

## RESEARCH ARTICLE

# Population dynamics inside cancer biomass driven by repeated hypoxia-reoxygenation cycles

Chi Zhang<sup>1</sup>, Sha Cao<sup>1</sup> and Ying Xu<sup>1,2,\*</sup>

<sup>1</sup> Computational Systems Biology Lab, Department of Biochemistry and Molecular Biology, Institute of Bioinformatics, University of Georgia, Athens, GA 30602, USA

<sup>2</sup> College of Computer Science and Technology, Jilin University, Changchun 130012, China

\* Correspondence: xyn@bmb.uga.edu

Received July 5, 2014; Revised October 9, 2014; Accepted October 12, 2014

**A computational analysis of genome-scale transcriptomic data collected on ~1,700 tissue samples of three cancer types: breast carcinoma, colon adenocarcinoma and lung adenocarcinoma, revealed that each tissue consists of (at least) two major subpopulations of cancer cells with different capabilities to handle fluctuating O<sub>2</sub> levels. The two populations have distinct genomic and transcriptomic characteristics, one accelerating its proliferation under hypoxic conditions and the other proliferating faster with higher O<sub>2</sub> levels, referred to as the hypoxia and the reoxygenation subpopulations, respectively. The proportions of the two subpopulations within a cancer tissue change as the average O<sub>2</sub> level changes. They both contribute to cancer development but in a complementary manner. The hypoxia subpopulation tends to have higher proliferation rates than the reoxygenation one as well as higher apoptosis rates; and it is largely responsible for the acidic environment that enables tissue invasion and provides protection against attacks from T-cells. In comparison, the reoxygenation subpopulation generates new extracellular matrices in support of further growth of the tumor and strengthens cell-cell adhesion to provide scaffolds to keep all the cells connected. This subpopulation also serves as the major source of growth factors for tissue growth. These data and observations strongly suggest that these two major subpopulations within each tumor work together in a conjugative relationship to allow the tumor to overcome stresses associated with the constantly changing O<sub>2</sub> level due to repeated growth and angiogenesis. The analysis results not only reveal new insights about the population dynamics within a tumor but also have implications to our understanding of possible causes of different cancer phenotypes such as diffused *versus* more tightly connected tumor tissues.**

**Keywords:** cancer population dynamics; intratumor heterogeneity; cancer cell subpopulations; hypoxia; reoxygenation; cancer evolution

## INTRODUCTION

It is well established that malignant tumor tissues have heterogeneous cell populations, determined based on genomic mutation data [1–3] and morphological information of cells within a tumor [4]. However, very little is known about (i) the typical composition of the subpopulations, if there is such a thing, within a cancer tissue; (ii) how the subpopulations change throughout the development of a cancer; and (iii) if different subpopula-

tions may contribute in different ways to the development of a cancer. Gaining a detailed understanding about these issues will not only offer new insights about the mechanisms of a cancer's development but also potentially lead to improved capabilities for treating the disease.

The key biological question addressed here is: how do cancer cells cope with the fluctuation in the O<sub>2</sub> level, as a result of repeated cycles of hypoxia and reoxygenation [5] caused by tumor growth and angiogenesis throughout the

entire development of a cancer. A natural hypothesis is to test if each tumor may consist of different subpopulations with different capabilities to cope with the fluctuating  $O_2$  levels.

Detection and characterization of distinct subpopulations within a cancer tissue experimentally represent a very challenging problem. While single-cell genome sequencing can provide some information about different cell types within a tissue [3,4,6], it is currently not feasible to derive subpopulation-level information within a tissue and mention how such subpopulation composition changes with time. Fortunately, the availability of genome-scale omic data collected on multiple cancer tissues has made it possible to tackle this problem using computational techniques.

We have recently carried out a computational study to infer and characterize the major subpopulations in terms of their capabilities to deal with the constantly changing  $O_2$  levels within each tissue sample of three cancer types: breast, colon and lung cancers, which all have large numbers of both transcriptomic and matching genomic data in the public domain. Specifically by de-convoluting the gene-expression data of hypoxia and angiogenesis related marker genes collected on tissue samples into contributions by individual subpopulations each consisting of cells with similar expression patterns, we discovered that all tissue samples under consideration each consists of at least three substantial and distinct subpopulations. One subpopulation increases its fraction when the tissue becomes more hypoxic, referred to as the hypoxia population while another increases its proportion as the tissue has a higher  $O_2$  level, named the reoxygenation subpopulation. The third subpopulation is one with relatively longer hypoxia and reoxygenation cycles, suggesting slower growth. Of all three cancer types, the fraction of the hypoxia subpopulation within a cancer tissue generally increases as the underlying cancer advances. We focus on the first two subpopulations in this study.

A number of distinct characteristics are revealed about the two subpopulations through our analyses, which are consistent across all the samples of the three cancer types. Specifically, the hypoxia subpopulation has substantially higher genomic mutation rates than the reoxygenation one. It also grows faster and relies more heavily on glycolysis than the reoxygenation population. In comparison, the reoxygenation subpopulation has cell-cell and cell-ECM adhesion genes as well as growth factors substantially over-expressed.

To the best of our knowledge, all the results reported here represent novel discoveries about the subpopulation composition and dynamics within tumors throughout their development, which may have important implications to development of improved treatments for cancers.

## MATERIALS AND METHODS

### Data used in the analyses

We applied our analysis on the TCGA RNAseqV2 and somatic mutation data. Three cancer types, namely breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD) and lung adenocarcinoma (LUAD), were selected and analyzed with the following consideration: (i) they all have normal control tissue samples; (ii) their sample sizes are relatively large with 994 samples for the BRCA set, 233 samples for the COAD set, and 470 samples for the LUAD set [7,8]; and (iii) the large variance in their expression levels of the hypoxia and reoxygenation marker genes suggests possible repeated cycles of hypoxia and reoxygenation.

Microarray gene-expression data of CTCs and their matching primary cancer samples were collected from the GEO database. Data set GSE18670 (pancreatic cancer, Affymetrix Human Genome U133 Plus 2.0 Array) and GSE38495 (melanoma, Illumina Genome Analyzer IIX) were analyzed to get gene markers of CTCs. Both of the two sets have 6 CTC samples, along with 6 and 20 matching primary cancer samples, respectively.

### Hypoxia, angiogenesis and reoxygenation marker genes

It has been reported that repeated hypoxia-reoxygenation cycles may directly regulate the cell cycle, glucose metabolism and angiogenesis [9]. 242 genes related to various biological processes such as cell cycle, hypoxia response, glycolysis, glucose and lactate transport, and marker genes of angiogenesis signaling, angiogenesis process, endothelial cells, pericytes and reoxygenation are collected from the literatures and public databases such as David [10], KEGG [11] and MSigDB [12], which represent the list of all directly relevant processes to hypoxia and reoxygenation. The detailed gene list is given in Supplementary Table 1.

### Biclustering analysis to identify repeated hypoxia-reoxygenation process

Repeated hypoxia has been observed using electron-paramagnetic resonance in cancer tissues [13]. Transcriptomic level alterations of several key maker genes have been observed in hypoxia and reoxygenation experiments [14–17]. In this analysis, we assume that each tissue sample is collected at one time point from the repeated hypoxia and reoxygenation process, their gene expression data can reflect the hypoxia level of a tissue when it was collected, and are comparable across different samples. The current understanding is that the typical response

time of hypoxia and reoxygenation is approximately a few days to a few weeks, which is much longer than the hours long response time of transcriptomic alterations [18,19]. Hence the transcriptomic level response to repeated hypoxia and reoxygenation cycle can be reflected by the transcriptomic data collected on the tissue samples.

We applied our in-house biclustering software QUBIC 1.0 that can simultaneously cluster the samples and the gene-expression data for identification of genes sharing similar expression patterns across some to-be-identified subset of samples, and find all such patterns above some similarity and size thresholds [20]. We have applied this tool to the sample-expression matrices for all the three cancer types for identification of all major biclusters. The tool is executed using the default parameters.

### Nonnegative matrix factorization (NMF) of a transcriptomic profile of multiple samples

NMF has been applied to cancer stratification [21,22], in which the transcriptomic profile of selected genes across multiple samples of a selected cancer type is modeled as follows:

$$X_{m \times n} = S_{m \times k} \cdot P_{k \times n} + E, \quad k \ll \{m, n\},$$

in which  $X_{m \times n}$  denotes a gene-expression matrix with  $m$  genes and  $n$  samples;  $S_{m \times k}$  denotes a nonnegative signature matrix with  $m$  rows and  $k$  columns, each representing one signature;  $P_{k \times n}$  is the nonnegative proportion matrix with  $k$  rows representing the proportion of each signature; and  $E$  is the residual matrix, the 2-norm of which is denoted by  $\|E\|^2$ . The goal is find  $S_{m \times k}$  and  $P_{k \times n}$  that minimize  $\|E\|^2$  for given  $k \ll m, n$ .

