Gene expression signatures as biomarker of tumour hypoxia

Benjamin H.L Harris*, Alessandro Barberis*, Catharine M.L. West, Francesca M. Buffa

*joint first

**Author information**

Benjamin Howell Lole Harris, MBBCh (Hons)
University of Oxford, Department of Oncology, Old Road Campus Research Building, Oxford, United Kingdom, OX3 7DQ.
Email: bhlharris@gmail.com

Alessandro Barberis, PhD
University of Oxford, Department of Oncology, Old Road Campus Research Building, Oxford, United Kingdom, OX3 7DQ.
Email: alessandro.barberis@oncology.ox.ac.uk

Catharine M.L. West, PhD
Translational Radiobiology Group, Institute of Cancer Sciences, University of Manchester, Christie Hospital, Manchester, United Kingdom, M13 9PT.
Email: Catharine.West@manchester.ac.uk

Francesca Meteora Buffa, PhD
University of Oxford, Department of Oncology, Old Road Campus Research Building, Oxford, United Kingdom, OX3 7DQ.
Tel: +44(0)1865 617406
Email: francesca.buffa@oncology.ox.ac.uk

**Article Type:** Overview

**Conflict of Interest:**

C.M.L. West and F.M. Buffa have a patent on a hypoxia signature (WO 011/076895). No potential conflicts of interest were disclosed by the other authors.

**Corresponding Author:**
Francesca Meteora Buffa, PhD
University of Oxford, Department of Oncology, Old Road Campus Research Building, Oxford, United Kingdom, OX3 7DQ.
Tel: +44(0)1865 617406
Email: francesca.buffa@oncology.ox.ac.uk
Abstract:

Hypoxia is a feature of most solid tumours and is associated with a poor prognosis. The hypoxic environment can reduce the efficacy of radiotherapy and some chemotherapeutics, and has been investigated extensively as a therapeutic target. The clinical use of hypoxia-targeting treatment will benefit from the development of a biomarker to assess tumour hypoxia. There are several possible techniques which measure either the level of oxygen or the tumour molecular response to hypoxia. The latter includes gene expression profiling, which measures the transcriptional response of a tumour to its hypoxic microenvironment.

A systematic review identified 32 published hypoxia gene expression signatures. The methods used for their derivation varied, but are broadly classified as: (1) identifying genes with significantly higher or lower expression in cancer cells cultured under hypoxic versus normoxic conditions; (2) using either previously characterised hypoxia regulated genes/biomarkers to define hypoxic tumours and then identifying other genes that are over- or under-expressed in the hypoxic tumours. Both generated gene signatures useful in furthering our understanding of hypoxia biology. However, signatures derived using the second method appear to be superior in terms of providing prognostic information.

Here we summarise all 32 published hypoxia signatures, discuss their commonalities and differences, and highlight their strengths and limitations. This review also highlights the importance of reproducibility and gene annotation, which must be accounted for to transfer signatures robustly for clinical application as biomarkers.

Keywords: Hypoxia, gene signature, radiation, microarray, annotation, treatment stratification.
Introduction

Hypoxia, or lack of oxygen, is a feature of most solid tumours. The hypoxic environment in tumours is the result of proliferation and oxygen demand outstripping angiogenesis, where badly arranged and leaky new vessels fail to provide enough oxygen for a tumour’s expanding mass [1]. Falling oxygen tensions place a selective pressure on cells to either adapt or die, and initiate rapid transcriptional and post-transcriptional responses. One example is the activation of transcription factors (e.g. the hypoxia inducible factors, HIF-1α and 2α [2]) and consequent increased expression of their target genes, and non-coding RNAs that modulate gene expression post-transcriptionally (e.g. microRNAs miR-182 [3] and miR-210 [4]). Hypoxia is therefore unsurprisingly linked to several of the key and emerging hallmarks of cancer [5], including sustained angiogenesis [6], metastasis and invasion [7, 8], evasion of apoptosis [9], suppression of the immune response [10] and genomic instability [3, 11].

Cancer cells that adapt to the selective pressure of hypoxia and adopt a ‘hypoxic phenotype’ tend to form more aggressive and invasive tumours that are resistant to traditional therapies [6, 12]. Hypoxia has in fact been associated with a poor prognosis in different cancer types including head and neck [13], oesophageal [14], gastric [15], liver [16], lung [17], breast [18, 19], bladder and prostate [20].

Tumour hypoxia and resistance to therapy

The poor prognosis of hypoxic tumours is due not only to the above described phenotypic changes and consequent increased aggressiveness, but also a number of other factors.

For sparsely ionising radiations typically used in radiotherapy, incident photons ionise intracellular molecules. Although some of these ionisations occur directly in and damage DNA (direct effects), most are in water because of its abundance within cells (indirect effects). The free radicals produced by these ionisations are stabilised in the presence of oxygen to ‘fix’ the damage caused by radiation [21]. Hypoxic cells are about 3-fold more resistant to the effects of sparsely ionising radiation.

