embryonic cells, preferentially use glycolysis even under aerobic conditions (Warburg, 1956). Besides glycolysis, an arm of metabolism that results in sarcosine production has also been implicated in prostate cancer (Sreekumar et al., 2009). These data suggest that metabolic reprogramming is crucial for tumorigenesis, and much remains to be uncovered.

Here we show that glycine metabolism and the metabolic enzyme glycine decarboxylase (GLDC) drive TICs and tumorigenesis in NSCLC. Using CD166 as a surface marker and NOD/SCID Il2rg−/− mice as xenotransplantation recipients, we isolated lung TICs from a broad range of primary NSCLC tumors (stages I–III). Primary lung TICs express high levels of LIN28B, GLDC, and many other glycine/serine metabolism enzymes. Both LIN28B and GLDC were required for lung TIC proliferation and tumor growth. Overexpression of GLDC alone, and other glycine/serine enzymes, promotes cellular transformation both in vitro and in vivo. Metabolomic analysis shows that GLDC overexpression induces dramatic changes in glycolysis and glycine metabolism, leading to changes in pyrimidine metabolism for cancer cell proliferation. In human patients, aberrant upregulation of GLDC is significantly associated with higher mortality from lung cancer, and aberrant GLDC expression is observed in multiple cancer types. Our findings establish a link between glycine metabolism and tumorigenesis and may provide novel targets for advancing anticancer therapy.

RESULTS

TICs in Lung Cancer

To assess the cellular heterogeneity within NSCLC, we obtained freshly resected lung tumors from 36 human patients with a broad range of stage I–III primary NSCLC (Table S1 available online). Patient lung cancer cells were directly transplanted subcutaneously into NOD/SCID Il2rg−/− mice with Matrigel (Quintana et al., 2008). Using this maximally sensitive assay, we estimated by limiting dilution analysis that lung TICs exist with a low frequency of 1 in 4 × 105 cells in unsorted NSCLC tumor cells (n = 36 patients; Figure 1A), consistent with published findings (Ishizawa et al., 2010).

To profile the surface phenotype of this subpopulation of lung TICs, we fractionated the NSCLC tumors by fluorescence-activated cell sorting (FACS; Figure S1A). After excluding hematopoietic and endothelial cells (Lin−), we tested a panel of cell-surface markers, including CD166, CD44, CD133, and EpCAM (Figure 1B). We found that CD166 was the most robust marker for enriching the lung TIC subpopulation, compared to CD133, CD44, or EpCAM, allowing us to reliably enrich lung TICs by nearly 100-fold (Figures 1A and 1B). In 12 out of 12 NSCLC patient tumors (lung adenocarcinoma), the CD166+ Lin− fraction contained cells that consistently initiated lung tumor formation in vivo. In contrast, CD166− Lin− tumor cells generally failed to initiate lung tumor formation even after 8 months of observation, although they also expressed carcinoembryonic antigen (CEA), a tumor-specific marker not expressed in normal adult lung cells (Figures 1A, 1B, and S1B). Similar results were observed in lung squamous cell carcinoma and large cell carcinoma (Figure S1C). Although CD166 expression varied across the NSCLC tumors we examined, CD166 was consistently higher in lung tumors than in normal adjacent lung tissues (n = 25 patients; Figures S1D and S1E).

CD166+ lung TICs demonstrate a capacity for self-renewal and differentiation in vitro. Serial transplantations showed that only the CD166+ fraction was able to self-renew and initiate primary and secondary xenograft tumors (Figures 1A and S1F). Upon transplantation, CD166+ lung TICs differentiated to form xenograft tumors that phenocopy the complex cytoarchitecture of their parental patient tumors, sharing similar histological morphology by hematoxylin-eosin (H&E) staining and similar tissue distributions of CD166, cytokeratin, E-cadherin, vimentin, smooth muscle actin, and synaptophysin (Figures 1C, S1G, and S1H). Furthermore, we found that transplants with more TICs grow more rapidly, suggesting that lung TIC frequency is correlated with tumor growth rate (Figures 1D and S1).

The self-renewal capacity of CD166+ lung TICs is further corroborated by in vitro assays. We tested the CD166+ fraction for the ability to form tumor spheres, a widely used in vitro technique for assessing self-renewal capacity (Dontu et al., 2003). Although both primary CD166+ and CD166− cells remained viable in vitro, only primary CD166+ but not CD166− cells were able to form compact self-renewing spheres (n = 9 patients; Figures 1E, 1F, and S1J). Using immunofluorescence and flow cytometry, we found that the lung tumor spheres retained high levels of CD166 expression but undetectable CD133 expression.
in contrast (Figures S1K and S1L). When primary lung tumor spheres were dissociated into single cells and transplanted into NOD/SCID Il2rγ−/− mice in vivo, we found that as few as 1–5 single cells consistently initiated tumorigenesis (Figures 1G and S1M).