The ‘‘Brunet’’ method in the NMF package [22] of  $R$  is applied with default setting to calculate the NMF on a given gene-expression matrix over 62 hypoxia-reoxygenation marker (HRM) genes. In detail, with an input matrix  $X$  and a given rank  $k$ , the  $S$  matrix and the  $P$  matrix are iteratively calculated using the following recurrence relationship:

$$\begin{cases} P_{ak} \leftarrow P_{ak} \frac{\sum_{i=1}^m S_{ia} X_{ik}}{\sum_{i=1}^m S_{ia}} \\ S_{ka} \leftarrow S_{ka} \frac{\sum_{j=1}^n P_{aj} X_{kj}}{\sum_{j=1}^n P_{aj}} \end{cases}$$

The stopping criterion of this iterative process is when the algorithm reaches convergence. Rather than directly evaluating the convergence of  $\|E\|^2$ , the Brunet method evaluates the stability of a connectivity matrix  $C$  of size  $m \times m$ , in which  $C_{ij} = 1$  if samples  $i$  and  $j$  belong to the

same cluster, and  $C_{ij} = 0$  if they belong to the different clusters [22]. We normalize  $P$  using  $P = \frac{P_{\text{estimated}} * J}{\sum_{ij} P_{i \times j}}$  to ensure the constraint that  $\sum_i P_{i \times j}$  is approximately 1.0 for each  $1 \leq j \leq n$ . The distribution of  $\sum_i P_{i \times j}$  for each cancer type is shown in Supplementary Figure S1.

The rank,  $k$ , of the signature and proportion matrices is determined based on the number of significant principle components of the gene-expression matrix  $X$ , i.e., the number of significant eigenvalues of the variance matrix of  $X$  [23]. Supplementary Figure S2 shows the distribution of the eigenvalues of each cancer type. In all the three cancer types,  $k = 3$  is selected in our analyses based on the eigenvalue distributions.

To determine whether the identified biclusters are due to different cancer subtypes or the changing proportions of different subpopulations within tissues in response to the micro-environmental conditions, we have examined the proportion of each signature, i.e., each column of the matrix  $P$ . If the expression-sample matrix is composed of samples from different cancer subtypes, proportions of different cancer subtypes should in general follow different distributions and hence result in multimodal distributions of the fitted curves for the proportion distributions across all the samples for each cancer type, as shown in Supplementary Figure S3 (derived from data set GSE32646). Generally speaking, if such a proportion distribution is unimodal, the samples are most likely from one population [24]. Based on this argument, we can predict whether the observed expression patterns of the HRM genes are due to cancer subtypes or population dynamics of the same set of subpopulations.

### Basic tools used in our data analyses

We have used Pearson correlation coefficients between gene-expression levels and the estimated proportion of each subpopulation to assess the level of correlation between the expression patterns of specific genes and the sizes of different subpopulations. Student t distribution is assumed when testing the statistical significance of each derived correlation coefficient, using 0.05 as the significance threshold. In the current study, we focus mainly on the hypoxia subpopulation and the reoxygenation subpopulation.

A gene is considered as up-regulated in one subpopulation if and only if the gene’s expression level is significantly positively correlate with the size of the subpopulation and negatively correlate with the size of the other subpopulation. Similarly defined are down-regulated genes.

Mutation rate is defined by the # mutations/24,981, where 24,981 is the total number of mutated genes found in all TCGA data. For each cluster, the mutation rate is defined by the average mutation rate of the samples in the cluster.

For categorical data of cancer stages, Spearman correlation coefficients are used to assess correlations between gene expression levels and cancer stages; and the permutation test is used to assess the statistical significance of each correlation coefficient, using 0.05 as the significance threshold.

Pathway enrichment is assessed using a hypergeometric test (statistical significance cutoff = 0.001) [25]. 2,775 pathways from the Msigdb database including the GO terms and the Msigdb canonical gene sets are used in our pathway-enrichment analysis.

Differentially expressed genes in CTCs *versus* their matching primary cancers are determined using the SAM's differential expression analysis method [26] (statistical significance cutoff = 0.05).

## RESULTS

Transcriptomic data analyses of three cancer types: BRCA, COAD and LUAD, have been carried out to gain a detailed understanding about the subpopulation composition and dynamics within cancer tissues as they progress. It is generally understood that a cancer tumor goes through multiple rounds of hypoxia and reoxygenation due to tumor growth and angiogenesis throughout a cancer's development. Specifically as a tumor grows, its cells tend to become increasingly more hypoxic; once the level of hypoxia exceeds certain thresholds, cells start to release signals for angiogenesis [27,28], which can lead to increased blood flow and hence oxygen into the tissue. Then the cells will go through a growth phase under relatively normoxic conditions before they become hypoxic again, and repeat this cycle. In this paper, we examine how the repeated hypoxia-reoxygenation cycle affect the evolution of cancer subpopulations with distinct characteristics.

### **Biclustering analysis of transcriptomic data reveals sizes of subpopulations change in accordance with O<sub>2</sub> levels**

We have carried out a biclustering analysis of gene-expression data collected on tissue samples of three cancer types, based on 242 marker genes related to hypoxia, angiogenesis signaling, angiogenesis process, markers of endothelial cells (blood vessel cells), and cell proliferation (see Supplementary Table 1 for the gene names). Our goal is to determine if some substantial

subsets of the tissue samples of each cancer type may share similar expression patterns among some (to be determined) subsets of the 242 genes. Our hypothesis is that samples in the same phase of the hypoxia-reoxygenation cycle may share similar expression patterns of some hypoxia and reoxygenation related genes. If true, a biclustering analysis over all the samples of the same cancer type vs the 242 genes should be able to detect such subsets of samples sharing similar expression patterns. Our in-house software QUBIC [20] is used for the analysis, where all samples of each cancer type are represented as a two dimensional matrix with each row representing one of the 242 genes, each column representing one sample and each entry being the expression level of the corresponding gene in the relevant sample (see METHODS for details).

Four substantial biclusters are observed consistently across all three cancer types with statistical significance  $< p = 1E-50$ , given by the biclustering method. A permutation test is conducted, which supports the identified biclusters with the estimated statistical significance levels  $< p = 1E-5$  (Supplementary Figure S4A). These biclusters cover at least 40% samples for each cancer type. Further analyses of the expression patterns and functions of the involved genes in the four clusters revealed that one cluster of samples generally has hypoxia-associated genes up-regulated, hence referred to as the hypoxia cluster; another cluster tends to have angiogenesis-signal receptor and endothelial marker genes up-regulated, hence called the reoxygenation cluster; and the third cluster shows up-regulation in both hypoxia-associated and angiogenesis-signaling genes, referred to as the angiogenesis cluster while the fourth cluster shows no expression changes in any substantial subset of the 242 genes with statistical significance. We speculate that the samples in the fourth cluster may have relatively long hypoxia-reoxygenation cycles, i.e., being slow growing tumors, rather than no such cycles since it is known that all cancers go through angiogenesis; we call these samples as the background group. Supplementary Figure S4B illustrates the biclusters identified in each cancer type.