Hypoxia can also impact on the success of conventional chemotherapy. As for radiation, some DNA damaging agents are less effective when oxygen levels fall. For other drugs there is an indirect relationship. Hypoxia occurs when tumour cell proliferation outstrips blood supply because oxygen can only diffuse about 160 μm in tissue, and in a similar way drugs are unable to reach cells distant from the blood supply. In immunotherapy, high molecular weight compounds reach cells via the interstitial space rather than the vascular compartment [22]. The high interstitial pressure found in tumours drives fluid out of the interstitial space and so interferes with delivery [23].

Therefore, hypoxia has been a target of therapeutic interest since the 1950s, especially within the radiotherapy field [21].

Therapeutic strategies to target hypoxic tumours

Various therapeutic strategies have been explored to target hypoxia in patients having radiotherapy. Several approaches have focused on alleviating hypoxia immediately before or during radiotherapy by increasing oxygen delivery to tumours, including the concurrent administration of oxygen mimetics such as nimorazole [24], erythroid stimulating agents [25-27], carbogen (98% oxygen + 2% carbon dioxide) with nicotinamide [28-31] or hyperbaric oxygen [32-34].

There have been some notable successes. The Danish Head and Neck Cancer (DAHANCA) group showed a significant improvement in loco-regional control of patients with supraglottic larynx and
pharynx carcinoma who underwent radiotherapy with the radiosensitiser nimorazole versus a placebo (49% vs. 33% locoregional control rate, p = 0.002) [35]. Also, the combination of radiotherapy with carbogen and nicotinamide has shown promise in laryngeal and bladder cancer [29, 30, 36].

Other approaches have attempted to exploit the hypoxic environment to deliver targeted treatments. Bacterial spores [37-39], viral vectors [40, 41], hypoxia activated pro-drugs [42, 43] and drugs targeted at specific hypoxia induced cellular pathways [44] have all been investigated.

Of the various strategies studied, only nimorazole is used routinely in Denmark and Norway [45] to radiosensitise head and neck tumours, and the approach is now being evaluated in the UK [46]. The lack of routine clinical use despite a high level of evidence that hypoxia modifying treatments improve outcomes following radiotherapy reflects in part problems with transferring promising preclinical findings into clinical benefit. It also reflects the fact that success in patients depends on several biological and clinical factors and there is a need to select the correct patient population to receive hypoxia modifying treatments, namely patients whose tumour are characterised by high levels of hypoxia. Therefore a key question is what is the best way to measure hypoxia in order to determine which tumours are likely to benefit from therapeutic strategies targeting hypoxia?

Measuring tumour hypoxia

A variety of approaches are available for studying the degree of tumour hypoxia, several have been used in clinical studies and trials, but none are routinely used in the clinic.

A direct measure of tumour oxygen tension is possible in accessible tumour sites where an oxygen electrode can be used. This has been successful in a number of clinical studies [47, 48]. Current challenges of this technique include not only its invasiveness and the high level of tumour heterogeneity [49, 50], but also the inability to distinguish tumour areas with viable hypoxic cells from regions of necrosis that do not contribute to tumour growth.

Various cross sectional imaging approaches have been proposed, such as magnetic resonance imaging detection of deoxyhaemoglobin (BOLD MRI) [51] but the biological meaning of the parameters used to reflect hypoxia and their link with tumour hypoxia phenotype has yet to be established clearly [52].

Exogenous markers of hypoxia (nitroimidazole compounds) have been successfully used to denote areas of hypoxia within tumours [53]. The compounds are administered systemically and converted into stable protein adducts in hypoxic areas. These can be detected by immunohistochemistry (IHC) on tumour biopsy or using in-vivo imaging, such as 18F-fluoromisonidazole (18F-FMISO) positron emission tomography (PET) [54]. However, there is debate whether these compounds are sensitive enough to detect an intermediate hypoxic tumour phenotype as HIF-1 is stabilised at higher oxygen tensions that those at which nitroimidazole adducts are formed [55].

Genes that are up- or down-regulated in response to hypoxia reflect the hypoxic phenotype and can provide an indirect measure of the level of hypoxia. Their expression can be assessed at the protein level, using IHC, or at the mRNA level, using for example gene expression arrays [56-58]. Genes which have now been validated in multiple studies are for example HIF1A, the main regulator of the transcriptional response to hypoxia, and its downstream target carbonic anhydrase-9 [59-61]. IHC can be used to identify the location of expression (e.g. membrane, nuclear, cytoplasmic) which in turn can reveal the functional status of the protein; it also allows for estimation of the heterogeneity in expression levels between different cells or areas. However, IHC is not fully quantitative, requiring
some degree of subjective interpretation even in the most automated protocols. Furthermore, antibody availability, specificity and sensitivity can vary, limiting this technique to a relatively small number of biomarkers. In some cases, proteins that are directly or indirectly regulated by hypoxia can be measured in blood samples. Circulating biomarkers of hypoxia that have been evaluated include osteopontin, hepatocyte growth factor, and interleukin-8 [62-64].

One limitation of assessing hypoxia by measuring cellular adaptation is that genes/proteins can be under the regulation of other mechanisms as well as hypoxia (e.g. glucose concentration [52]) and therefore may not be a specific measure of hypoxia across tumours. Another issue can be somatic mutations interfering with their expression. For example, mutations in the Von Hippel-Lindau (VHL) gene cause a decreased degradation of HIF-1α which can result in a hypoxia phenotype in the absence of hypoxia.