The increased tumor-initiating frequency of lung tumor sphere cells suggests that they are even more highly enriched for lung TICs than the patient tumor CD166+ fraction, and that lung TICs expanded during in vitro culture to form tumor spheres.

To test whether CD166 drives tumorigenicity in lung TICs, we...
knocked down CD166 in two lines of NSCLC patient-derived tumor spheres by retroviral shRNA (Figure S1 N). We found that the tumorigenicity of lung TICs in the tumor spheres was not significantly affected by CD166 shRNA, demonstrating that CD166 is an inert cell-surface marker that enriches for lung TICs (Figures S1 O–S1Q).

### Lung TICs Express High levels of Glycine/Serine Metabolism Enzymes

To gain a deeper understanding of the molecular basis for the TIC state and its tumorigenic capacity, we sought to obtain a molecular signature for lung TICs. To do this, we performed genome-wide transcriptome analysis on CD166\(^+\)/Lin\(^-\) tumor cells, CD166\(^+\) Lin\(^-\) tumor cells, and lung tumor spheres, in increasing order of lung TIC frequency (Figure 2A). As a negative control, we also profiled CD166\(^+\) versus CD166\(^-\) cells from normal adjacent lung tissues (n = 3 patients; Table S1). This led us to a profile of genes that are upregulated and downregulated in lung TICs, compared to non-TICs (Figure 2B).

Lung TIC-associated genes include the oncogenic stem cell factor LIN28B, embryonic lung transcription factors like PEA3 and the trachealless homolog NPAS1 (Viswanathan et al., 2009;...
show that all three oncogenes induce GLDC to fuel one-carbon metabolism (Kume et al., 1991). High expression of GLDC (Liu et al., 2003; Levesque et al., 2007), as well as cell-cycle regulators like CCNB1 and GADD45G (Figure 2C). The highest-ranking genes were validated by qRT–PCR (Figure 2S2A), KEGG pathway analysis of the lung TIC-gene profile showed that the top enriched pathways were “cell cycle,” “DNA replication,” “glycine, serine, and threonine metabolism,” “pyrimidine metabolism,” “MAPK signaling pathway,” and “p53 signaling pathway” (Figure 2D). Within the glycine, serine, and threonine metabolism pathway, we found that glycine/serine metabolism enzymes like GLDC, glycine C-acetyltransferase (GCAT), serine hydroxymethyltransferase (SHMT1), phosphoserine phosphatase (PSPH), and phosphoserine aminotransferase (PSAT1) were all upregulated in lung TICs (Figures 2E and S2B–S2D). In particular, GLDC was one of the most highly upregulated genes in multiple analyses of lung TIC-enriched populations, at both the mRNA and protein levels (Figures 2C and S2C). GLDC is a key component of the highly conserved glycine cleavage system in amino acid metabolism that catalyzes the breakdown of glycine to form CO₂, NH₃, and 5,10-methylene-tetrahydrofolate (CH₂-THF) to fuel one-carbon metabolism (Kume et al., 1991).

**GLDC Is an Oncogene That Promotes Tumorigenesis and Cellular Transformation**

High expression of GLDC and LIN28B in lung TIC-enriched populations, but not in CD166⁺ normal lung cells, suggests that these two genes drive tumorigenicity in lung TICs. To test this hypothesis, we knocked down GLDC and LIN28B in lung tumor spheres with shRNAs (Figure S3A) and compared their growth both in vitro and in vivo. We found that both GLDC and LIN28B were necessary for cell proliferation in sphere culture, as well as anchorage-independent colony formation in soft agar (Figures 3A and S3B). Importantly, tumorigenicity was also significantly reduced upon knockdown of either GLDC or LIN28B (Figures 3B and S3C). A549 lung adenocarcinoma cells showed similar results (Figures S3D–S3G). Our results suggest that lung TICs and lung tumorigenesis are dependent on GLDC. This led us to ask what oncogenes upregulate GLDC. Because the E2F pathway upregulates many metabolic genes during cell proliferation, we examined the expression of GLDC over the course of the cell cycle in both normal human lung fibroblasts (HLFs) and the transformed A549 cells after synchronization by serum starvation. Our results showed that GLDC is insensitive to cell-cycle progression in both normal HLFs and transformed A549 cells, suggesting that GLDC is not regulated by cell-cycle or E2F signals (Figure 3C). We then examined GLDC levels in MCF10A cells after transformation by oncogenic KRASG12D, PIK3CAE545K, and MYCT58A. Our results show that all three oncogenes induce GLDC by ~20-fold, suggesting that oncogene-induced GLDC transcription is common to the cellular transformation process mediated by oncogenic Ras, PI3K, and Myc (Figure 3D).

