APPENDIX

Supplement Part I: HSQC metabolite identification and quantification

Part I of this Supplement presents the 2D NMR data collected and analyzed from the MCF10 tumor progression model. A strength of this methodology is that we are able to identify and quantify a number of metabolites in each sample without prior separation. Thus peak identification in the HSQC spectra is key as analysis depends upon extracting accurate signal volume and multiplet intensities from a spectrum of a mixture. Table S1 lists the carbons observed for each metabolite and the corresponding multiplets and shifts in the proton and carbon dimensions. Table S2 shows the cross peak integration for the ^{13}C multiplet components. Table S3 displays the relative pool sizes of thirteen metabolites which were not explicitly presented in the main text.

The assignments of metabolites in the HSQC spectra were made according to published data, by comparison with spectra of authentic compounds, by spiking the samples and by observing the ^{13}C-^{13}C scalar coupling fine structures. Amino acids alanine, arginine, glycine, isoleucine, leucine, lysine, proline, and valine as well as lactate were readily identified. The cross peaks for succinate, taurine, *myo*-inositol, creatine and phosphocreatine were verified by recording the spectra of individual compounds. Resonances of creatine and phosphocreatine can not be distinguished and thus represent the total creatine pool. Resonances of γ-glutamyl-glutathione (GSH), glutamine, choline, phosphocholine, glycerophosphocholine, *N*-acetyl-glucosamine (GlcNAc), *N*-acetyl-galactosamine (GalNAc), UDP-GlcNAc, and UDP-GalNAc were assigned using the

spiking method. The cross peaks for the acetyl group of UDP-GlcNAc and UDP-GalNAc are overlapped, and the acetyl moiety of GlcNAc and GalNAc is not resolved.

Table S1. ^{1}H and ^{13}C resonances from metabolites observed by HSCQ

Serial no.	Metabolites	Carbon position	^{1}H multiplet	^{13}C multiplet	^{1}H chemical shift (ppm)	^{13}C chemical shift (ppm)
Amino acids						
1	Alanine	C2	q	m	3.79	53.24
		C3	d	m	1.49	18.89
2	Arginine	C5	t	s	3.26	43.26
3	GSH (Glu)	C3	m	m	2.18	28.91
		C4	t	m	2.57	34.14
	GSH (Cys)	C2	t	s	4.58	58.40
		C3	m	s	2.97	28.25
	GSH (Gly)	C2	s	s / m	3.79	46.10
4	Glutamate	C2	t	m	3.79	56.90
		C3	m	m	2.07, 2.13	29.58
		C4	t	m	2.36	36.22
5	Glutamine	C4	m	m	2.46	33.56
6	Glycine	C2	s	s / m	3.57	44.20
7	Isoleucine	C4-H_3	d	s	1.02	17.39
		C5-H_3	t	s	0.95	13.83
8	Leucine	C5	d	s	0.98	24.75
		C5′	d	s	0.97	23.67

9	Lysine	C5	m	s	1.74	29.12
		C6	t	s	3.02	41.95
10	Proline	C4	m	s / m	2.02	26.49
		C5	m	s / m	3.37, 3.44	48.85
11	Valine	C4	d	s	1.05	20.68
		C4′	d	s	1.00	19.39
Organic acids						
12	Lactate	CH	q	m	4.12	71.22
		CH_3	d	m	1.33	22.78
13	Succinate	CH_2-CH_2	s	m	2.42	36.90
Sugars						
14	*m*-Inositol	C1H, C3H	dd	s	3.55	73.87
		C2H	t	s	4.07	74.92
		C4H, C6H	dd	s	3.63	75.15
		C5H	t	s	3.29	77.09
15	GlcNAc / GalNAc	CH_3	d	m	2.01	24.62
16	UDP-GlcNAc /UDP-GalNAc	CH_3	d	m	2.08	24.83
Membrane components						
17	Choline	CH_2OH	m	s	4.07	58.35
		NCH_2	m	s	3.52	70.17
18	Phosphocholine	CH_2-PO_3^-	m	s	4.18	60.79
		NCH_2	m	s	3.60	69.23
		$N(CH_3)_3$	s	s	3.23	56.75