**Hypoxia gene expression signatures**

Several studies have used gene expression microarrays to determine a global transcriptional response to hypoxia (Table 1). This has typically been done in cell lines exposed to hypoxia versus normoxia. Genes found to be significantly upregulated, or exceeding a defined threshold from baseline normoxic expression, are typically grouped together and have been referred to as a ‘hypoxia gene expression signature’ or ‘hypoxia signature’. This approach has been used in multiple fields to investigate the effects of hypoxia on the vasculature [65, 66] and the immune system [67], as well as on tumour cells [68, 69]. Here, we refer to this approach as the in-vitro ‘traditional’ approach.

Translating such signatures directly to the clinic is challenging and controversial as the transcriptional response in cell line in-vitro culture might be, in part, artefactual and unrepresentative of the hypoxia phenotype observed in tumours.

To address this limitation, alternative approaches have integrated an in-vitro traditional approach with analysis of gene expression in clinical samples (Table 1). These methods typically use either gene signatures obtained from the in-vitro traditional approach or previous knowledge of hypoxia regulated gene as a starting point to derive a gene expression signature of the tumour hypoxia phenotype. The first step typically establishes if the in-vitro derived signature genes are expressed and regulated as expected in the clinical samples; further steps can then use computational methods to aggregate additional genes with the aim of deriving a new signature more representative of the tumour hypoxia phenotype. These different approaches are outlined schematically in Figure 1.

**Hypoxia signatures commonalities and differences**

The evaluation of individual genes or common pathways in hypoxia signatures is likely to lead to the continued identification of potential therapeutic targets.

A systematic interrogation of the National Center for Biotechnology Information database (Pubmed) [70], Scopus [71], Web of Science [72], hypoxiadb [73] and genesigdb [74], identified 32 published hypoxia signatures to date (Table 1). The search terms where "hypoxia signature" or "hypoxia" AND/+ "signature" in each database. The majority of the signatures identified (27/32) were derived using the in-vitro ‘traditional’ approach. The size of the signatures, that is the number of genes, varied from 4 to 759. The median number of genes in signatures derived from the in-vitro traditional
approach was 85, compared to 61 in signatures derived with approaches taking into account expression in clinical samples.
Table 1. All 32 hypoxia gene expression signatures published to date. Studies highlighted in bold derived signatures using approaches involving analysis of clinical samples. Please note that there are interdependencies between signatures (Figure 2). Starred (*) signatures contain both up- and down-regulated genes.