To test whether aberrant GLDC upregulation is sufficient to drive cellular transformation, as has been shown for LIN28B (Viswanathan et al., 2009), we overexpressed GLDC in NIH/3T3 cells (Figure S3H). We found that GLDC overexpression significantly increased colony formation by 3T3 cells under normal culture conditions (Figures 3E and 3F). To test for cellular transformation in vitro, we cultured the 3T3 cells overexpressing GLDC under anchorage-independent conditions and found that GLDC transforms 3T3 cells readily with a rate exceeding that of LIN28B (Figures 3G and S3I). Upon transplantation into NOD/SCID Il2γ⁻/⁻ mice, 3T3 cells overexpressing GLDC consistently formed tumors in 6/6 transplants, and 3T3 cells overexpressing LIN28B formed tumors in 3/6 transplants, whereas 3T3 cells overexpressing the empty control vector never formed tumors (Figures S3J–S3L).

To test whether GLDC can also transform normal primary HLF and normal primary human bronchial epithelial (NHBE) cells, we overexpressed GLDC in HLF and NHBE cells (Figures S3M and S3O). Both HLF and NHBE cells showed a dramatic increase in cell proliferation upon overexpression of GLDC alone (Figures 3H–3J and S3P). Surprisingly we found that GLDC also transforms HLF and NHBE cells readily in vitro (Figures 3K and S3Q). However, perhaps because primary adult HLF and NHBE cells are not immortalized, GLDC-overexpressing HLF and NHBE cells do not form tumors upon transplantation (Figures S3N and S3R). In contrast, CD166⁺ lung tumor cells, which also could not form tumors in vivo, could now initiate tumorigenesis at a low frequency upon overexpression of GLDC (Figure 3L). Collectively, our results show that GLDC is an oncogene that is both necessary and sufficient to promote tumorigenesis.

**GLDC Promotes Tumorigenesis through Its Metabolic Activity**

Although GLDC is a metabolic enzyme, it remained unclear whether GLDC promotes tumorigenesis through a metabolism-dependent or -independent mechanism. To address this question, we engineered a series of four point mutations within or near the evolutionarily conserved catalytic active site of the GLDC enzyme to disrupt its metabolic activity (Figure 4A). These point mutations comprised three nonlethal GLDC mutations found in human patients with nonketotic hyperglycinemia (H753P, P769L, G771R; Figures S4A and S4B) and one mutation K754A that is predicted to abrogate the covalent bond with the critical pyridoxal-5'-phosphate cofactor (Nakai et al., 2005; Kure et al., 2006). When we overexpressed these four GLDC mutants in 3T3 cells, none of them could lead to tumorigenesis in vivo, whereas wild-type GLDC could, even though all of them were expressed at high levels similar to that in transformed A549 cells (Figure 4B). Thus the metabolic activity of GLDC is required for its tumorigenic function.

In addition, the upregulation of many other upstream enzymes in the glycine/serine pathway in lung TICs further supports the idea that metabolic activity in the glycine/serine pathway is responsible for promoting tumorigenesis (Figure 2E). To test this idea, we also overexpressed PSAT1, PSPH, SHMT1, SHMT2, and GCAT in 3T3 cells and transplanted them in vivo to test for cellular transformation and tumorigenesis (Figure 4C). By 2 months, we found that three other glycine/serine enzymes—PSAT1, PSPH, and SHMT2—could also transform 3T3 cells to form tumors in vivo (Figure 4D). Interestingly, we noted that PSAT1, PSPH, and SHMT2 overexpression only led to a slight upregulation of GLDC protein (Figure 4E), suggesting that their tumorigenic activity is due to increased glycine/serine metabolism, rather than indirect upregulation of GLDC. These findings indicate that increased metabolism in the glycine/serine pathway...
Figure 3. GLDC and LIN28B Are Necessary and Sufficient for Malignant Growth

(A) Proliferation curve of tumor sphere (TS) cells with shRNA knockdown of either GLDC (GD-sh) or LIN28B (28B-sh).