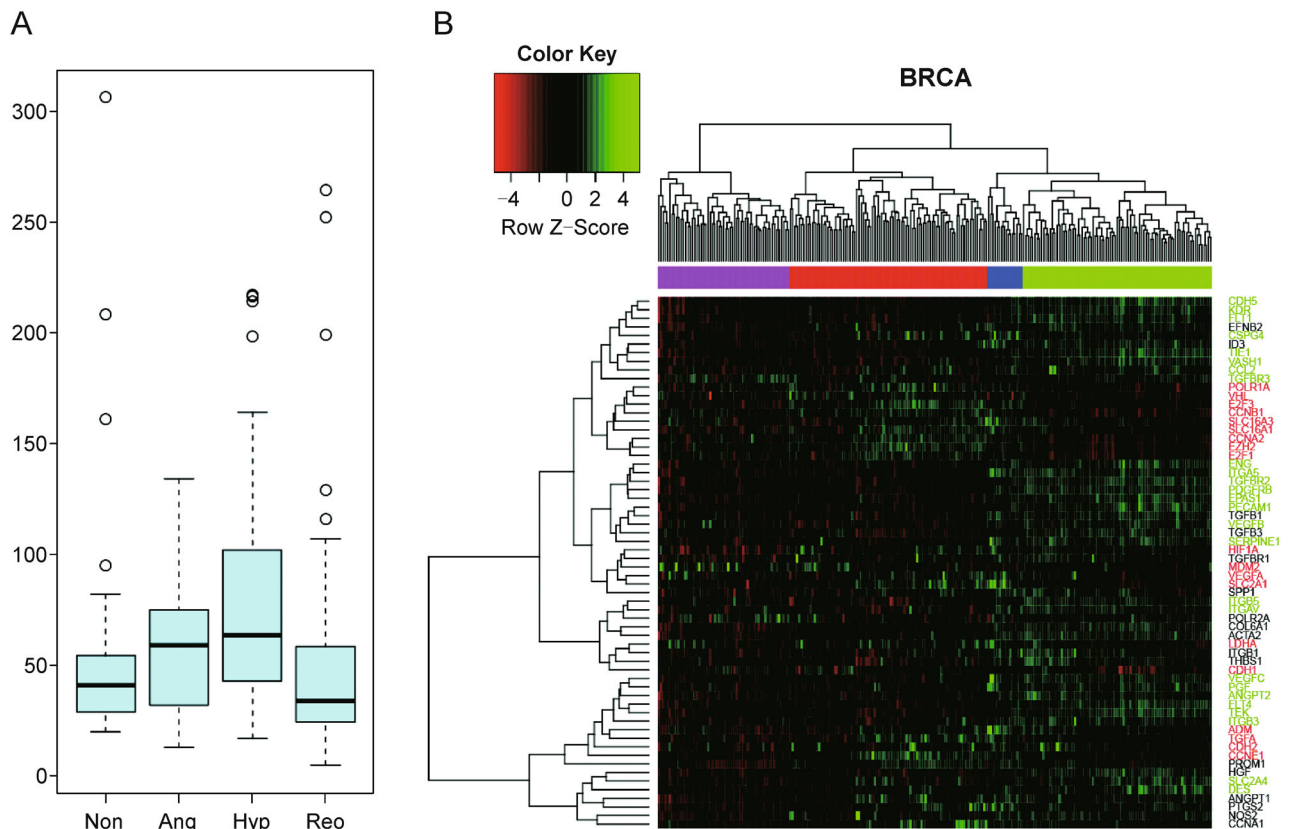
The above result is consistent with our general understanding about the repeated hypoxia-reoxygenation cycles that a tumor tends to go through as it advances [29,30]. In the remaining of the paper, we will carry out further analyses of the first two clusters. Our first question is: do the two clusters of samples each represent the same cell population displaying different gene-expression patterns under different conditions, or each represent a mixture of multiple subpopulations of different cell types?

To rule out the possibility that the two clusters are samples in different stages, a simple statistical test using Fisher exact test was conducted, which revealed that

neither of the two clusters of samples shows preference to any specific stage ( $p = 0.1718, 0.1799, \text{ and } 0.2791$  for breast, colon and lung cancer respectively). This is conducted through building a 2 by 4 table with rows representing the two clusters and the four columns representing four cancer stages, and each entry denoting the number of samples in a specific stage in each cluster (Supplementary Table 2), hence ruling out the possibility of the two clusters being stage dependent. We then checked the genomic mutation rates in the two clusters, and found that the hypoxia cluster has substantially higher mutation rates than the reoxygenation cluster in breast cancer ( $p = 4.612e-06$ , by Mann-Whitney test), as shown in the boxplots of Figure 1A, and colon cancer ( $p = 0.02215$ ) and lung cancer ( $p = 5.433e-07$ ), as shown in Supplementary Figures S5A and S6A. Hence we conclude that the two clusters of samples represent (at least) two different cell types, one having substantially higher genomic mutations than the other, which leaves two possibilities: the two clusters each represent a distinct

(homogeneous) subtype of the disease, or samples in each cluster all consist of a mixture of multiple cell types and the fraction of each cell type is approximately the same across all such samples (hence giving rise to their similar expression patterns across samples in the same cluster) but different for samples in different clusters. A mathematical analysis is carried out to determine which of these two may better represent the available transcriptomic and genomic data.

Before we move on, we noted a clear correspondence among the identified clusters across the three cancer types with all the corresponding clusters sharing similar expression patterns. These shared expression patterns across each set of corresponding clusters are particularly outstanding and discerning across non-corresponding clusters over 62 out of the 242 genes. Hence we use this reduced set of 62 genes, referred to as HRM genes, for the following analysis. See Supplementary Table 3 for the names of these genes. Figure 1B and Supplementary Figures S5B and S6B show the expression patterns of the



**Figure 1.** (A) Distribution of the numbers of mutations in the four major biclustering classes of breast cancer, where the four boxplots, from left to right, are for the background, angiogenesis, hypoxia, and reoxygenation cluster, respectively. (B) Gene expression profiles of the HRM genes in the four major biclustering classes of breast cancer. Samples from the background, hypoxia, angiogenesis, and reoxygenation clusters are color-coded using the pink, red, blue and green bars on top of the figure, respectively. Hypoxia and reoxygenation markers are labeled by red and green, respectively.

HRM genes in the identified biclusters in different cancer types.

### Subpopulations of different cancer cell types are consistently observed across all cancer tissues

#### (i) NMF reveals the existence subpopulations within each tissue sample

A number of studies have reported gene-expression patterns associated with hypoxia and reoxygenation conditions [14–17]. However, no reports have been published, to the best of our knowledge, regarding how such patterns change as a cancer advances and the implications of such changes to a cancer's development. We have carried out a nonnegative factorization of the sample-expression matrix to determine if all tissues of a specific cancer type may each consist of subpopulations of different cell types, each of which has its distinct characteristics. Specifically, we check if the expression matrix for each cancer type can be represented as the product of two nonnegative matrices with one representing the gene-expression signatures of the to-be-identified cell subpopulations across all the samples and the other representing the (to-be-determined) proportion of each cell type in the total cell population in each sample.

This problem can be formulated as a NMF problem of the sample-expression data matrix  $X_{m \times n}$  for each cancer type, i.e.,  $X_{m \times n}$  can be approximately represented as the product of two nonnegative matrices,  $S_{m \times k} \cdot P_{k \times n}$  so the 2-norm of the residual,  $\|X - SP\|^2$ , is minimized, with  $k \ll \min\{m, n\}$  being a to-be identified value, under the constraint that  $\sum_i P_{i \times j}$  is approximately 1.0 for each  $1 \leq j$

$\leq n$ , where  $m$  is the number of marker genes (i.e.,  $m = 62$ ) and  $n$  is the number of samples;  $S$  denotes a signature matrix with  $k$  columns, each representing one signature;  $P$  is a proportion matrix with  $k$  rows representing the proportion of each signature; and  $\|X\|^2$  represents the 2-norm of vector  $X$ . A solution strategy is given in METHODS for this constrained optimization problem, which is used to solve the above for a fixed  $k > 0$ .

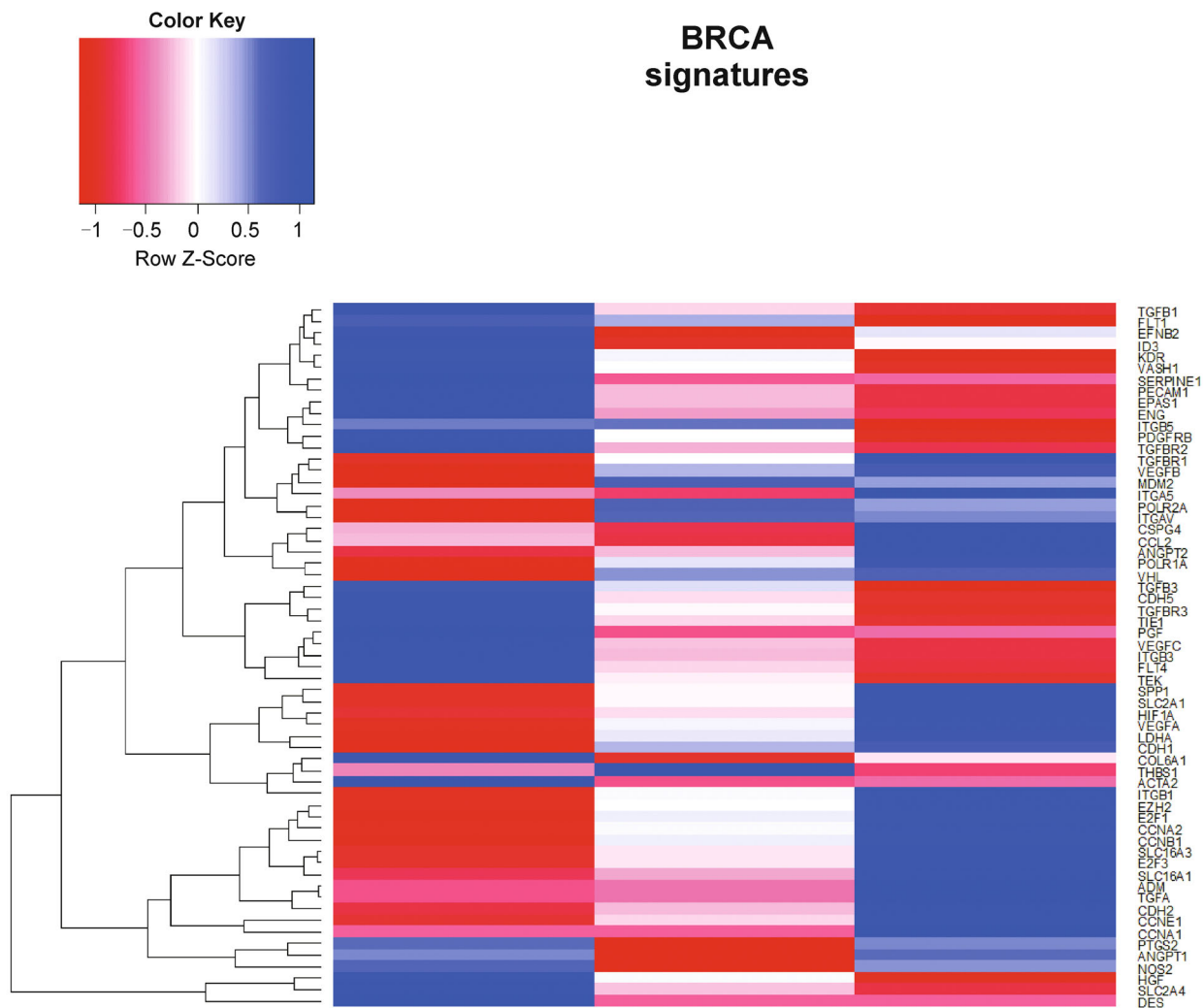
We noted that when  $k = 2$ , the residual matrix  $E$  is quite significant, but drops substantially when  $k = 3$  and further reduction is minimal when  $k$  goes to 4 and beyond (see Supplementary Figure S2) for each of the three cancer types. This strongly suggests that there are three independent, significant and consistent signatures of three cell subpopulations within each cancer tissue, across all samples for each cancer type.