<table>
<thead>
<tr>
<th>Pubmed ID</th>
<th>Author &amp; publication date</th>
<th>Number of gene symbols</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>10706099</td>
<td>Koong et al, 2000</td>
<td>10</td>
<td>[75]</td>
</tr>
<tr>
<td>12947397</td>
<td>Denko et al, 2003</td>
<td>80</td>
<td>[76]</td>
</tr>
<tr>
<td>15093745</td>
<td>Jogi et al, 2004</td>
<td>107</td>
<td>[68]</td>
</tr>
<tr>
<td>15100389</td>
<td>Ning et al, 2004</td>
<td>104</td>
<td>[66]</td>
</tr>
<tr>
<td>15374877</td>
<td>Manalo et al, 2005</td>
<td>107</td>
<td>[77]</td>
</tr>
<tr>
<td>15833863</td>
<td>Wang et al, 2005</td>
<td>56</td>
<td>[78]</td>
</tr>
<tr>
<td>15994966</td>
<td>Detwiller et al, 2005</td>
<td>27</td>
<td>[79]</td>
</tr>
<tr>
<td>16849508</td>
<td>Bosco et al, 2006</td>
<td>177</td>
<td>[67]</td>
</tr>
<tr>
<td>16507782</td>
<td>Mense et al, 2006</td>
<td>111</td>
<td>[80]</td>
</tr>
<tr>
<td>16740701</td>
<td>Aprelikova et al, 2006</td>
<td>236</td>
<td>[69]</td>
</tr>
<tr>
<td>16417408</td>
<td>Chi et al, 2006</td>
<td>111</td>
<td>[81]</td>
</tr>
<tr>
<td>16565084</td>
<td>Elvidge et al, 2006</td>
<td>181</td>
<td>[82]</td>
</tr>
<tr>
<td>16595741</td>
<td>Peters et al, 2006</td>
<td>159</td>
<td>[65]</td>
</tr>
<tr>
<td>17532074</td>
<td>Seigneuric et al, 2007 (early 0% O_2)</td>
<td>71</td>
<td>[83]</td>
</tr>
<tr>
<td>17532074</td>
<td>Seigneuric et al 2007, (early 2% O_2)</td>
<td>34</td>
<td>[83]</td>
</tr>
<tr>
<td>17532074</td>
<td>Seigneuric et al, 2007 (common genes between early 0% and 2% O_2 signatures)</td>
<td>14</td>
<td>[83]</td>
</tr>
<tr>
<td>17409455</td>
<td>Winter et al, 2007</td>
<td>99</td>
<td>[58]</td>
</tr>
<tr>
<td>17187782</td>
<td>Shi et al, 2007</td>
<td>32</td>
<td>[84]</td>
</tr>
<tr>
<td>17320280</td>
<td>Sung et al, 2007</td>
<td>90</td>
<td>[85]</td>
</tr>
<tr>
<td>18984585</td>
<td>Beyer et al, 2008</td>
<td>159</td>
<td>[86]</td>
</tr>
<tr>
<td>20592013</td>
<td>Van Malenstein et al, 2009*</td>
<td>7 (3*)</td>
<td>[87]</td>
</tr>
<tr>
<td>19491311</td>
<td>Benita et al, 2009*</td>
<td>81 (57*)</td>
<td>[88]</td>
</tr>
<tr>
<td>19832978</td>
<td>Fardin et al, 2009</td>
<td>8</td>
<td>[89]</td>
</tr>
<tr>
<td>19291283</td>
<td>Hu et al, 2009</td>
<td>13</td>
<td>[90]</td>
</tr>
<tr>
<td>20652058</td>
<td>Fardin et al, 2010</td>
<td>35</td>
<td>[91]</td>
</tr>
<tr>
<td>20429727</td>
<td>Sorensen et al, 2010</td>
<td>27</td>
<td>[92]</td>
</tr>
<tr>
<td>20087356</td>
<td>Buffa et al, 2010</td>
<td>51</td>
<td>[56]</td>
</tr>
<tr>
<td>20416888</td>
<td>Ghorbel et al, 2010</td>
<td>166</td>
<td>[93]</td>
</tr>
<tr>
<td>21846821</td>
<td>Toustrup et al, 2011</td>
<td>15</td>
<td>[94]</td>
</tr>
<tr>
<td>21325071</td>
<td>Ghazouli et al, 2011</td>
<td>70</td>
<td>[57]</td>
</tr>
<tr>
<td>22356756</td>
<td>Starmans et al, 2012</td>
<td>759</td>
<td>[95]</td>
</tr>
<tr>
<td>23820108</td>
<td>Eustace et al, 2013</td>
<td>26</td>
<td>[96]</td>
</tr>
</tbody>
</table>
A comparison of the 32 signatures showed that no gene was found in all signatures. This could reflect both the heterogeneity of the methods used, but also real different between cell lines and clinical samples from different tumour types. However, a number of genes were more prevalent. Table 2 outlines the most commonly found genes. The left hand table shows the 20 most frequent genes occurring in all signatures when using the published symbols as the reference annotation. The right hand table shows the 20 most frequent microarray probesets and their annotated gene symbols, using the most reliable published annotation (e.g. probesets or ensemble ID) and BioMart to retrieve the array probesets. Whilst this comparison shows a significant overlap, it is notable that the two approaches do not produce the same output even if the starting point is the same published signature. This highlights the importance of using a reliable annotation. Symbol annotation changes as our biological understanding of gene function matures and old names are typically kept as aliases. Whilst gene symbols are unique, aliases are not, making re-annotation challenging and prone to mistakes. Conversely, the probeset measuring the expression level, with its target sequence, is the identifiable physical object that measured the expression and, for a given gene expression array platform can always be cross referenced. Thus, in this respect it is the most stable reference, and the one which should always be provided to allow for re-annotation of symbols and other gene identifiers as biological understanding evolves.

Gene frequencies from a simple gene symbol tally are constrained with those generated from the best available annotation (including probesets or ensembl IDs). The results are striking. For instance, the original frequency count of AK3 working from symbols alone was 14 but on re-annotation from original probesets or other gene IDs this dropped to three; the dropped probesets originally annotated as AK3 matched AK4 on re-annotation. On the other hand, re-annotation can also cause signatures to shrink and data can be lost in the process: for example when re-annotating Bosco et al, 2006 from Genbank IDs to symbols, the signature shrank from 177 genes to 57 genes. Therefore, all manuscripts, should submit a minimum of three annotations but more if available (probeset ID, ensembl ID and gene symbol). The most robust annotation for microarrays is the microarray probeset as this corresponds directly to a target DNA sequence. If a custom microarray is used, the sequence of the probesets in the hypoxia signature must be included.

It is important to note that not all signatures were generated independently. Some groups have derived multiple hypoxia signatures and the lineages of these signatures are outlined in Figure 2.
When an array probeset was not listed in the original publication, another ID was used (Accession, Ensembl, Entrez or Unigene) and the matching probesets were retrieved. If only the symbol was present in the original publication, it was re-annotated to the official current symbol, and the matching probesets retrieved.