19	Glycero-phosphocholine	CH_2-PO_3^-	m	s	4.33	62.17
		NCH_2	m	s	3.68	68.74
		CH_2-O-P	m	m	3.89, 3.95	69.25
		CHOH	m	m	3.91	73.38
		CH_2OH	m	m	3.62, 3.68	64.76
Nucleotides						
20	UTP / UDP	C1H	d	m	6.00	91.14
		C2H	t	m	4.39	76.49
		C3H	t	m	4.38	72.39
		C4H	t	m	4.30	85.96
		$C5H_2$	m	m	4.24	67.89
Other compounds						
21	Creatine	CH_2	s	s	3.94	56.63
		CH_3	s	s	3.05	39.69
22	Taurine	N-CH_2	t	s	3.44	38.15
		S-CH_2	t	s	3.28	50.23

GSH, glutathione; GlcNAc, *N*-acetylglucosamine; GalNAc, *N*-acetylgalactosamine; s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet.

^{13}C multiplets were obtained based on [U-13C]glucose labeling and 2D [^{13}C, ^{1}H] HSQC spectroscopy.

Table S2. Relative intensities of ^{13}C multiplet components for metabolites observed by HSCQ

Carbon position	Isotopomer populations	Multiplet components	MCF-10A	MCF-10AT	MCF-10AT1	MCF-10CA1a
Alanine-C2	2-^{13}C	s	0.27	0.20	0.19	0.15
	2,3-$^{13}C_2$	d	0.01	0.13	0.15	0.12
	1,2-$^{13}C_2$	d*	0.01	0.01	0.02	0.01
	1,2,3-$^{13}C_3$	dd	0.71	0.66	0.64	0.72
Alanine-C3	3-^{13}C	s	0.28	0.21	0.22	0.18
	2,3-$^{13}C_2$	d	0.72	0.79	0.78	0.82
Lactate-C3	3-^{13}C	s	0.16	0.21	0.22	0.14
	2,3-$^{13}C_2$	d	0.84	0.79	0.78	0.86
Acetyl-CoA (GlcNAc/GalNAc)-C2	2-^{13}C	s	0.29	0.23	0.22	0.24
	1,2-$^{13}C_2$	d	0.71	0.77	0.78	0.76
Glutamine-C4	4-^{13}C	s	0.50	0.29	0.28	0.29
	3,4-$^{13}C_2$	d	0.01	0.01	0.01	0.01
	4,5-$^{13}C_2$	d*	0.48	0.67	0.68	0.67
	3,4,5-$^{13}C_3$	dd	0.01	0.03	0.03	0.03
Glutamate-C3	3-^{13}C	s	0.73	0.65	0.64	0.59
	2,3-$^{13}C_2$ / 3,4-$^{13}C_2$	d	0.27	0.35	0.35	0.40
	2,3,4-$^{13}C_3$	t	0	0	0.01	0.01
Glutamate-C4	4-^{13}C	s	0.50	0.21	0.32	0.30
	3,4-$^{13}C_2$	d	0.01	0.01	0.01	0.01
	4,5-$^{13}C_2$	d*	0.48	0.72	0.64	0.66

	3,4,5-$^{13}C_3$	dd	0.01	0.06	0.03	0.03
Glu (GSH)-C3	3-^{13}C	s	0.67	0.58	0.55	0.53
	2,3-$^{13}C_2$ / 3,4-$^{13}C_2$	d	0.32	0.41	0.41	0.46
	2,3,4-$^{13}C_3$	t	0.01	0.01	0.04	0.01
Glu (GSH)-C4	4-^{13}C	s	0.24	0.12	0.13	0.14
	3,4-$^{13}C_2$	d	0.02	0.01	0.01	0.02
	4,5-$^{13}C_2$	d*	0.70	0.77	0.76	0.75
	3,4,5-$^{13}C_3$	dd	0.04	0.10	0.10	0.09
Gly (GSH)-C2	2-^{13}C	s	0.88	0.58	0.48	0.44
	1,2-$^{13}C_2$	d	0.12	0.42	0.52	0.56
Glycine-C2	2-^{13}C	s	0.86	0.57	0.48	0.44
	1,2-$^{13}C_2$	d	0.14	0.43	0.52	0.56
Proline-C4	4-^{13}C	s	1.00	1.00	1.00	0.42
	4,5-$^{13}C_2$	d	0.00	0.00	0.00	0.56
	3,4,5-$^{13}C_3$	t	0.00	0.00	0.00	0.02
Proline-C5	5-^{13}C	s	1.00	1.00	1.00	0.39
	4,5-$^{13}C_2$	d	0.00	0.00	0.00	0.61
UDP-C2	2-^{13}C	s	0.24	0.16	0.15	0.12
	1,2-$^{13}C_2$	d	0.62	0.58	0.59	0.59
	1,2,3-$^{13}C_3$	t	0.14	0.26	0.26	0.29