We have examined the distribution of the proportions of each cell type across all samples of each cancer type, and

found that three distributions are unimodal for all three cell types, as shown in Supplementary Figure S7. This strongly suggests that the cell-type specific fraction across all samples of the same cancer type forms one continuous distribution rather than multiple ones, hence suggesting that each sample consists of three subpopulations of distinct cell types, whose fractions change as the hypoxia level changes.

Interestingly, the three cancer types under consideration share very similar expression patterns in each of their three signatures: one signature with highly expressed hypoxia markers, one signature having highly expressed reoxygenation markers and a signature with low expression levels across the vast majority of the HRM genes as shown in Figure 2 and Supplementary Figure S8. Hence we call the three cell types the hypoxia, reoxygenation and background cells, respectively. Further analyses show that there is a strong correspondence between the three cell subpopulations and the biclustering results given in the previous section. Specifically, samples in the hypoxia and reoxygenation biclusters have substantially higher fractions of the hypoxia and the reoxygenation cell subpopulation, respectively, compared to the other clusters. And the angiogenesis samples tend to have more balanced fractions between the hypoxia subpopulation and the reoxygenation subpopulation, respectively.

As expected, we noted a significant negative (positive) correlation between the proportion of the hypoxia (reoxygenation) subpopulation in a tissue and its expression levels of the oxygen-consuming enzymes (excluding any from the 62 HRM genes) (Supplementary Figures S9A and S9B). This revealed a dynamics among the relative sizes of the three subpopulations within a cancer tissue, as a function of the oxygen level, which changes due to repeated growth and angiogenesis. Specifically, the hypoxia subpopulation increases as the hypoxia level goes up while the other two subpopulations clearly shrink due to cell death possibly induced by the incompatible  $O_2$  condition and/or by cell-cell competition within the same growing tissue [31], which has been widely observed in *Drosophila* where more viable cells can directly induce cell death in less fit cells in the neighborhood and replace them by the daughter cells of their own [32]. The same can be said about the fraction decreases of the hypoxia subpopulation in tissues as they become increasingly less hypoxic. This dynamics within a tissue is not apparent from the outside since the overall size of a tumor may not change when the relative proportions of different subpopulations change. One direct supporting evidence of the proposed change of subpopulation proportions is that there is a strong correlation between the proportion of the hypoxia subpopulation and the genomic mutation rate, as discussed in the previous section that the hypoxia cell type has higher genomic mutation rates, as shown in



**Figure 2. Gene-expression signatures for three cell subpopulations in breast cancer.** In the figure, the columns represent the gene expression pattern of the reoxygenation, background and hypoxia subpopulation, from left to right. The color code represents down or up-regulation with the detailed definition given in the top of the figure.

Figure 1 and another figure to be referred afterwards, and Supplementary Figures S5 and S6.

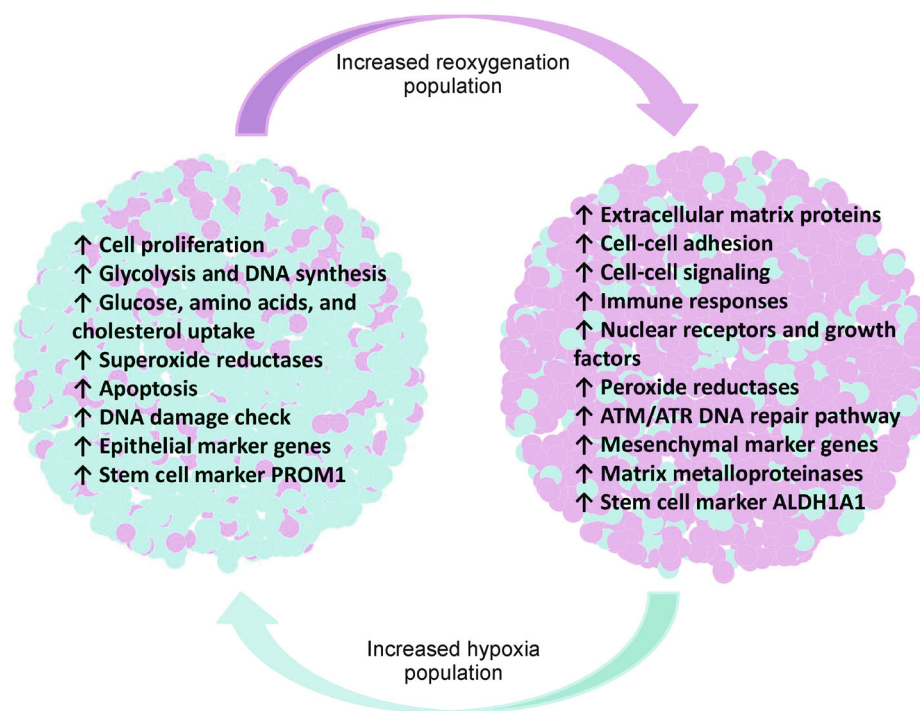
**(ii) Biological processes specific to different subpopulations**

We noted that the following HRM genes are consistently overly expressed in the hypoxia subpopulations across all three cancer types: (i) hypoxia markers *HIF1A* and *VHL*; (ii) hypoxia-response genes *VEGFA*, *SLC2A1*, *ADM*, *TGFA*, *LDHA*, *SLC16A1* and *SLC16A3*; and (iii) cell-cycle genes *CCNE1*, *CCNA2*, *CCNB1*, *E2F1*, *POLR1A*, *E2F3*, *EZH2*, while the following HRM genes are consistently up-regulated in the reoxygenation subpopulation across all cancer types: angiogenic factor receptors

and related genes such as *TIE1*, *DES*, *SERPINE1*, *HGF* and *FLT4*. A detailed list of these genes is given in Supplementary Table 4.

A pathway enrichment analysis was conducted over the up- or down-regulated genes, respectively, in the hypoxia and the reoxygenation subpopulations, separately, with the detailed lists of differentially expressed genes in each subpopulation and in each cancer type given in Supplementary Table 5. The enriched pathways are shown in Figure 3 and Supplementary Table 6.

**Cell proliferation, glycolysis, DNA synthesis, amino acid transporters and cholesterol transporter genes are up-regulated in the hypoxia subpopulation:** The following genes show strong positive correlations between their expression levels and the proportion of the hypoxia subpopulation: *CCNA2*, *CCNB1*, *CCND1*,



**Figure 3. Key characteristics of the two subpopulations.**

*CCNE1*, and *CCNE2*, DNA polymerase *POLR1A*, cyclin dependent kinases *CDK1* and *CDK2*, and cell cycle related transcription factor *E2F1*, *E2F2*, and *E2F3*, as well as glucose transporters *SLC2A1*, *SLC2A4*, glycolysis enzymes *ENO1*, *GAPDH*, *GPI* and *HK2*, lactate dehydrogenase *LDHA*, amino acid transporter genes, and cholesterol sensor and transporter genes *SCARB1* and *LDLR*, hence indicating accelerated cell proliferation by the hypoxia subpopulation with the increase in the hypoxia level. These observations are consistent with published studies showing that more hypoxic cancers tend to be more aggressive [33,34]. In addition, acidity-response genes *CA9* and *CA12*, lactate transporter genes *SLC16A1* and *SLC16A3*, sodium/hydrogen exchanger genes *SLC9A3* and *SLC9A6*, and a number of V-ATPase genes exhibit positive correlations with the proportion of the hypoxia subpopulation in a tissue, indicating increased acidity level in the vicinity of the hypoxia subpopulation.