(B) Quantitative mass analysis of xenograft tumors formed 30 days after transplanting 100,000 tumor sphere cells with either GLDC knockdown (GD-sh) or LIN28B knockdown (28B-sh).

(C) Western blot analysis of endogenous GLDC during the cell cycle in synchronized normal human HLFs and transformed A549 cells. HLF or A549 cells were serum-starved for 72 hr followed by release into serum-containing medium with samples collected at indicated time points. Expression of GLDC, FOS (early...
due to GLDC or other glycine/serine enzymes can exert a potent tumorigenic effect.

**GLDC Regulates Glycine Metabolism, with Effects on Glycolysis and Pyrimidines**

Given that GLDC promotes tumorigenesis through a metabolism-dependent mechanism, we performed metabolomic analysis to gain deeper mechanistic insights into the GLDC-driven metabolism changes that lead to tumorigenesis. We used liquid chromatography-mass spectrometry (LC-MS) to perform metabolomics profiling of HLF cells and 3T3 cells overexpressing GLDC, as well as A549 lung adenocarcinoma cells with retroviral knockdown of GLDC, relative to empty vector controls. We found that glycine-related metabolites, glycolysis intermediates, and CDK1 (E2F target) were tested. Normal growing, unsynchronized cells (Cyc) were used as a control. HSP90 was used as a loading control.

<table>
<thead>
<tr>
<th>Group</th>
<th>No. tumors/No. injections</th>
</tr>
</thead>
<tbody>
<tr>
<td>3T3-GLDC</td>
<td>6/6</td>
</tr>
<tr>
<td>3T3-PSAT1</td>
<td>3/6</td>
</tr>
<tr>
<td>3T3-PSPH</td>
<td>2/6</td>
</tr>
<tr>
<td>3T3-SHMT2</td>
<td>1/6</td>
</tr>
<tr>
<td>3T3-SHMT1</td>
<td>0/6</td>
</tr>
<tr>
<td>3T3-GCAT</td>
<td>0/6</td>
</tr>
<tr>
<td>3T3</td>
<td>0/6</td>
</tr>
</tbody>
</table>

In all panels, error bars represent SEM. *p < 0.05, **p < 0.01. See also Figure S4.
and many pyrimidines were significantly perturbed by both GLDC overexpression and knockdown (p < 0.05; Figures 5A–5D). For glycine-related metabolites, we found that sarcosine (N-methylglycine) levels increased significantly upon GLDC overexpression and dropped significantly upon GLDC knockdown, indicating that GLDC promotes sarcosine synthesis or accumulation (Figure 5A). Consistent with this observation, betaine aldehyde in the betaine-sarcosine-glycine pathway for glycine synthesis also showed the same pattern of changes (Figure 5A). Glycine levels decrease with GLDC overexpression and increase with GLDC knockdown, in agreement with the fact that GLDC breaks down glycine irreversibly. In contrast, serine levels increase with GLDC overexpression and decrease with GLDC knockdown, suggesting that GLDC promotes serine synthesis or uptake (Figure 5A).

Surprisingly, GLDC perturbation also led to dramatic changes in glycolysis and other amino acids (Figures 5B, 5C, and 5S). Our data suggest that GLDC promotes glycolysis, leading to the increased synthesis or accumulation of glucose-1-phosphate, phosphoenolpyruvate, pyruvate, and lactate (Figures 5B and 5C). In fact many of the upstream glycine/serine metabolism enzymes that we found upregulated in lung TICs, such as PSAT1, PSPH, and SHMT1/2, channel glycolytic intermediates into de novo serine and glycine biosynthesis (Figure 2E), suggesting that GLDC is working in a concerted fashion with these enzymes to promote the glycolysis-serine-glycine flux. This is supported by our finding that GLDC does not significantly promote glycine uptake (Figures 5SB and 5SC) but promotes glycine uptake instead (Figures 5B and 5C).

Finally, our metabolomics analysis also revealed that GLDC promotes the synthesis or accumulation of pyrimidines, including thymidine, deoxyuridine, thymine, uracil, and cytosine (Figure 5D). The GLDC-catalyzed reaction converts glycine into CH₃-THF (Kume et al., 1991). CH₃-THF contains the methyl group that fuels de novo thymidine synthesis from deoxyuridine in concert with pyrimidine biosynthesis and hence nucleotide synthesis during cell proliferation (Tibbetts and Appling, 2010). Recent studies suggest that early oncogenesis involves aberrant activation of cell proliferation, which then leads to a crisis of nucleotide deficiency and replication stress (Bester et al., 2011). Our observations on pyrimidine synthesis suggest that upregulation of GLDC could promote cellular transformation by overcoming this deficiency to progress onward in early oncogenesis.