Table 2 The top 20 most frequently appearing genes across 32 published hypoxia signatures. The left hand table was generated when using published symbol annotation. The right hand table was generated using the most reliable annotation available (e.g. array probeset or ensemble ID) and then mapping them to array probesets.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Description</th>
<th>Frequency</th>
<th>Affymetrix probeset</th>
<th>Gene symbol</th>
<th>Description</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNIP3</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3</td>
<td>31</td>
<td>200632_s_at</td>
<td>NDRG1</td>
<td>N-myc downstream regulated 1</td>
<td>18</td>
</tr>
<tr>
<td>F3</td>
<td>coagulation factor III</td>
<td>25</td>
<td>207543_s_at</td>
<td>P4HA1</td>
<td>prolyl 4-hydroxylyase, alpha polypeptide I</td>
<td>17</td>
</tr>
<tr>
<td>LOX</td>
<td>lysyl oxidase</td>
<td>21</td>
<td>221478_at</td>
<td>BNIP3L</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3-like</td>
<td>17</td>
</tr>
<tr>
<td>TNF</td>
<td>tumour necrosis factor</td>
<td>18</td>
<td>221479_s_at</td>
<td>BNIP3L</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3-like</td>
<td>16</td>
</tr>
<tr>
<td>TH</td>
<td>tyrosine hydroxylase</td>
<td>18</td>
<td>200737_at</td>
<td>PGK1</td>
<td>phosphoglycerate kinase 1</td>
<td>15</td>
</tr>
<tr>
<td>SLC2A1</td>
<td>solute carrier family 2 (facilitated glucose transporter), member 1</td>
<td>17</td>
<td>202912_at</td>
<td>ADM</td>
<td>adrenomedullin</td>
<td>15</td>
</tr>
<tr>
<td>PGK1</td>
<td>phosphoglycerate kinase 1</td>
<td>17</td>
<td>201849_at</td>
<td>BNIP3</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3</td>
<td>15</td>
</tr>
<tr>
<td>NDRG1</td>
<td>N-myc downstream regulated 1</td>
<td>17</td>
<td>201848_s_at</td>
<td>BNIP3P1</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3 pseudogene</td>
<td>15</td>
</tr>
<tr>
<td>GAL</td>
<td>galanin/GMAP prepropeptide</td>
<td>16</td>
<td>202022_at</td>
<td>ALDOC</td>
<td>aldolase C, fructose-bisphosphate</td>
<td>13</td>
</tr>
<tr>
<td>BNIP3L</td>
<td>BCL2/adenovirus E1B 19kDa interacting protein 3-like</td>
<td>16</td>
<td>202620_s_at</td>
<td>PLOD2</td>
<td>procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2</td>
<td>13</td>
</tr>
<tr>
<td>ANG</td>
<td>angiogenin, ribonuclease, RNase A family, 5</td>
<td>16</td>
<td>200738_s_at</td>
<td>PGK1</td>
<td>phosphoglycerate kinase 1</td>
<td>13</td>
</tr>
<tr>
<td>P4HA1</td>
<td>prolyl 4-hydroxylase, alpha polypeptide I</td>
<td>15</td>
<td>202733_at</td>
<td>P4HA2</td>
<td>prolyl 4-hydroxylase, alpha polypeptide I</td>
<td>13</td>
</tr>
<tr>
<td>ADM</td>
<td>adrenomedullin</td>
<td>15</td>
<td>202887_s_at</td>
<td>DDIT4</td>
<td>DNA-damage-inducible transcript 4</td>
<td>13</td>
</tr>
<tr>
<td>AK3</td>
<td>adenylate kinase 3</td>
<td>14</td>
<td>202364_at</td>
<td>MXI1</td>
<td>MAX interactor 1, dimerization protein</td>
<td>12</td>
</tr>
<tr>
<td>PDK1</td>
<td>pyruvate dehydrogenase kinase, isozyme 1</td>
<td>13</td>
<td>202619_s_at</td>
<td>PLOD2</td>
<td>procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2</td>
<td>12</td>
</tr>
<tr>
<td>ERO1L</td>
<td>ERO1-like (S. cerevisiae)</td>
<td>13</td>
<td>218498_s_at</td>
<td>ERO1L</td>
<td>ERO1-like (S. cerevisiae)</td>
<td>12</td>
</tr>
<tr>
<td>ALDOC</td>
<td>aldolase C, fructose-bisphosphate</td>
<td>13</td>
<td>201250_s_at</td>
<td>SLC2A1</td>
<td>solute carrier family 2 (facilitated glucose transporter), member 1</td>
<td>12</td>
</tr>
<tr>
<td>PLOD2</td>
<td>procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2</td>
<td>12</td>
<td>217356_s_at</td>
<td>PGK1</td>
<td>phosphoglycerate kinase 1</td>
<td>12</td>
</tr>
<tr>
<td>P4HA2</td>
<td>prolyl 4-hydroxylase, alpha polypeptide II</td>
<td>12</td>
<td>227068_at</td>
<td>PGK1</td>
<td>phosphoglycerate kinase 1</td>
<td>12</td>
</tr>
<tr>
<td>MXI1</td>
<td>MAX interactor 1, dimerization protein</td>
<td>12</td>
<td>221009_s_at</td>
<td>ANGPT L4</td>
<td>angiopoietin-like 4</td>
<td>12</td>
</tr>
</tbody>
</table>

1 When an array probeset was not listed in the original publication, another ID was used (Accession, Ensembl, Entrez or Unigene) and the matching probesets were retrieved. If only the symbol was present in the original publication, it was re-annotated to the official current symbol, and the matching probesets retrieved.
 Uses of hypoxia signatures

Hypoxia signatures, and the methods used to generate them, have been exploited successfully not only to identify new hypoxia-induced genes and related cellular pathways [97-100], but also importantly for i) characterisation of a patient’s prognosis and ii) prediction of benefit from the addition of hypoxia modifying agents to radiotherapy (patient stratification).