**Over-expressions in cell-matrix interaction and cell-cell adhesion, cell-cell signaling, immune responses, nuclear receptors and growth factor genes are distinct characteristics of the reoxygenation subpopulation:**

The following genes show strong positive correlations with the proportion of the reoxygenation subpopulation in a cancer tissue: the ECM-component proteins and cell-adhesion proteins such as cell-matrix adhesion proteins, collagens, angiogenesis proteins, extracellular proteins

and cell surface proteins, as well as multiple nuclear receptors such as *NR1D2*, *NR1H2*, *NR2F2*, *NR3C1*, *NR4A1*, *NR4A2*, *NR4A3*, and *NR5A2*, growth factors such as *CLEC11A*, *IGF1*, *FGF1*, *FGF13*, *FGF7* and *PDGFC*, and cell signaling proteins, plus immune response genes such as killer T-cell marker genes *CD1D*, *CD8* and *CD28*, helper T-cell marker *CD4*, and tumor associated macrophage (TAM) markers *CD68*, *CD163* and *CD206* (*MRC1*). In addition, oxygen-consuming enzymes in steroid hormone synthesis pathways are up-regulated, which may produce ligands that can bind to the over-expressed nuclear receptors to activate growth factors for cell proliferation [35,36]. It is worth noting that the helper T-cells and TAMs can release growth factors and regulate angiogenesis, ECM components, adaptive immunity and metastasis [37].

Based on this information, we posit that the two subpopulations play distinct and complementary roles in supporting tumor development and expansion. Specifically, the reoxygenation subpopulation provides growth signals; make new ECMs as the foundation for further tumor growth; enhance cell-cell and cell-ECM interactions to serve as scaffolds for keeping the tumor cells connected; and bring in new blood supply in support of further growth while the hypoxia subpopulation releases angiogenesis signals as a result of increased hypoxia and create an acidic environment through releasing lactic acids to facilitate tumor cells to encroach into the

neighboring areas [38] and to weaken T-cells attack [39]. We suspect that the reoxygenation cells may serve a role to keep all the cells connected since they have increased cell-cell adhesion activities while the hypoxia cells have decreased cell-cell adhesions. For this reason, we call the hypoxia and reoxygenation types the loose and the sticky type, respectively.

From our gene-expression data analyses, the reoxygenation subpopulation is clearly the slower growing subpopulation in comparison with the hypoxia cells, confirming a general evolutionary hypothesis: greater pressures drive organisms to evolve faster [40–42] as the hypoxia group is clearly under greater pressures.

**Over-expressed superoxide reductases in hypoxia subpopulation versus peroxide reductases in reoxygenation subpopulation suggest that the two subpopulations have different sources of reactive oxygen species (ROS):** It has been established that ROS can be produced by different mechanisms including repeated cycles of hypoxia followed by reoxygenation and cell proliferation [43]. Interestingly the two cell populations have different sets of antioxidant enzymes up-regulated, namely anti-superoxide enzymes such as *XDH*, *SOD2* and *IDO1* in the hypoxia subpopulation and peroxide reductases such as *GPX2*, *GPX3*, and *GPX5* in the reoxygenation population. It is known that superoxide is mainly produced by the leakage of mitochondria under hypoxic conditions while peroxide can be produced by repeated cycles of hypoxia-reoxygenation [44–46]. Hence we posit that as a cancer tissue evolves, superoxides and peroxides are alternatively produced by the two dominating subpopulations, leading to the over-expression of different classes of anti-oxidants.

**Over-expressed mesenchymal marker and matrix metalloproteinase genes reveal the invasiveness of the reoxygenation subpopulation:** We have checked the expressions of the mesenchymal marker genes in the two subpopulations, and found that mesenchymal marker genes *CDH2*, *FN1*, *VIM* and *VTN* and matrix metalloproteinase genes *MMP2*, *MMP3*, *MMP14*, *MMP16*, *MMP17*, *MMP19* and *MMP28* are consistently up-regulated in the reoxygenation subpopulation but not in the hypoxia population across all tissue samples, suggesting that the reoxygenation cells represent the more invasive population between the two. Clearly, this, in conjunction with the earlier data, revealed that faster growing and more invasive are independent behaviors of a cancer.

**Up-regulated apoptosis genes in the hypoxia population and the *ATM/ATR* DNA repair pathway in the reoxygenation population are observed:** Apoptosis (activation) genes such as *BAK1*, *CASP3*, *CASP9*, *CYCS*, *DIABLO* and proteasome genes are up-regulated in the hypoxia population. This, in conjunction with an

earlier discussion, reveals that the hypoxia subpopulation has both high proliferation and apoptosis rates. In comparison, the reoxygenation population has both lower proliferation and apoptosis rates.

The observed up-regulation of the DNA repair genes *ATM* and *ATR* suggests more activated DNA repair activities in the reoxygenation population, which is consistent with previous reports [47].

Over-expressed steroid hormone receptors are observed in the reoxygenation population in all three cancer types, such as *ESR1*, *AR* and *PGR* (Supplementary Figure S10A). In contrast, these genes are down-regulated in the hypoxia population. A possible explanation is that their ligand steroid hormones are produced through cholesterol oxidation, which has reduced activities due to reduced oxygen levels. At the tissue level, breast cancer samples fall into three groups each having a distinct expression pattern by these genes (Supplementary Figure S10B). Interestingly the other two cancer types do not have such a grouping. Knowing that the expression levels of *ESR1* and *PGR* are the defining features of breast cancer subtypes, we are tempted to speculate that breast cancer subtypes may be related to the oxygen level of the tumor, which warrants further investigation. Figure 3 summarizes the main distinct as well as common characteristics of the two subpopulations, as detailed in the above.

### (iii) Population dynamics over cancer stages reveal cancer evolution possibly driven by repeated hypoxia-reoxygenation cycles

The impacts of hypoxia and tumor angiogenesis on cancer evolution have been discussed in previous studies, including increased genome instability [30], drug resistance [48] and selection of specific characteristics of cancer stem cells [49]. It has been established that a number of cancer micro-environmental factors such as the hypoxic level, acidic level, oxidative stress, immune responses and ECM composition could be altered in response to the repeated hypoxia-angiogenesis cycles [50–52], leading to the selection of cell subpopulations with certain characteristics, which better fit the micro-environments, hence driving a cancer's evolution.

**Selection of genomic mutations by the repeated hypoxia-reoxygenation cycles:** By comparing the mutation rates of cancer tissue samples at different stages, we noted that the mutation rate does not increase significantly as a cancer advances (Supplementary Figure S11). It is interesting to note that highly mutated cancer genomes are rarely observed in advanced cancers when compared with early stage cancers. Hence we speculate that while increased genomic mutations may provide larger pools for natural selection and hence for cell survival, cancer

cells may have a maximal level of tolerance in the level of mutations; that is, cells with highly mutated genomes are eliminated to keep the cells viable and capable to cope with the changing and complex microenvironments.

To examine the possible roles played by the repeated hypoxia-reoxygenation on cancer genome evolution, we have carried out a regression analysis to correlate the proportions of subpopulations in a cancer tissue to the mutation rates across all samples of each cancer type. The regression result, as shown in Figure 4, revealed that the mutation rate is positively correlated with the proportion of the hypoxia subpopulation and negatively correlated with that of the reoxygenation subpopulation in a tissue sample. This indicates that the mutation rate increases when a tissue becomes more hypoxic and the mutation rate decreases when the tissue is more reoxygenated, which is true for all the cancer types under consideration. Overall, we speculate that there may be a general relationship between the mutation rate and the cell subpopulation with a specific optimal  $O_2$  level, possibly adjusted by the organs; and each subpopulation has its intrinsic mutation rate. Hence when the proportion of a subpopulation changes, the mutation rate of a whole cancer tissue changes accordingly. Detailed regression parameters are listed in Supplementary table 7.

**Biological processes selected by repeated hypoxia-reoxygenation:** We have carried out the following analysis to determine which biological processes may be positively selected by cancer evolution, i.e., defined as those that are strongly correlated with the proportion of a specific subpopulation (see METHODS for details).