To test whether any of the metabolite changes induced by GLDC can mimic GLDC’s effects on cancer cells, we analyzed whether an increased exogenous supply of specific metabolites could rescue GLDC retroviral knockdown in A549 cells. We found that 10 μM of sarcosine could significantly rescue the proliferation defect upon GLDC knockdown, with little effect on control A549 cells (Figure 5F), indicating that increased sarcosine-glycine metabolism flux can rescue the effects of reduced GLDC enzyme. To further test whether the production of CH₃-THF is necessary for GLDC’s effects on proliferation, we tested whether the antifolate drug methotrexate could specifically abrogate GLDC-induced proliferation by reducing the tetrahydrofolate (THF) cofactor needed to produce CH₃-THF for pyrimidine synthesis. Our results show that low doses of methotrexate specifically abrogated GLDC-induced proliferation in 3T3 and HLF cells, with little effect on control 3T3 and HLF cells (Figure 5E). Furthermore, methotrexate in combination with GLDC shRNA killed transformed A549 cells much more effectively than either alone (Figure 5E), suggesting that a combination of antifolates with a GLDC inhibitor could completely shut off glycine catabolism to treat cancer cells more effectively. Using these metabolic data, we constructed a model of how aberrant GLDC expression might reprogram glycolysis and glycine metabolism fluxes in cancer cells to promote cancer cell proliferation and tumorigenesis (Figure 5G).

**Prognostic Significance of Aberrant GLDC Expression in NSCLC Patients**
To assess whether our experimental findings on GLDC are relevant to human lung cancer patients in the clinic, we used tissue microarray immunohistochemistry to examine the prognostic significance of GLDC expression, tumor size, tumor grade, and cancer stage in clinical tumor samples from cohorts of NSCLC patients (n = 143) (Figures 6A and S6; Table S2). Subdistribution hazard ratio (SHR) analysis showed that patients with high or grade 3+ GLDC expression have a 3-fold higher risk of lung cancer mortality compared to patients with low or grade 0 GLDC expression, even when adjusted for cancer stage (SHR = 3.01, 95% confidence interval [CI]; 1.48–6.10, p = 0.002) (Figure 6B). Cumulative mortality analysis also showed that high GLDC expression (grade 3+) is significantly associated with higher cumulative incidence of mortality across 143 NSCLC lung cancer patients, even when adjusted for cancer stage (p = 0.005) (Figure 6C). CD166 expression was not significantly associated with higher mortality in lung cancer patients—which was not unexpected given that only 1 in 5 × 10⁶ CD166⁺ cells are tumorigenic (Figures S6A–S6C). Indeed, common staining of lung tumors revealed that GLDC⁺ cells mostly form a subset of CD166⁺ cells, and that not all CD166⁺ cells are GLDC⁺ (Figures 6D and S6D). LIN28B immunohistochemistry staining was also not significantly correlated with lung cancer patient mortality (Figures S6A–S6C), although western blots revealed that lung TICs express a second LIN28B isoform that is indiscernible from immunohistochemistry staining, thus rendering our LIN28B staining results inconclusive (Figure S6E). Our immunohistochemistry results on clinical tumor samples are consistent with the idea that lung TICs constitute the bulk of the tumors in late stages of malignancy and demonstrate that aberrant activation of GLDC is significantly associated with human mortality in NSCLC patients—further supporting its role as a metabolic oncogene in human NSCLC.

**Aberrant GLDC Expression in Other Cancers**
To check whether aberrant GLDC expression is specific to NSCLC, we examined a variety of other cancers. Surprisingly GLDC is also aberrantly upregulated in subsets of primary tumors from other cancers, especially ovarian and germ cell tumors (Figure 7A; Table S3). Further analysis of 606 human cancer cell lines showed that 158 (26.1%) cancer cell lines overexpress GLDC, including lines from ovarian, germ cell, cervical, lung, lymphoma, prostate, bladder, and colon cancer (Figure 7B; Table S4). To test whether GLDC is also required for growth by one of these GLDC-overexpressing cell lines, we knocked
Figure 5. Metabolomics of Cells upon GLDC Overexpression and Knockdown

(A–D) Relative fold change in levels of (A) glycine-related metabolites, (B) glycolysis intermediates, and (D) pyrimidines in 3T3 cells with GLDC overexpression (3T3-GD/Ctrl), HLF cells with GLDC overexpression (HLF-GD/Ctrl), and A549 cells with GLDC knockdown (A549-GD-sh/Ctrl), as determined by LC-MS metabolomics. (C) Lactate production by 3T3 cells with GLDC overexpression or A549 cells with GLDC knockdown.