Table S3. Relative pool sizes of various metabolites[a]

Metabolites	Ratio MCF-10AT / MCF-10A	Ratio MCF-10AT1 / MCF-10A	Ratio MCF-10CA1a / MCF-10A
(Glc + Gal)NAc	> 3.74 **	> 5.02 **	> 3.95 **
UDP-(Glc + Gal)NAc	3.60 ± 0.89 **	3.29 ± 0.72 **	2.55 ± 0.54 **
m-Inositol	0.61 ± 0.08 *	0.71 ± 0.09 *	0.38 ± 0.04 ***
Free choline	0.54 ± 0.09 *	0.32 ± 0.05 **	< 0.07 ***
Phosphocholine	0.34 ± 0.05 **	0.26 ± 0.03 ***	0.23 ± 0.03 ***
Glycerophosphocholine	< 0.07 ***	< 0.07 ***	< 0.03 ***
Arginine	1.12 ± 0.17	0.84 ± 0.13	1.25 ± 0.18
Isoleucine	0.88 ± 0.15	0.71 ± 0.12	0.85 ± 0.10
Leucine	1.09 ± 0.11	0.76 ± 0.10	1.04 ± 0.10
Lysine	1.04 ± 0.23	0.69 ± 0.12	0.96 ± 0.15
Valine	0.91 ± 0.13	0.66 ± 0.12	0.83 ± 0.12
Creatine	0.47 ± 0.04 **	0.49 ± 0.05 **	0.64 ± 0.06 **
Taurine	0.31 ± 0.04 ***	0.26 ± 0.03 ***	0.14 ± 0.02 ***

[a] Relative amount of the compounds were obtained by normalizing the peak size to the response of internal reference compounds, and further, to 1 mg of total protein amount, and then by comparing with the results of MCF-10A sample. Quantitation based on natural isotopic abundance (a direct measure of metabolite concentrations). Values represent mean ± SD (n=4); * indicates $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ for comparison to MCF10-A.

Supplement Part II: Isotopomer analysis and modeling

^{13}C tracer experiments and NMR spectroscopy provide a wealth of intuitive information on the cell intermediary metabolism. To account rigorously for this information, development of a mathematical model of cell metabolism is required. Comprehensive isotopomer models have been widely used for analysis of the metabolism in microorganism. Nevertheless, few models have been constructed for the metabolism in human cells or tissues. Here a mathematical model of human cell intermediary metabolism is developed. It describes the distribution of the ^{13}C isotopomers of metabolites in cells fed with ^{13}C-labeled substrates. The model allows the determination of fluxes through different metabolic pathways using the ^{13}C multiplet pattern of various metabolites obtained from HSQC or ^{13}C-NMR spectroscopy. The considered metabolic network includes glycolysis, the pentose phosphate pathway, the tricarboxylic acid cycle, anaplerotic reaction, and glycine and proline synthesis as follows. The model was used for calculating metabolic fluxes or flux ratios in a progression series of breast cancer cells, incubated with [U-$^{13}C_6$]glucose.