The following pathways are positively selected by the repeated hypoxia-reoxygenation: apoptosis, cell proliferation, DNA synthesis, proteasome and WNT signaling, while cell-matrix junction, ECM components, immune responses such as T- and B-cell activation, cell defense system, cell-cell signaling, steroid hormone receptors, regulation of cell proliferation, platelet homeostasis, collagen and lipid rafts are negatively selected (See details in Figure 5). These findings indicate that cancer evolution in general positively selects the main characteristics of the hypoxia subpopulation and negatively select key features of the reoxygenation subpopulation.

**Proportion distributions of the two subpopulations versus cancer stages:** We have examined changes of proportions of subpopulations in a cancer tissue *versus* cancer stages and the T, N, M status, representing cancer progression level, the primary tumor size, the number of regional lymph nodes that cancer has spread to, and the state of distant metastasis, respectively. Interestingly, the proportion of the hypoxic subpopulation in a tumor increases monotonously with respect to the cancer stage N, as shown in Figure 6, indicating a possible correlation between the relative fractions of subpopulations and the

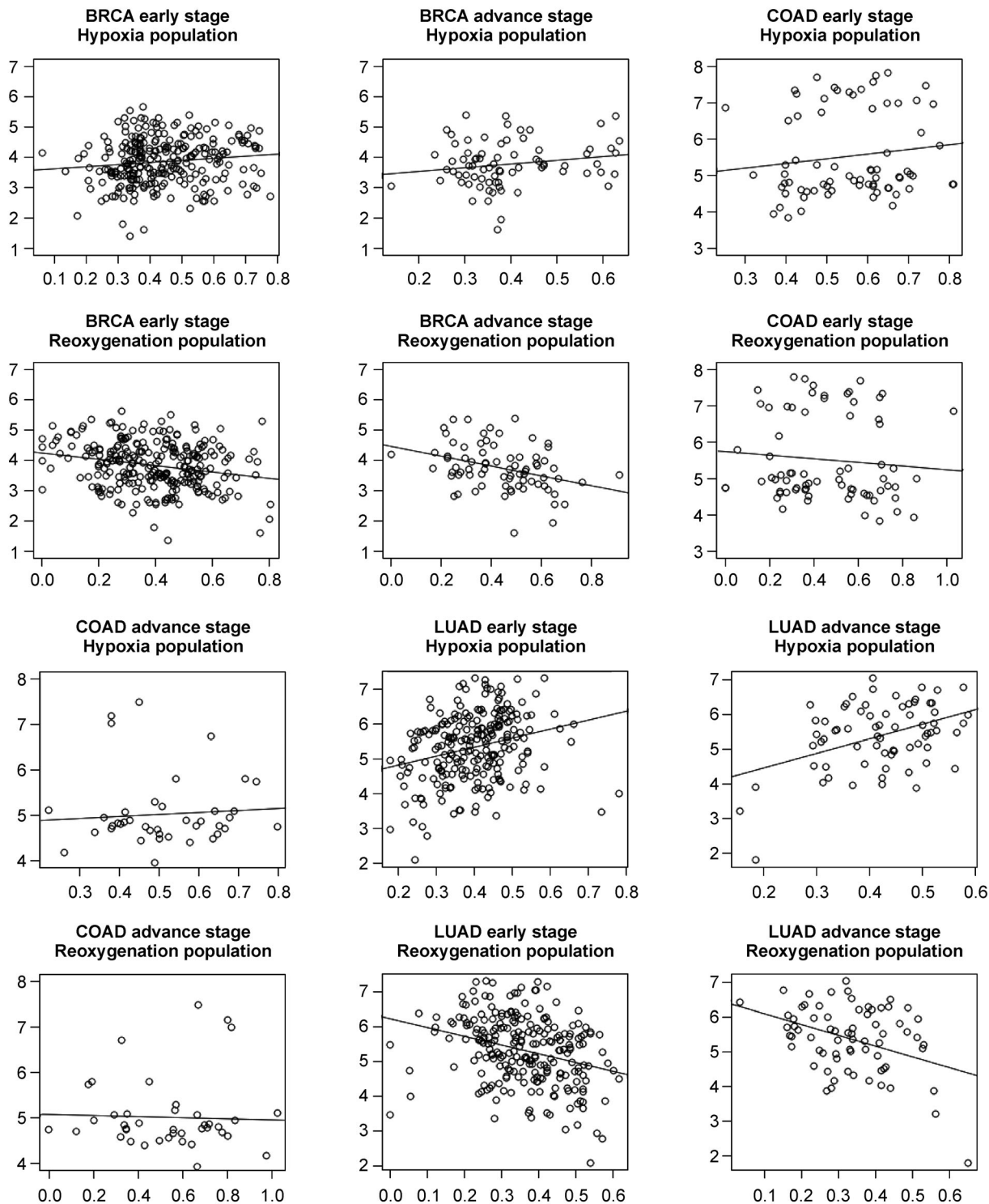
number of local lymph nodes that have been infected.

**Relationship between the two subpopulations and circulating tumor cells:** To understand if the metastasized cells may have come from one of or both the subpopulations, we have compared the gene-expression signatures of the two subpopulations with those of circulating tumor cells (CTC) [53,54] collected from cancer patients' circulation. Specifically we have compared the signatures of CTCs originated from melanoma and pancreatic cancers (the only CTC data we can find and use here) with those of breast, colon and lung. Our analysis found that the cell proliferation, DNA synthesis, reoxygenation signatures, oxygen consuming enzymes, and cell-cell adhesion genes are over-expressed while apoptosis marker and hypoxia marker genes are under-expressed in CTCs compared to their corresponding primary cancers, indicating that the up-regulated genes in the CTCs are consistent with those of the reoxygenation subpopulation in multiple gene categories while they show over-expressed cell-proliferation genes, which resemble more the hypoxia subpopulation. It is clear that additional information such as genomic mutation data is needed to draw a conclusion of whether the CTCs are predominantly from either, both or none of the two major subpopulations in a cancer tissue.

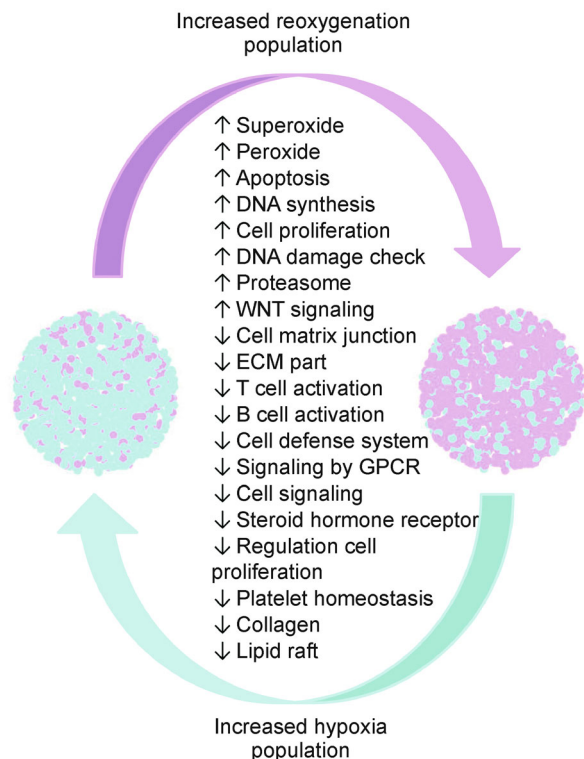
## DISCUSSION

An application of the NMF method to gene-expression data collected on a large number of cancer tissue samples allows us to probe the internal structure of a cancer tissue in terms of the subpopulation composition and dynamics and study how the relative sizes of the subpopulations change with the changing  $O_2$  level, which is due to a repeated cycle of tissue growth and tumor angiogenesis throughout the entire process of a cancer development. It is noteworthy that there could be other ways to define subpopulations within a cancer tissue when examined from other angles. Here we look at the subpopulation composition and dynamics from the perspective of how the cells respond to repeated cycles of hypoxia and angiogenesis.

While it is not intuitively straightforward, our analysis provides strong evidence that the internal composition of the subpopulations, specifically the hypoxia and the reoxygenation subpopulations, in a cancer tissue changes. Various known mechanisms can be used to explain the predicted changes in the subpopulation dynamics in a tissue. This includes (i) increased apoptosis rates of one subpopulation under unfavorable  $O_2$  conditions, while the dead cells can be replaced by new cells of another subpopulation, which generally has increased growth rates under the same condition as our data have shown; and (ii) cell-cell competition as discussed earlier can



**Figure 4. Gene mutation rate versus the proportion of a specific cell type.** In each panel, the x-axis represents the proportion of a subpopulation and the y-axis represents the logarithm of the number of mutations averaged over all samples. The slope of the regression line suggests the correlation between the proportion of a subpopulation and the corresponding mutation rate.