From a hypoxia signature to a clinically useful hypoxia signature score

The first issue in trying to translate hypoxia signatures into a clinical test is to address which tumours are hypoxic, and which methods and thresholds identify them with the highest confidence. Specifically, even if a gene signature contained genes that validated in terms of prognostication and correlation with other measures of hypoxia, there is still a need to establish how to classify a tumour robustly as hypoxic, i.e. determine a cut-off value for patient stratification and testing/applying prospectively. Some studies have used a continuous score to rank the clinical samples within a cohort from the least to the most hypoxic [56]; although this approach is correct until a threshold is defined this poses some practical issues for clinical implementation. Another method is to train a weighted classifier using an external measure of hypoxia (e.g. oxygen electrodes) [94]; however this will change depending on the independent method chosen and since there is no agreement on a gold standard to measure hypoxia in the clinical setting this might be biased and not reflect the clinically relevant hypoxia, or the tumour hypoxia phenotype. Another method is to standardize the log2 gene expression of the signature and classify patients with a score greater than zero as “more hypoxic” and those less than zero as “less hypoxic” [81]. Other groups use K-means clustering on a hypoxia signature score to allocate “more hypoxic” and “less hypoxic” groups [91]. However there is no consensus on the ideal method to define tumours as hypoxic and there is no methodological study to date assessing and comparing the performance of the different methods.

It has also been shown that subtle changes to methods used in pre-processing of expression data, i.e. the method to remove the background noise gene expression, may reduce the ability to replicate prognostic findings, for both single biomarkers [101] and multi-gene hypoxia signatures [102]. Fox et al. showed that an ensemble classification technique gives the best stability and most reproducible size of effect, and can help distinguish good signatures (ones that validate across cohorts) from poor signatures (ones that do not perform well in additional datasets) [102].

Prognostic hypoxia signatures

A number of hypoxia signatures have been shown to be independent prognostic factors for patients with various tumour types. In general, signatures that involve the analysis of clinical samples at some stage of derivation have proved to be more robust prognostic factors.

Signatures derived using the in-vitro ‘traditional’ approach

The first to investigate the prognostic value of these signatures was Chi et al. 2006 [81]. Here, four normal cell lines were subjected to 0% and 2% oxygen and an ‘epithelial hypoxia signature’ was derived by considering genes commonly upregulated across cell lines. The expression of this signature was then used to cluster clinical samples from retrospective patient cohorts, namely two breast cancer datasets (total N= 373) and one ovarian cancer patient series (N=72). The clustering revealed two distinct patient groups, one with high expression of the epithelial hypoxia signature
and one with low expression. Patients with tumours characterised by high expression of this signature had significantly decreased relapse-free and overall survival, which was maintained after adjustment for clinical covariates.

As a follow up to Chi’s work, Seigneuric et al. 2007 investigated the differential prognostic potential of the temporal response to hypoxia of human mammalian epithelial cells [83]. Specifically, they found that the early transcriptional responses to hypoxia (0% and 2% oxygen) yielded 15 common upregulated genes. The 15 gene signature was prognostic in survival analysis, including clinical covariates, in 251 patients with breast cancer. Surprisingly, the hypoxia signature from cells left in hypoxia for longer (the ‘late hypoxia signatures’) was not prognostic. The same group subsequently identified seven gene clusters, defining seven different signatures, upregulated at different times in three cancer cell lines subjected to hypoxic conditions [95]. None of these signatures were prognostic in a cohort of 2312 breast cancer patients. Whereas on the same dataset, other signatures including Winter et al. [58], Hu et al. [90] and Buffa et al. [56] all had prognostic significance.

Sorensen et al. 2010 identified a subset of genes upregulated under hypoxia in at least four tumour cell lines (3 oral carcinoma, 1 hypopharyngeal and 1 cervical carcinoma) [92, 94]. These genes were subsequently used by Toustrup et al. to develop a signature in head and neck squamous cell carcinoma (HNSCC) clinical samples, discussed in the next section.

Fardin et al. 2010 investigated the gene expression of neuroblastoma (NB) cell lines under hypoxic conditions to generate a NB hypoxia signature [91]. The NB hypoxia signature was prognostic in a dataset of 88 NB patients. Patients with a higher expression of the genes in this signature had significantly worse event free and overall survival, independent of other clinical factors. Fardin also tested other hypoxia signatures on this dataset and, of these, only the Winter signature was prognostic (log rank, p=0.006).

Approaches involving optimization of the signature in clinical samples

Toustrup et al, 2011 built on the signature derived by Sorensen et al. 2010 in cancer cell lines. This signature was reduced into a 15 gene ‘hypoxia gene expression classifier’ by means of comparison with other biomarkers of hypoxia first in xenografts, using FAZA as a marker of hypoxia, and then in HNSCC clinical samples, using oxygen electrode measurements. The resultant 15-gene classifier was an independent prognostic factor in a dataset of 323 head and neck cancer patients (Danish Head and Neck Cancer group (DAHANCA) 5 cohort) [94].