(E) Effects of the antifolate drug methotrexate on colony formation after GLDC overexpression or knockdown. 3T3 and HLF cells overexpressing GLDC were plated at clonal density and exposed to varying concentrations of methotrexate for 8 days. A549 cells with GLDC knockdown were plated in soft agar at clonal density and exposed to varying concentrations of methotrexate for 14 days.

(F) Effects of sarcosine on soft agar colony formation after GLDC knockdown in A549 cells. One thousand cells were seeded in soft agar at clonal density and exposed to 10 μM sarcosine for 14 days.

(G) Model of metabolic flux changes induced by GLDC.

In all panels, error bars represent SEM. See also Figure S5.
Figure 6. GLDC Is a Prognostic Indicator for Mortality in NSCLC Patients

(A) GLDC immunohistochemistry staining in a NSCLC tumor microarray (n = 143). Representative images shown for human primary lung adenocarcinomas (AdC) immunostained with GLDC. Staining intensity grade is indicated in the upper right corner. The boxed regions in the upper images are shown at higher magnification in the lower images. Scale bar, 100 μm.

(B) Subdistribution hazard ratios for each GLDC staining intensity grade, adjusted for American Joint Committee on Cancer (AJCC) staging. CI, confidence interval.

(C) Cumulative incidence of lung cancer mortality adjusted for AJCC staging, for patients with each GLDC staining intensity grade.

(D) Coimmunofluorescence staining of CD166 (red) and GLDC (green) on primary lung cancer patient tumors, counterstained with DAPI (blue). Representative cases with coexpression of high levels of CD166 and high levels of GLDC (left panel) and low levels of CD166 and low levels of GLDC (right panel) are shown. Higher magnification inset is shown in bottom left corner. Scale bar, 50 μm.

See also Figure S6 and Table S2.
Indeed we found that GLDC knockdown reduced their proliferation and tumorigenic potential upon transplantation, suggesting that GLDC may act as an oncogene in other cancer cells as well (Figures 7C–7E). To examine the possibility that GLDC is a housekeeping gene for cell proliferation, we also knocked down GLDC in normal HLFs (Figure S7A). We found that HLF proliferation was unaffected by retroviral knockdown of GLDC (Figures S7 B and S7C). Furthermore we observed that GLDC is highly expressed only in a few normal tissues, including postmitotic liver cells, kidney cells, placenta cells, and olfactory bulb neurons (Figure S7 D). Altogether our observations in both experimental and clinical settings suggest that human GLDC is not a housekeeping gene required for cell proliferation but rather an oncogenic metabolic enzyme aberrantly upregulated in NSCLC and possibly several other human cancers.

**DISCUSSION**

**TIC State in Lung Cancer**

Our work sheds new light on the nature of the TIC state and the role of metabolic reprogramming in tumorigenesis. In this study, we isolated a subpopulation of TICs from NSCLC patients using the marker CD166 and showed that both the oncogenic stem cell
factor LIN28B and the glycine metabolism enzyme GLDC drive the tumorigenicity of lung cancer TICs.

Our data showed that CD166 enriched for TICs in primary NSCLC, and that CD166 served as an inert surface marker. In contrast, our results on CD133 are different from the results reported by Eramo et al. (2008) even though both studies used the same CD133 antibody. This is most likely due to differences in the xenotransplantation assays, which tend to underestimate the true frequency of TICs. We employed a more sensitive mouse xenotransplantation assay using NOD/SCID il2rg−/− mice instead of SCID mice, and we directly transplanted primary tumor cells with Matrigel instead of expanding the tumor cells in vitro. Previous studies have demonstrated that using a more sensitive mouse xenotransplantation assay dramatically improves our understanding of TICs (Quintana et al., 2008). Our present study supports this notion, leading us to CD166 as a new marker for the lung TIC-containing fraction. In normal physiology, CD166 is expressed predominantly during embryonic development, including the embryonic upper airway, primary cardiac cells, and mesenchymal stem cells (Avril-Delplanque et al., 2005; Murakami et al., 2007; Hennrick et al., 2007; Sabatini et al., 2005). Expression of CD166 in the embryonic lung is consistent with our observation that CD166+ lung TICs express high levels of embryonic lung transcription factors like PEA3 and the trachealess homolog NPAS1, as well as the oncogenic stem cell factor LIN28B (Liu et al., 2003; Levesque et al., 2007; Viswanathan et al., 2009). Interestingly, mouse Lin28 is also expressed in the embryonic lung during normal development (Yang and Moss, 2003). These observations suggest that the TIC state in lung cancer is similar to the embryonic lung progenitor state in many aspects.