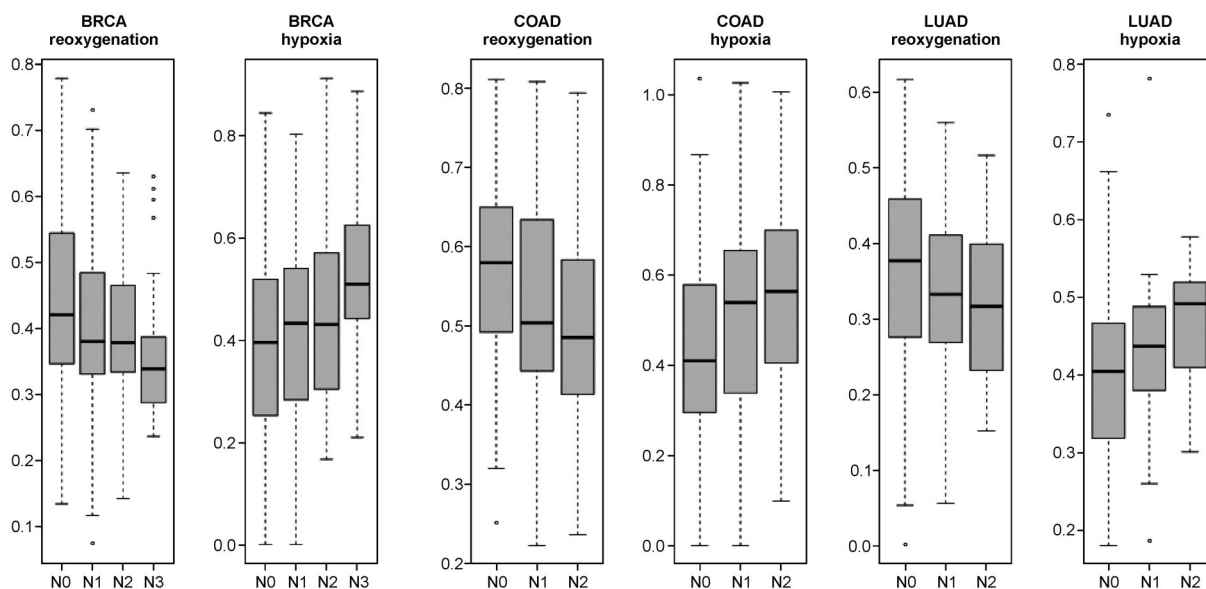


**Figure 5. Biological processes selected by repeated hypoxia-reoxygenation.**

directly replace viable but less fit cells (of the opposite subpopulation) by daughters of the more fit cells.

While the higher genomic mutation rates provide a solid evidence for the changing fraction of the hypoxia subpopulation in a cancer tissue as a function of the  $O_2$  level, it remains to be fully understood how exactly the high mutation rates benefit or are even needed by the development of a cancer tissue. One possible benefit, as discussed earlier, could be that the higher mutations may provide a richer pool of survival pathways, of course at the possible expense of higher death rates of such cells. More in-depth studies are needed to further investigate this issue.

The discovery of the conjugative relationship between the hypoxia and reoxygenation subpopulations, as discussed earlier, clearly adds a new dimension to cancer study. Even though previous studies have found that cancer tissues generally have heterogeneous cells, virtually nothing is known about their mutually beneficial relationship, potentially mutually dependent relationship among the subpopulations as we aim to understand in our future study. One potential direction is to investigate how the ratios between the sizes of the hypoxia and the reoxygenation subpopulations may be relevant to or even affect the form of a cancer biomass, such as diffused *versus* all neoplastic cells sticking together as one or a few tumors, as we would intuitively imagine that the higher percentage of the loose cells of the hypoxia subpopulation in a cancer, the more diffused the cancer cells may become in an organ, which clearly requires further study and validation. Our observation also suggests possible links between the activation of the steroidal hormone



**Figure 6. Distributions of proportions of the hypoxia and the reoxygenation subpopulations versus the N stages.** In each panel, each bar from left to right represents the distribution of the proportion of a subpopulation in stage N0, N1 and N2 (and N3 in breast cancer) samples.

receptors and the hypoxia level, which also requires further analyses, particularly in relevance to breast cancer subtypes.

The discovery made on cancer subpopulation composition and dynamics could also have important implications to the study of the metastasis process since one may intuitively imagine that the loose cells of the hypoxia subpopulation may have higher chances to escape from their bases.

## CONCLUSION

Our computation-based analyses of transcriptomic data of a total of 1,697 tissue samples of three cancer types revealed that the alternating hypoxia and reoxygenation cycles, plus the transitions between them resulted from tumor growth and repeated tumor angiogenesis, drive the changes in the relative proportions of two major subpopulations of cancer cells within a cancer tissue, i. e., the hypoxia and the reoxygenation subpopulations, which provide complementary beneficial conditions and material to facilitate a sustained growth of a neoplastic tumor under substantially different  $O_2$  levels. The new knowledge gained here, along with the computational techniques employed, should help to open new doors in cancer research, allowing studies of cancer development processes and mechanisms in a more realistic setting.

## SUPPLEMENTARY MATERIALS

The supplementary materials can be found online with this article at DOI 10.1007/s40484-014-0032-8.

## ACKNOWLEDGEMENT

Authors would like to thanks Dr. Wenxuan Zhong, Dr. Ping Ma and Mr. Xin Xing from the Department of Statistics in the University of Georgia for their helpful discussion in the NMF method of this research project. Authors would like to thanks Dr. Qin Ma and Mr. Yu Shang from Computational System Biology Lab in the University of Georgia for their helpful discussion in the biclustering analysis of this research project. Authors would also like to thank the anonymous reviewers for their critical reviews and constructive suggestions. YX would like to thank the continuous financial support from the Eminent Scholar Program of the Georgia Research Alliance.

## COMPLIANCE WITH ETHICS GUIDELINES

The authors Chi Zhang, Sha Cao and Ying Xu declare that they have no conflict of interests.

This article does not contain any studies with human or animal subjects performed by any of the authors.

## REFERENCES

- Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K., Wu, H., et al. (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell*, 148, 886–895
- Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.*, 366, 883–892
- Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., Li, F., Wu, K., Liang, J., Shao, D., et al. (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell*, 148, 873–885
- Axelson, H., Fredlund, E., Ovenberger, M., Landberg, G. and Pählman, S. (2005) Hypoxia-induced dedifferentiation of tumor cells—a mechanism behind heterogeneity and aggressiveness of solid tumors. *Semin. Cell Dev. Biol.*, 16, 554–563
- Malec, V., Gottschald, O. R., Li, S., Rose, F., Seeger, W. and Hänze, J. (2010) HIF-1 alpha signaling is augmented during intermittent hypoxia by induction of the Nrf2 pathway in NOX1-expressing adenocarcinoma A549 cells. *Free Radic. Biol. Med.*, 48, 1626–1635
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011) Tumour evolution inferred by single-cell sequencing. *Nature*, 472, 90–94
- The Cancer Genome Atlas Network. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487, 330–337
- The Cancer Genome Atlas Network. (2012) Comprehensive molecular portraits of human breast tumours. *Nature*, 490, 61–70
- Cui, J., Mao, X., Olman, V., Hastings, P. J. and Xu, Y. (2012) Hypoxia and miscoupling between reduced energy efficiency and signaling to cell proliferation drive cancer to grow increasingly faster. *J. Mol. Cell Biol.*, 4, 174–176
- Huang, W., Sherman, B. T. and Lempicki, R. A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, 4, 44–57
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.*, 42, D199–D205
- 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
- Matsumoto, S., Yasui, H., Mitchell, J. B. and Krishna, M. C. (2010) Imaging cycling tumor hypoxia. *Cancer Res.*, 70, 10019–10023
- Dewhirst, M. W. (2009) Relationships between cycling hypoxia, HIF-1, angiogenesis and oxidative stress. *Radiat. Res.*, 172, 653–665
- Polotsky, V. Y., Savransky, V., Bevans-Fonti, S., Reinke, C., Li, J., Grigoryev, D. N. and Shimoda, L. A. (2010) Intermittent and sustained hypoxia induce a similar gene expression profile in human aortic endothelial cells. *Physiol. Genomics*, 41, 306–314
- Dewhirst, M. W. (2007) Intermittent hypoxia furthers the rationale for hypoxia-inducible factor-1 targeting. *Cancer Res.*, 67, 854–855
- Toffoli, S. and Michiels, C. (2008) Intermittent hypoxia is a key regulator of cancer cell and endothelial cell interplay in tumours. *FEBS J.*, 275, 2991–3002