Contribution of Pentose Phosphate Pathway to Pyruvate and Ribose formation

The relative flux of glucose carbons through glycolysis and the PPP to pyruvate relies on the analysis of ^{13}C multiplets of alanine C2. The observed relative multiplet intensities were transformed to the relative abundances of intact carbon fragments. According to the carbon rearrangements that govern the PP pathway, the interconversion of three molecules pentose composed of intact C1-C2-C3-C4-C5 fragments eventually yields five molecules of pyruvate, three of which retain an intact C1-C2-C3 fragment, while two of

them carry a C2-C3 fragment originating from a single glucose molecule but the C1-C2 carbon bonds have been cleaved. The latter fragments are denoted as fraction $f^{(2)}$\{Ala-C2\}. The fraction of pyruvate derived from the PP pathway can then be estimated from the expression 5/2 * $f^{(2)}$\{Ala-C2\}. The fraction $f^{(2)}$\{Ala-C2\}is the main contributor to the isotopomer population of [2,3-$^{13}C_2$]alanine, which is observed as the contribution of doublet ($^1J_{CC}$ = 35 Hz) to the total multiplets of alanine C2.

Ribose for nucleotides is synthesized directly from ribose-5-phosphate (R5P). R5P can be synthesized via both the oxidative or non-oxidative branches of the PPP. To determine the relative contributions of oxidative and non-oxidative PPP to R5P production, the ^{13}C scalar coupling mulitplets of C2 of UTP or UDP were used to calculate the relative abundances of intact carbon fragments in R5P. R5P formation through oxidation of glucose-6-phosphate yields intact C_5 fragments that is represented by $f^{(3)}$(UTP/UDP-C2). However, if R5P is synthesized via the non-oxidative steps of PPP, the reactions catalyzed by the reversible transketolase and transaldolase (EC 2.2.1.2) lead to the cleavage of C2-C3 and C1-C2 connectivities in R5P, respectively. Therefore, the relative contribution of non-oxidative PPP to R5P production can be estimated from the sum of fractions $f^{(1)}$ and $f^{(2)}$(UTP/UDP-C2).

Anaplerotic Flux and TCA Cycle Flux

In the MCF10 cell lines, there are two points-of-entry for pyruvate into the TCA cycle, pyruvate dehydrogenase (PDH) and pyruvate carboxylase (PC). (Entry via malic enzyme is also possible but not observed in our system.) The entry of [U-$^{13}C_3$]pyruvate into the

TCA cycle through PDH or PC generates different labeling patterns of α-ketoglutarate. The relative abundances of these patterns are observed in glutamate labeling, a relatively abundant compound in rapid equilibrium with the low abundance α-ketoglutarate. TCA cycle replenishment by [1,2-$^{13}C_2$]acetyl-CoA with the activity of pyruvate dehydrogenase generates [4,5-$^{13}C_2$]α-ketoglutarate. Consequently, the isotopomer population of [4,5-$^{13}C_2$]glutamate reflects the flux through pyruvate dehydrogenase, which equals the TCA cycle (citrate synthase) flux when the acetyl-CoA synthetase flux is zero.

However, if pyruvate enters the TCA cycle via pyruvate carboxylase then U-$^{13}C_3$]pyruvate and unlabeled CO_2 generate two ^{13}C isotopomers of oxalacetate, [1,2,3-$^{13}C_3$] and [2,3,4-$^{13}C_3$]oxaloacetate. Two isotopomers are formed because oxaloacetate is reversibly converted into symmetric succinate (or fumarate). Thus [1,2,3-$^{13}C_3$] and [2,3-$^{13}C_2$]glutamate is formed due to the activity of pyruvate carboxylase. The relative activity of pyruvate carboxylase versus pyruvate dehydrogenase was calculated from the ^{13}C multiplet components of glutamate at carbon positions C3 and C4: d{Glu-C3} / (d*{Glu-C4} + dd{Glu-C4}), where d{Glu-C3} is the contribution of doublet ($^1J_{CC}$ = 35 Hz) to the total multiplets of glutamate C3, and d*{Glu-C4} and dd{Glu-C4} are the relative contributions of the doublet with $^1J_{CC}$ coupling constant of 55 Hz and the doublet-of-doublets to the glutamate C4 multiplets, respectively.

Based on the analysis of the ^{13}C multiplets of acetyl-CoA C2, glutamate C4, and C4 of γ-glutamyl moiety of glutathione, we found that the C4 and C5 of γ-glutamyl of glutathione are solely derived from acetyl-CoA, whereas an isotopic dilution originating from a non-

enriched carbon source occured to glutamate. Hence, the ^{13}C multiplet pattern of γ-glutamyl moiety of glutathione was used to determine the relative activity of the anaplerotic pathway and the TCA cycle. In metabolic steady-state, the anaplerotic flux equals the flux of removal of TCA cycle intermediates towards cell synthesis.