**GLDC Is a Metabolic Oncogene**

Our results demonstrate that multiple components in the glycine-serine pathway are also oncogenes. In addition to embryonic lung factors, lung TICs also express high levels of GLDC, GCAT, SHMT1/2, PSPH, and PSAT1, suggesting that TICs rely on glycine-serine metabolism for tumorigenesis. Overexpression of catalytically active GLDC, as well as PSAT1, PSPH, and SHMT2, could induce cellular transformation in 3T3 cells to form tumors, whereas retroviral knockdown of GLDC significantly reduces the tumorigenicity of lung cancer cells. We further observed that GLDC+ cells mostly form a subset of CD166+ cells in lung tumors.

PSAT1, PSPH, and SHMT1/2 lie upstream of GLDC in the glycine-serine pathway, diverting glycolytic flux from 3-phosphoglycerate through serine to glycine. GLDC is an oxidoreductase that catalyzes the irreversible rate-limiting step of glycine catabolism, by breaking down each glycine molecule in the glycine cleavage system to produce NADH, CO2, NH3, and CH2-THF (Kume et al., 1991). CH2-THF fuels the one-carbon/fo late metabolism pool, which in turn supplies methylene groups for biosynthesis (Tibbetts and Appling, 2010). Consistent with these facts, we found that GLDC regulates many metabolites in glycolysis and the glycine-serine pathway, leading to specific changes in pyrimidine synthesis. Pyrimidine derivatives like thymidine, in turn, are required for nucleotide synthesis in cell proliferation. Recent studies suggest that early oncogenesis involves aberrant activation of cell proliferation, which then leads to a crisis of nucleotide deficiency and replication stress (Bester et al., 2011)—a crisis that GLDC upregulation could overcome for continued progression in tumorigenesis. Interestingly, we found that GLDC also increases the levels of N-methylglycine or sarcosine, an oncometabolite implicated in prostate cancer (Sreekumar et al., 2009). Furthermore, we observed that GLDC promotes glycolysis. Combined with our findings on LIN28, which has been shown to promote glucose uptake and glycolysis (Zhu et al., 2010, 2011), GLDC might be cooperating with LIN28 as well as PSAT1, PSPH, and SHMT1/2 to divert the glycolytic flux to glycine and produce CH2-THF. These observations support the notion that the Warburg effect promotes biosynthesis for tumorigenesis (Hsu and Sabatini, 2008; Vander Heiden et al., 2009).

**GLDC and Glycine Metabolism Are Relevant to Human Cancer Patients**

From the prognostic perspective, aberrant GLDC expression is significantly correlated with the survival rates of NSCLC patients. This is consistent with the model that TIC clones expand to constitute the bulk of the tumor in advanced stages of malignancy (Boiko et al., 2010). Aberrantly increased GLDC is also widespread in many other human cancers, including lymphoma, ovarian, germ cell, cervical, prostate, bladder, and colon cancer, whereas most normal adult human tissues express very low levels of GLDC. Our experimental data further suggest that in cancers that rely on GLDC and glycine metabolism, the highly toxic antifolate drug methotrexate might be initially effective because it targets TICs, although our data suggest an even more effective chemotherapy could be potentially achieved by combining an antifolate drug with a GLDC inhibitor or by searching for a glycine cleavage complex-specific antifolate drug—much like the search for kinase-specific inhibitors in targeted cancer therapy.

Our study links a glycine metabolism enzyme to lung cancer and tumorigenesis. Recently several metabolic enzymes have been linked to cancer in patients, supporting the status of metabolic reprogramming as a new hallmark of cancer (Hanahan and Weinberg, 2011). In particular, the pyruvate kinase M2 isoform PKM2, isocitrate dehydrogenase IDH1/2, and phosphoglycerate dehydrogenase PHGDH have been implicated in multiple cancers (Christofk et al., 2008; Parsons et al., 2008; Dang et al., 2009; Locasale et al., 2011; Possemato et al., 2011). Regardless of the controversy over the frequency of TICs at different stages of malignancy, our approach shows that characterizing the unique molecular basis that defines cancer cells with tumorigenic capacity may nevertheless provide novel drug targets for advancing cancer therapy.