Buffa and colleagues developed a pathway discovery method [56] that can computationally discover new relationships between bona fide gene candidates (so called ‘seeds’) of a tumour phenotype, hypoxia in this case, and other genes not necessarily previously linked with the phenotype. The seeds are typically genes extracted from hypoxia signatures previously derived using the in-vitro traditional approach or by summarization of existing knowledge. The method then builds co-expression networks in hundreds of clinical samples by starting from the seeds and finding other genes that have the highest expression similarity to the seeds across samples. This defines a hypoxia signature for the clinical samples. The hypothesis is that this method not only tests the usefulness of the in-vitro derived genes (the seeds) in clinical samples by testing their expression and co-
expression, but also discovers new genes regulated by hypoxia which would not be observed in in-vitro experiments [100].

This approach provided a first signature, the Winter et al. 2007 signature, which was used gene expression array data in 59 HNSCC samples to identify a 99-gene hypoxia signature (or ‘hypoxia metagene’ as named in the publication) [58]. The signature comprised upregulated genes and was shown to have independent prognostic significance in cohorts of 60 HNSCC and 295 breast cancers (the NKI dataset, as previously used by Chi). Buffa and colleagues refined the signature by repeating the process in multiple HNSCC and breast cancer cohorts to derive meta-signatures: a ‘common hypoxia metagene’, a head and neck meta-signature and a breast meta-signature [56]. The common hypoxia metagene was prognostic across three cancer types (breast lung and HNSCC) using four independent datasets comprising of 295 breast, 216 lung and 80 HNSCC patients [56, 98]. This was subsequently confirmed by independent studies [95, 101, 102]. As a smaller number of genes (~26) is useful for platform development for clinical exploitation, reduced signatures were also derived and shown to be prognostic across the datasets.

Hu et al. 2009 developed a 13-gene profile based on the vascular endothelial growth factor (VEGF) signalling pathway, and used this to divide breast cancer patients into three groups (low, medium and high VEGF profile scorers). The VEGF profile showed significant prognostic ability in three independent datasets (295 breast cancer, 111 lung cancer and 50 glioblastoma patients)[90].

**Treatment stratification and prediction of response to treatment**

Stratification using hypoxia signatures is a fledging area. Although this has been mentioned in several papers as an ultimate goal, only two groups have investigated the predictive value of hypoxia signatures.

Toustrup et al. 2011 studied the effectiveness of their 15-gene ‘hypoxia gene expression classifier’ on 323 HNSCC patients randomly assigned to nimorazole or placebo with radiotherapy (DAHANCA 5[63]) [94]. Toustrup found that those retrospectively selected to be in the ‘more hypoxic’ group by their ‘hypoxia gene expression classifier’ benefited from nimorazole (the hypoxic modifier, p = 0.0001). The ‘less hypoxic’ patients treated with nimorazole showed no such improvement in the cumulative incidence of locoregional tumour failure after 5 years.

The Buffa et al. reduced hypoxia signature[56] was taken forward using a Taqman Low Density Array Approach and shown to be a repeatable and reproducible biomarker for clinical application and to be prognostic in a prospective HNSCC cohort [103]. The signature was measured under Good Clinical Laboratory Practice methods in two trials of hypoxia modification: the Dutch Accelerated Radiotherapy with Carbogen and Nicotinamide (ARCON) trial in laryngeal cancer [30] and the UK Bladder Carbogen and Nicotinamide (BCON) trial in bladder cancer [29, 31]. In laryngeal cancer, those patients with a hypoxia score at or above the pre-defined median cut-off (the ‘more hypoxic’ group) had significantly improved regional control when treated with carbogen and nicotinamide in addition to accelerated radiotherapy rather than accelerated radiotherapy alone. Although the same effect was not seen in bladder cancer, the head and neck signature had borderline significance (p=0.06) for prognosis in bladder cancer and an exploratory cut-off using the upper quartile of expression to define hypoxic tumours was prognostic (p=0.02). The work highlighted the need to derive tumour site specific meta-signatures.
These studies highlight the potential for using gene signatures to stratify cancers into more and less hypoxic tumours to optimise treatment. Both the Toustrup and the reduced head and neck signatures are now undergoing prospective qualification as biomarkers in trials of hypoxia modification of radiotherapy using nimorazole in HNSCC [46].

The next development will be to prospectively allocate patients to hypoxia modifying therapy based on their hypoxia signature score. Ultimately, the aim is to use hypoxic signature scores to routinely allocate patients to hypoxia modifying treatments.

Challenges and future work

Hypoxia signatures show great variety but some commonality. Currently, there is no consensus as to the best way to derive hypoxia signatures. A potential goal could be to identify a core hypoxia signature that can be derived from the 32 hypoxia signatures. However, it is possible that tumours from different tissues are too variable in their hypoxia responses for such generalisation. Figure 3 shows an example of a gene expression heatmap of the top 65 (from a total of 4278) probesets found in at least eight out of 32 signatures and their related gene annotations in one breast cancer cohort (GSE2990, n=189), a bladder cancer cohort (GSE31684, n=93) and a HNSCC cohort (GSE6791, n=42). The genes were selected by a variety of approaches, most of which were in vitro methods. The genes have clear differences in their expression levels in the clinical samples, not only because some have low versus high expression, but also because they are differentially expressed between the three tumour types. When gene expression is standardized, the genes show different abilities to separate the samples. This shows the importance of not only assessing and refining the in vitro derived signatures in clinical samples, but also reassessing genes in each new cancer type to confirm whether they still represent the hypoxia phenotype in the specific cancer. The next step in completing this comparison requires inclusion of more tumours and testing the prognostic significance and/or predicative ability of the most frequently occurring genes in all the signatures.