Glycine Biosynthesis

Glycine could be synthesized via serine from an intermediate of glycolysis, 3-phosphoglycerate, or obtained directly from media components. Due to the absence of the isotopomer population of [1,2-$^{13}C_2$]alanine, we determined that no backflow from the TCA cycle to glycolysis (*e.g.* the conversion of malate to pyruvate catalyzed by malic enzyme) occured. Therefore all the C_3 intermediates of glycolysis have the same isotopomer distribution and thus the labeling state of 3-phosphoglycerate is equivalent to that of pyruvate, which is assessed via alanine.

We obtain the isotopomer balance equation (SE1).

$$P_{Gly-C2} \cdot \begin{bmatrix} s \\ d \end{bmatrix}_{Gly-C2} = X^{syn} \cdot P_{Ala-C2} \cdot \begin{bmatrix} s+d \\ d*+q \end{bmatrix}_{Ala-C2} + (1-X^{syn}) \cdot P_n \cdot \begin{bmatrix} 1-P_n \\ P_n \end{bmatrix} \qquad \text{(SE1)}$$

In this equation, X^{syn} represents the fraction of glycine derived from glycolysis, *s*, *d*, *d**, and *q* correspond to the relative contributions of singlet, doublet, doublet with a larger coupling constant and quartet, respectively. P_n is the natural ^{13}C abundance ($P_n = 0.011$), and P_{Gly-C2} and P_{Ala-C2} are the specific enrichments of glycine C2 and alanine C2. P_{Gly-C2} can be calculated from X^{syn}, P_n and P_{Ala-C2} using the relation $P_{Gly-C2} = X^{syn} \cdot P_{Ala-C2} + (1 -$

X^{syn}) · P_n, thus X^{syn} is determined from the ^{13}C multiplets of alanine C2 and glycine C2 with Eq. (SE2).

$$X^{syn} = \frac{P_n \cdot \left(\begin{bmatrix} s \\ d \end{bmatrix}_{Gly-C2} - \begin{bmatrix} 1-P_n \\ P_n \end{bmatrix} \right)}{P_{Ala-C2} \cdot \left(\begin{bmatrix} s+d \\ d*+q \end{bmatrix}_{Ala-C2} - \begin{bmatrix} s \\ d \end{bmatrix}_{Gly-C2} \right) + P_n \cdot \left(\begin{bmatrix} s \\ d \end{bmatrix}_{Gly-C2} - \begin{bmatrix} 1-P_n \\ P_n \end{bmatrix} \right)} \quad \text{(SE2)}$$

Proline Biosynthesis

Proline could be synthesized via glutamate from glucose, or obtained directly from media components. The equation for calculating proline biosynthesis is derived similarly to that of glycine biosynthesis. Thus the fraction of proline derived from glucose, X^{syn}, was determined from the ^{13}C multiplets of glutamate C4 and proline C4 using Eq. (SE3).

$$X^{syn} = \frac{P_n \cdot \left(\begin{bmatrix} s \\ d \\ t \end{bmatrix}_{Pro-C4} - \begin{bmatrix} (1-P_n)^2 \\ 2 \cdot (1-P_n) \cdot P_n \\ P_n{}^2 \end{bmatrix} \right)}{P_{Glu-C4} \cdot \left(\begin{bmatrix} s \\ d+d* \\ q \end{bmatrix}_{Glu-C4} - \begin{bmatrix} s \\ d \\ t \end{bmatrix}_{Pro-C4} \right) + P_n \cdot \left(\begin{bmatrix} s \\ d \\ t \end{bmatrix}_{Pro-C4} - \begin{bmatrix} (1-P_n)^2 \\ 2 \cdot (1-P_n) \cdot P_n \\ P_n{}^2 \end{bmatrix} \right)} \quad \text{(SE3)}$$

In this equation, P_{Glu-C4} is the specific enrichment of glutamate C4, *s*, *d*, *d**, t, *q* correspond to the relative contributions of singlet, doublet, doublet with a larger coupling constant, triplet, and quartet, respectively.