**EXPERIMENTAL PROCEDURES**

**Tumor Cell Preparation**

NSCLC tumors were collected from patients according to protocols approved by the Ethics Committee of the National University of Singapore. Samples were washed, dissociated, and incubated in DNase and collagenase/dispsase. After incubation, cell clusters and red blood cells were removed. Then single cells were resuspended and ready for transplantation. See the Extended Experimental Procedures for more details.
Flow Cytometry
A list of antibodies used can be found in Table S5. Cells were FACS-sorted using a FACS Aria II (BD). Flow cytometry was performed using a LSR II flow cytometer, and data were analyzed with CELLQuest Pro software (BD).

Animals and Transplantation of Tumor Cells
NOD.Cg-Fkrl<ko> Il2rg<ko/ko>/SzJ mice (Jackson Laboratories) at 4–6 weeks old were subcutaneously transplanted with single-cell suspensions in serum-free medium and Matrigel (BD) (1:1).

Tumor Sphere Culture
Cells were grown in DMEM/F12 containing ITS (BD Biosciences) and supplemented with 50 ng/ml EGF and 20 ng/ml basic fibroblast growth factor (bFGF) (Invitrogen), using nontreated cell culture plates (Nunc). Fresh medium was replenished every 3 days.

cDNA Microarray Analysis
Total RNA was extracted by Trizol (Invitrogen) and purified by RNeasy Mini Kit (Qiagen). Lung primary tumors (one patient), xenografts (three patients), tumor spheres (four patients), and normal human adult lung tissues (three patients) were used. RNA was processed and hybridized to HumanRef-8 v3.0 Beadarrays (Illumina), and the microarray data were normalized and analyzed as described previously (Chua et al., 2006). A fold-change cut-off threshold of 1.5 was applied to generate the lung TIC gene signature after four comparisons: primary tumor CD166+/versus CD166- (P+/P-), xenograft tumor CD166+ versus CD166- (X+/X-), spheres versus xenograft tumor CD166+ (S/X+), and normal lung CD166+ versus CD166- (N+/N-). After intersecting the differentially expressed genes (DEGs) of P+/P-, X+/X-, and S/X+ and excluding DEGs intersecting with N+/N-, DAVID Bioinformatics Resources 6.7 was applied for KEGG pathway analysis of the final list of DEGs (Huang et al., 2009).

Metabolomics
Metabolites were extracted by centrifugation of culture media at 14,000 rpm for 30 min at 4 °C. Metabolomic profiling was performed through UPLC/MS using a Zorbax Eclipse Plus-C18 column on the Agilent 1200 RRLC and an Agilent 6530 Accurate Mass QTOF. Mass spectrometry was performed on an Agilent 6530 Accurate Mass Q-TOF mass spectrometer operating in positive ion mode with 2 GHz extended dynamic range mode. See Extended Experimental Procedures for more details.

Statistical Analysis
Differences were compared using two-tailed Student’s t test. p values < 0.05 were considered statistically significant. All analyses were performed with SPSS 18.0 (SPSS). Lung TIC frequencies were estimated using ELDA software (Hsu and Smyth, 2009). Fisher’s exact test was used to assess the association between GLDC, CD166, or LIN28B and clinicopathological parameters. The effect of GLDC, CD166, or LIN28B expressions on lung cancer mortality was modeled using competing risks regression and quantified based on the SHR (Fine and Gray, 1999).

ACCESSION NUMBERS
The GEO accession number for human datasets is GSE33198.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Extended Experimental Procedures, seven figures, and six tables and can be found with this article online at doi:10.1016/j.cell.2011.11.050.

ACKNOWLEDGMENTS
This work is supported by grants from Biomedical Research Council (BMRC) and Agency for Science, Technology and Research (A*STAR) (B.L.) and grants from the Singapore Cancer Syndicate and Singapore Lung Cancer Consortium (E.H.L.). We gratefully acknowledge assistance from Y.P.R. Yip, S.L. Long, and J. Xu for collection of lung tissues. We are grateful to the Biopolis Shared Facilities Histopathology Laboratory staff for their support with immunohistochemistry and image analysis. We thank Z.M. Li for assistance with making transformed breast cell lines. We greatly acknowledge V. Gaddemane, C.S. Gan, and T. Hennessy from Agilent Technologies, Singapore for their support in acquiring and analyzing the mass spectrometry data for the differential analysis of the metabolites. We thank W.L. Tam and F. McKeon for comments on the manuscript.

REFERENCES