18. Weis, S. M. and Cheresh, D. A. (2011) Tumor angiogenesis: molecular pathways and therapeutic targets. *Nat. Med.*, 17, 1359–1370
19. Carmeliet, P. and Jain, R. K. (2000) Angiogenesis in cancer and other diseases. *Nature*, 407, 249–257
20. Li, G., Ma, Q., Tang, H., Paterson, A. H. and Xu, Y. (2009) QUBIC: a qualitative biclustering algorithm for analyses of gene expression data. *Nucleic Acids Res.*, 37, e101
21. Gao, Y. and Church, G. (2005) Improving molecular cancer class discovery through sparse non-negative matrix factorization. *Bioinformatics*, 21, 3970–3975
22. Brunet, J. P., Tamayo, P., Golub, T. R. and Mesirov, J. P. (2004) Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl. Acad. Sci. USA*, 101, 4164–4169
23. Lee, D. D. and Seung, H. S. (1999) Learning the parts of objects by non-negative matrix factorization. *Nature*, 401, 788–791
24. Kong, X. Z., Zheng, C. H., and Wu, Y. Q. (2007) Molecular cancer class discovery using non-negative matrix factorization with sparseness constraint. *Advanced Intelligent Computing Theories and Applications: With Aspects of Theoretical and Methodological Issues*, 4681. Berlin : Springer-Verlag, 792–802
25. Evangelou, M., Rendon, A., Ouwehand, W. H., Wernisch, L. and Dudbridge, F. (2012) Comparison of methods for competitive tests of pathway analysis. *PLoS One*, 7, e41018
26. Tusher, V. G., Tibshirani, R. and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA*, 98, 5116–5121
27. Liao, D. and Johnson, R. S. (2007) Hypoxia: a key regulator of angiogenesis in cancer. *Cancer Metastasis Rev.*, 26, 281–290
28. Dewhirst, M. W., Cao, Y. and Moeller, B. (2008) Cycling hypoxia and free radicals regulate angiogenesis and radiotherapy response. *Nat. Rev. Cancer*, 8, 425–437
29. Hanahan, D. and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. *Cell*, 144, 646–674
30. Luoto, K. R., Kumareswaran, R. and Bristow, R. G. (2013) Tumor hypoxia as a driving force in genetic instability. *Genome Integr.*, 4, 5
31. Gatenby, R. A., Smallbone, K., Maini, P. K., Rose, F., Averill, J., Nagle, R. B., Worrall, L. and Gillies, R. J. (2007) Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. *Br. J. Cancer*, 97, 646–653
32. Bondar, T. and Medzhitov, R. (2010) p53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell*, 6, 309–322
33. Vaupel, P. and Mayer, A. (2007) Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev.*, 26, 225–239
34. Vaupel, P. (2008) Hypoxia and aggressive tumor phenotype: implications for therapy and prognosis. *Oncologist*, 13, 21–26
35. Lai, L. C. (2002) Role of steroid hormones and growth factors in breast cancer. *Clin. Chem. Lab. Med.*, 40, 969–974
36. Evangelou, A. I., Winter, S. F., Huss, W. J., Bok, R. A. and Greenberg, N. M. (2004) Steroid hormones, polypeptide growth factors, hormone refractory prostate cancer, and the neuroendocrine phenotype. *J. Cell. Biochem.*, 91, 671–683.
37. Quatromoni, J. G. and Eruslanov, E. (2012) Tumor-associated macrophages: function, phenotype, and link to prognosis in human lung cancer. *Am. J. Transl. Res.*, 4, 376–389
38. Hirschhaeuser, F., Sattler, U. G. and Mueller-Klieser, W. (2011) Lactate: a metabolic key player in cancer. *Cancer Res.*, 71, 6921–6925
39. Fischer, K., Hoffmann, P., Voelkl, S., Meidenbauer, N., Ammer, J., Edinger, M., Gottfried, E., Schwarz, S., Rothe, G., Hoves, S., et al. (2007) Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood*, 109, 3812–3819
40. Greaves, M. and Maley, C. C. (2012) Clonal evolution in cancer. *Nature*, 481, 306–313
41. Gutierrez, A., Laureti, L., Crussard, S., Abida, H., Rodríguez-Rojas, A., Blázquez, J., Baharoglu, Z., Mazel, D., Darfeuille, F., Vogel, J., et al. (2013)  $\beta$ -Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity. *Nat. Commun.*, 4, 1610
42. Tompkins, J. D., Nelson, J. L., Hazel, J. C., Leugers, S. L., Stumpf, J. D. and Foster, P. L. (2003) Error-prone polymerase, DNA polymerase IV, is responsible for transient hypermutation during adaptive mutation in *Escherichia coli*. *J. Bacteriol.*, 185, 3469–3472
43. Millar, T. M., Phan, V. and Tibbles, L. A. (2007) ROS generation in endothelial hypoxia and reoxygenation stimulates MAP kinase signaling and kinase-dependent neutrophil recruitment. *Free Radic. Biol. Med.*, 42, 1165–1177
44. Jastroch, M., Divakaruni, A. S., Mookerjee, S., Treberg, J. R. and Brand, M. D. (2010) Mitochondrial proton and electron leaks. *Essays Biochem.*, 47, 53–67
45. Zulueta, J. J., Yu, F. S., Hertig, I. A., Thannickal, V. J. and Hassoun, P. M. (1995) Release of hydrogen peroxide in response to hypoxia-reoxygenation: role of an NAD(P)H oxidase-like enzyme in endothelial cell plasma membrane. *Am. J. Respir. Cell Mol. Biol.*, 12, 41–49
46. Tas, F., Hansel, H., Belce, A., Ilvan, S., Argon, A., Camlica, H. and Topuz, E. (2005) Oxidative stress in breast cancer. *Med. Oncol.*, 22, 11–15
47. Kim, B. M., Choi, J. Y., Kim, Y. J., Woo, H. D. and Chung, H. W. (2007) Reoxygenation following hypoxia activates DNA-damage checkpoint signaling pathways that suppress cell-cycle progression in cultured human lymphocytes. *FEBS Lett.*, 581, 3005–3012
48. Sullivan, R., Paré, G. C., Frederiksen, L. J., Semenza, G. L. and Graham, C. H. (2008) Hypoxia-induced resistance to anticancer drugs is associated with decreased senescence and requires hypoxia-inducible factor-1 activity. *Mol. Cancer Ther.*, 7, 1961–1973
49. Louie, E., Nik, S., Chen, J. S., Schmidt, M., Song, B., Pacson, C., Chen, X. F., Park, S., Ju, J. and Chen, E. I. (2010) Identification of a stem-like cell population by exposing metastatic breast cancer cell lines to repetitive cycles of hypoxia and reoxygenation. *Breast Cancer Res.*, 12, R94
50. Kim, Y., Lin, Q., Glazer, P. M. and Yun, Z. (2009) Hypoxic tumor microenvironment and cancer cell differentiation. *Curr. Mol. Med.*, 9, 425–434
51. Teppo, S., Sundquist, E., Vered, M., Holappa, H., Parkkisenniemi, J., Rinaldi, T., Lehenkari, P., Grenman, R., Dayan, D., Risteli, J., et al. (2013) The hypoxic tumor microenvironment regulates invasion of aggressive oral carcinoma cells. *Exp. Cell Res.*, 319, 376–389
52. Weljie, A. M. and Jirik, F. R. (2011) Hypoxia-induced metabolic shifts in cancer cells: moving beyond the Warburg effect. *Int. J. Biochem. Cell Biol.*, 43, 981–989
53. Sergeant, G., van Eijnsden, R., Roskams, T., Van Duppen, V. and Topal, B. (2012) Pancreatic cancer circulating tumour cells express a cell motility gene signature that predicts survival after surgery. *BMC Cancer*, 12, 527
54. Ramsköld, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., Daniels, G. A., Khrebtukova, I., Loring, J. F., Laurent, L. C., et al. (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.*, 30, 777–782
55. Denko, N. C. (2008) Hypoxia, HIF1 and glucose metabolism in the

- solid tumour. *Nat. Rev. Cancer*, 8, 705–713
56. Gillies, R. J., Verduzco, D. and Gatenby, R. A. (2012) Evolutionary dynamics of carcinogenesis and why targeted therapy does not work. *Nat. Rev. Cancer*, 12, 487–493
57. Scott, B., Sun, C. L., Mao, X., Yu, C., Vohra, B. P., Milbrandt, J. and Crowder, C. M. (2013) Role of oxygen consumption in hypoxia protection by translation factor depletion. *J. Exp. Biol.*, 216, 2283–2292
58. Wheaton, W. W. and Chandel, N. S. (2011) Hypoxia. 2. Hypoxia regulates cellular metabolism. *Am. J. Physiol. Cell Physiol.*, 300, C385–C393
59. Yuan, G., Nanduri, J., Khan, S., Semenza, G. L. and Prabhakar, N. R. (2008) Induction of HIF-1alpha expression by intermittent hypoxia: involvement of NADPH oxidase,  $Ca^{2+}$  signaling, prolyl hydroxylases, and mTOR. *J. Cell. Physiol.*, 217, 674–685