A major challenge in summarising and combining signatures is gene annotation. In our systematic review of the literature, the only common annotation between all signatures was to use gene symbols. However, gene symbols change as our biological understanding evolves and therefore re-annotation from the original probesets is essential. The difference observed in this review when re-annotating existing studies using two different strategies were striking (Table 2). Summarisation would also benefit if there was curation of a central database of hypoxia signatures. Some progress along these lines has been made in the creation for example of a hypoxia database collecting and annotating genes that have been shown to be related to hypoxia (www.hypoxiadb.com), and databases collecting and re-annotating gene signatures (e.g. www.genesigdb.com).

A key difference in generating a hypoxia signature is whether to derive from cellular (in vitro) or human tumour tissue (ex vivo) data. Both approaches have advantages and disadvantages. For instance, in vitro approaches struggle to accurately reflect the tumour microenvironment, its heterogeneity, and are limited in discovering pathways associated with 3D growth. On the other hand, ex vivo biopsies are contaminated by stromal tissue which could identify false positive signals from pathways less relevant for tumours. Both approaches have generated prognostic and predictive signatures [94, 96], but increasingly as prognostic and predictive abilities are tested on
larger datasets, the in-vitro traditional approach appears to be less robust than those involving analysis of clinical samples in the derivation of hypoxia signatures. This was true both when the prognostic value of the signatures was considered in multiple independent studies [56, 95] and when the effect of slight differences in the data processing were considered [102].

Another challenge is to ascertain whether these signatures truly measure hypoxia, and even more importantly, the clinical relevant hypoxia. The gold standard method for measuring hypoxia is considered to be by oxygen electrode to measure oxygen tension directly. However, the approach takes measurements in areas of healthy cells and in areas of necrosis, which overestimates the number of viable hypoxic cells. One hypoxia classifier did use oxygen electrode measurements in its generation [94]. However, there is limited published work comparing in vivo oxygen electrode readings and hypoxia transcriptional regulation.

Gene expression signatures are being used increasingly to characterise the hypoxic status of a tumour sample. This involves measuring the expression profile of a signature in a given sample. The approach has proved to be a useful empirical tool as indicated by prognostic and predictive significance. However, in some tumours expression of these genes can be affected by factors other than hypoxia, e.g., mutations in VHL, which are common in renal cell carcinoma. The fraction of the incorrectly classified hypoxic tumours may lead to 1) inaccurate stratification and hence 2) ineffectual hypoxia modification treatment. Although this will only affect those cancers affected by such mutations, the combination of these signatures with data from mutation studies might provide a more reliable measure of hypoxia, and improve predictive ability in such cases.

Validation of signatures either as truly measuring hypoxia or maintaining prognostic and predictive significance is imperative if these signatures are to be used to stratify patients in the clinic. As new tumour expression and outcome data emerge from large retrospective cohorts and clinical trials, re-evaluation of existing hypoxia signatures as prognostic and predictive biomarkers in specific cancer types is essential. This again will help to improve the robustness of the available signatures and aid their successful translation into useful clinical tools.

**Conclusion**

Hypoxia is a key feature in many tumour microenvironments. Gene signatures can stratify patients in terms of prognosis, predict benefit from hypoxia-modifying therapies and increase understanding of the complexities of hypoxia biology. The challenges ahead include: stricter curation of signature submission and gene expression data, continued signature validation, and the crucial step of moving from retrospective to prospective classification of patients and their allocation to hypoxia modifying therapies in the clinic.
Figure 1. Schematic outlining the development of hypoxia signatures. The in-vitro ‘traditional’ approach is shown with the grey arrows. Black arrows denote alternative approaches aimed at establishing if in-vitro derived hypoxia signatures are representative of the hypoxia phenotype observed in clinical tumour samples. Double-line boxes denote steps at which the signature is trained or pruned.
Figure 2. Hypoxia signature dependencies. An arrow indicates that the signatures were related, and one was derived from the other. The first in the signature lineages is on the left. The system used for derivation is shown below the signatures.
Figure 3. Example of gene expression heatmap. Three cohorts have been considered, respectively HNSCC cohort (GSE6791, n=42, top panels), bladder cancer (GSE31684, n=93, mid panels) and breast cancer (GSE2990, n=189, lower panels). The x-axis shows the patients (samples) in each study. The y-axis reports the most occurred probesets over all those found in the 32 signatures. In particular, we chose the probesets occurred in at least 8 signatures. The vertical side bar ranks the probesets from
the most frequent (dark yellow) to the least ones (light yellow). On the left the actual expression values are shown; white corresponds to the median expression of the whole gene array and can provide an indicative threshold for calling a gene present, blue corresponds to values below the median expression (these are mostly measurements that fall below detection levels), and red above. On the right the expression is standardized per gene; white corresponds to zero, blue genes with expression below the mean down to 2 standard deviations, red above the mean up to 2 standard deviations.
Acknowledgement

This work received support from Cancer Research UK (Grant C5255/A15935) and EU project FP7-ICT-2009.5.3 (Grant 270089).

References:


