cloudy background mucosa (Figure 2D). Basal cell hyperplasia is shown in Figure 3. Proliferative cells labelled by MIB-1 were localized at the peri-IPCL. Basal cell palisading was preserved and no basal cell enlargement was detected. Spinous and surface layers were also preserved (Figure 3A–C). Figure 3 is presumed be an earlier manifestation of dysplasia, since thickened IPCLs with dilation and branching have shifted to the superficial side of the epithelium (Figure 3A). In contrast to inflammatory lesions, hyperplastic lesions showed small, well-demarcated brownish areas on NBI images with magnification (Figure 3D). Figure 4 shows mild dysplasia. Basal cell palisading was observed, though proliferative cells with enlarged nuclei, proliferating in a lamellar pattern, were limited to the lower third of the epithelial layer (Figure 4A–C). IPCL abnormalities such as IPCL extension upward, IPCL dilation and branching and IPCL diameter expansion were also recognized clearly, as well as basal cell hyperplasia (Figure 4A). The NBI image showed a wider brownish area than Figure 3D. In severe dysplasia (Figure 5), nuclear arrangement polarity was lost and nuclear density was severely increased throughout the intraepithelial layer, although there was maturation of the superficial portion of the epithelium. Microvascular irregularities were more severe than those in mild dysplasia. In the NBI image, dense, irregular IPCLs were recognized as a thicker and whitish epithelial lesion against a well-demarcated brownish area (Figure 5D). In SCC *in situ* (Figure 6), the basal cell palisading was obscure and the superficial maturation of the epithelium was lost. Prominent architectural disarray, marked cytological atypia and increased mitotic figures with pathological forms were recognized. Thickening of the intraepithelial

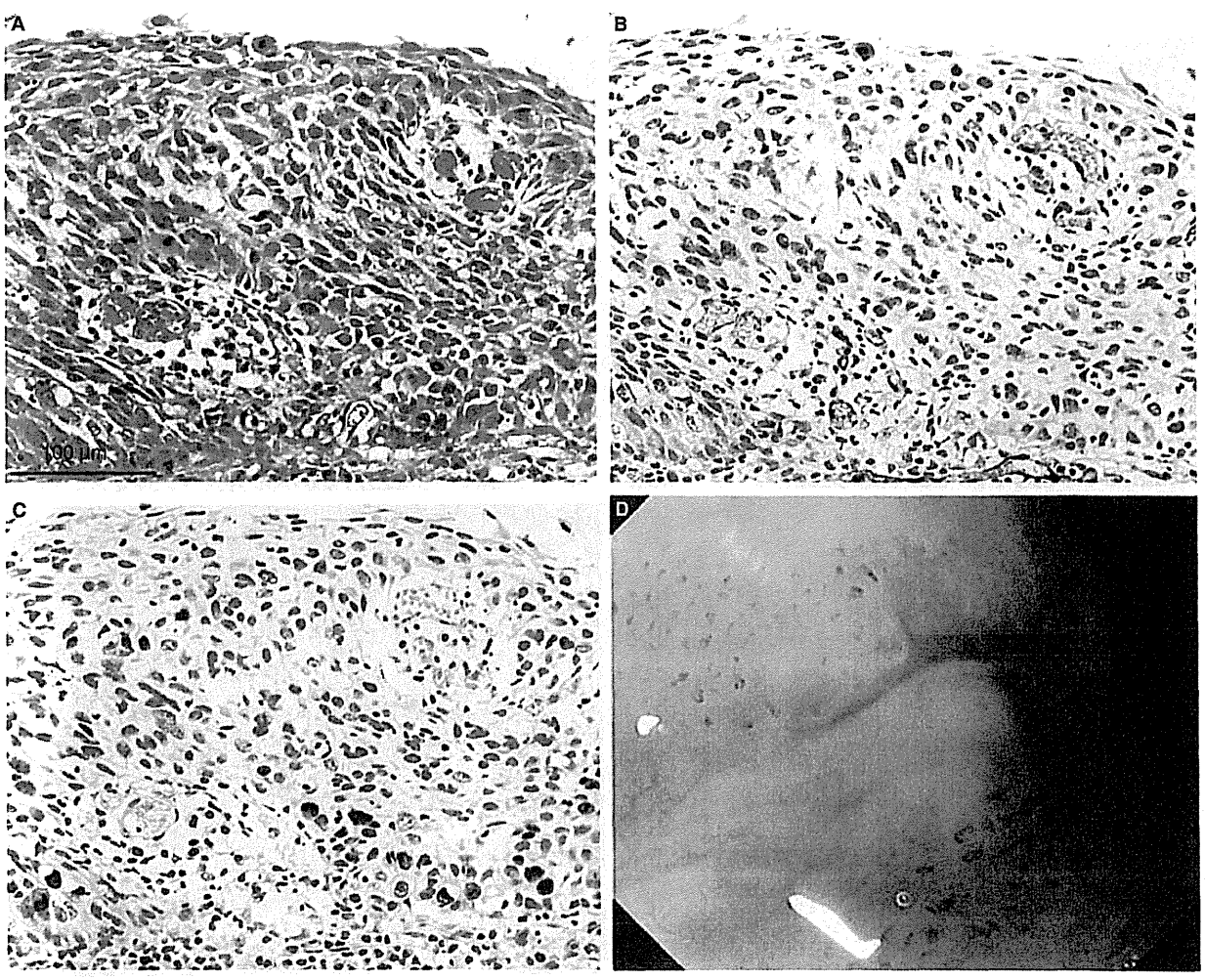

Figure 6. The histology of squamous cell carcinoma *in situ* from an endoscopically resected specimen. A, H&E. B, CD34. C, MIB-1. D, Narrow-band imaging of the hypopharynx (pyriform sinus).

lesion was also recognized as being different from that in severe dysplasia. IPCL dilation and IPCL branching-like petals were clearly recognized. The NBI image (Figure 6D) showed more dilated irregular IPCLs than Figures 4D and 5D. In invasive SCC (Figure 7), IESCC thickening and subepithelial solitary cell nests were recognized (Figure 7A–C). Dilated and branched IPCLs were recognized in the upper area of IESCC accompanied by thickening of IESCC. The NBI image showed a more voluminous and thickened brownish area (Figure 7D).

SUBEPITHELIAL INVASION, IESCC THICKNESSES AND TUMOUR THICKNESS OF INVASIVE SCC

The range of thickness of IESCC was broad: 125–1000 μm. The tumour thickness of invasive SCCs ranged from 300 to 3500 μm. The tumour thickness of all STPSCCs ranged from 125 to 3500 μm.

MVD

To calculate the MVD, immunohistochemistry using CD34 antibody was performed to detect microvessels clearly (Figures 1–7B). The values of MVD ranged from 0.1 to 17.82 (mean ± SD 4.04 ± 3.23).

RELATIONSHIPS BETWEEN HISTOLOGICAL PARAMETERS

There was a significant correlation between MVD and IESCC thickness ($P = 0.0115$, Spearman correlation test $r = 0.25$, Figure 8A). Interestingly, there was

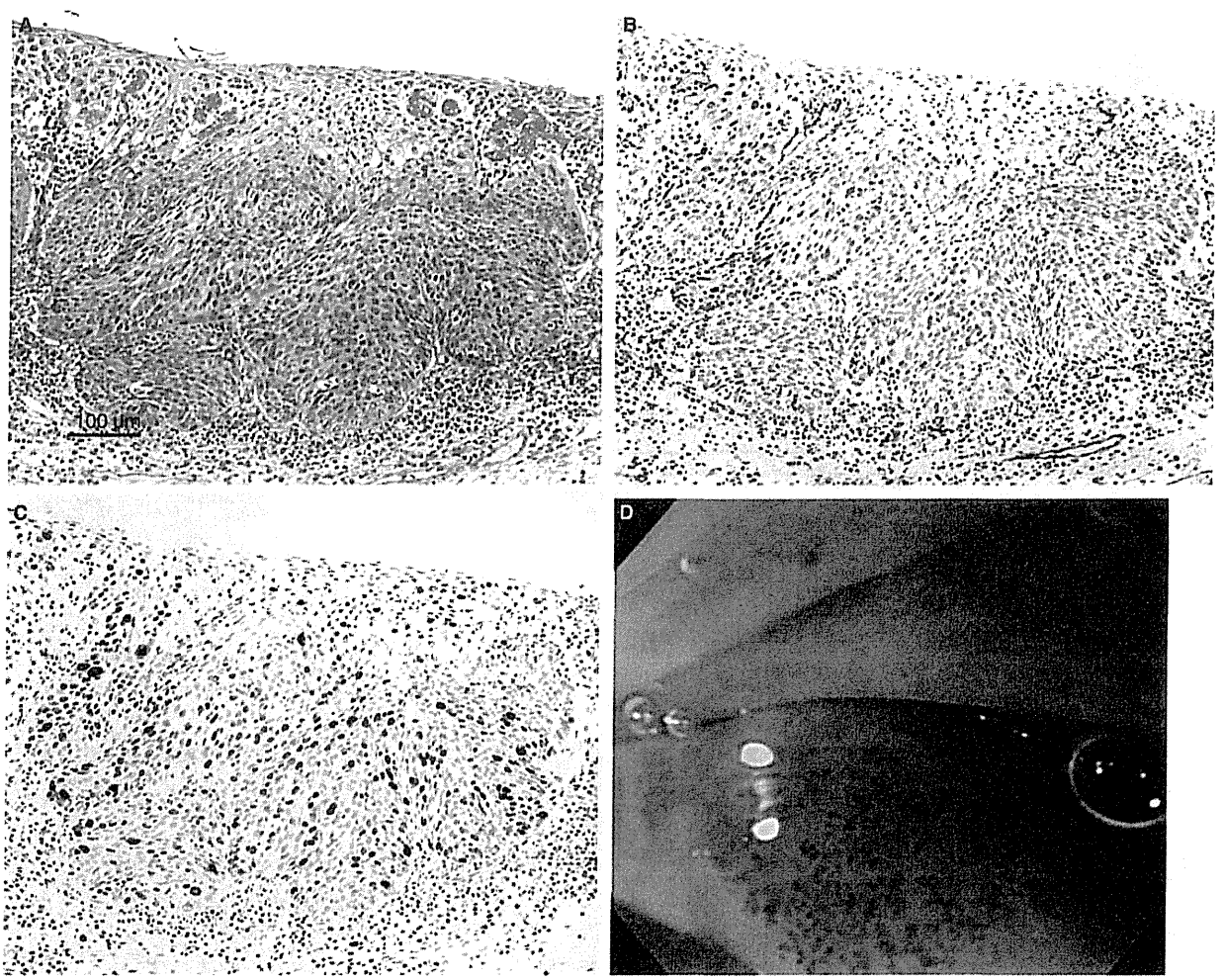

Figure 7. The histology of invasive squamous cell carcinoma from an endoscopically resected specimen. **A**, H&E. **B**, CD34. **C**, MIB-1. **D**, Narrow-band imaging of the oropharynx (lateral wall).

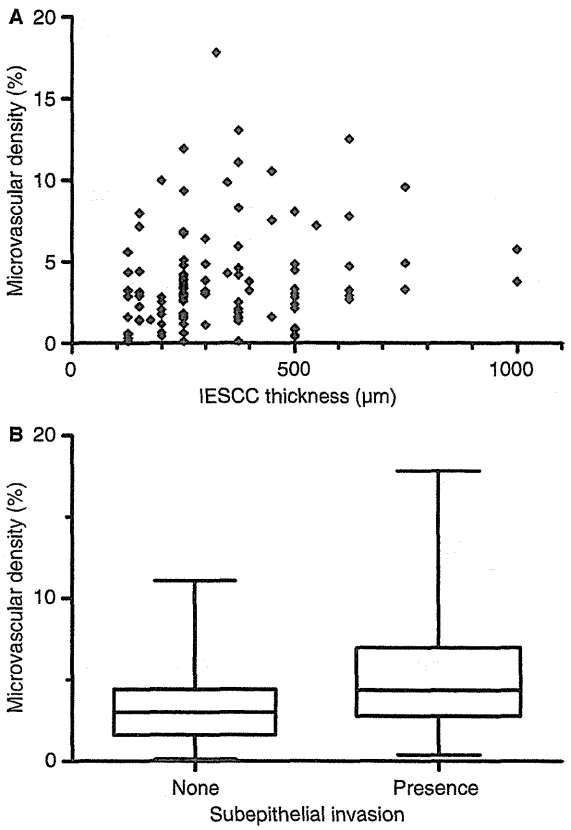

Figure 8. The relationships between intraepithelial squamous cell carcinoma (IESCC) thickness, microvascular density (MVD), and subepithelial invasion. **A**, The relationship between MVD and IESCC thickness. **B**, The relationship between MVD and subepithelial invasion. A *P*-value < 0.05 was taken to indicate a statistically significant difference.

significant correlation between MVD and subepithelial invasion ($P = 0.0078$, Figure 8B).

Next, we investigated the relationships between IESCC thickness, tumour thickness, tumour size and subepithelial invasion. There was significant correlation between IESCC thickness and subepithelial invasion ($P < 0.0001$, Figure 9A). There was significant correlation between tumour thickness and subepithelial invasion ($P < 0.0001$, Figure 9B). These correlations suggested that SCC detected by NBI endoscopy shows increased MVD due to microvascular irregularities, which then leads to greater IESCC thickness, resulting in subepithelial invasion.

The cases with subepithelial invasion had significantly larger tumours than those without subepithelial invasion ($P = 0.0013$, Figure 9C). There was significant correlation between tumour thickness and tumour size (Spearman $r = 0.4560$, $P < 0.0001$, Figure 9D).

As described above, there were significant correlations between tumour thickness, MVD and subepithelial invasion. We investigated the relationship between tumour thickness and vessel infiltration, including invasive and non-invasive carcinomas. As shown in Figure 9E, there was significant correlation between tumour thickness and vessel infiltration ($P < 0.0001$). Only one STPSCC (0.96%) showed lymph node metastasis as a late event (Table 2). The tumour thickness in this case was 1750 μm and both lymphatic and blood vessel infiltration were observed in the primary tumour. The mean tumour thickness of STPSCCs resected by EMR/ESD was 1040 μm. Three cases (2.9%) showed local recurrence after EMR/ESD. All three tumours were <20 mm in diameter but showed subepithelial invasion, and one of the three had blood vessel infiltration. The horizontal cut ends of specimens by EMR/ESD were positive for carcinoma cells.

Discussion

The dense microvascular proliferation caused by irregular branching of IPCLs, the upward shift and thickening of IPCLs, which reflect microvascular irregularities detected by NBI endoscopy, were observed pathologically in squamous epithelial lesions of the pharynx. The alterations of microvascular structures represented by IPCL irregularities occurred with architectural or cytological abnormalities in squamous epithelial lesions (Table 3). Based on the current results, we propose four steps to invasive SCC starting with basal cell hyperplasia. In the first step, the upward shift of IPCLs with dilation and branching occurs in basal cell hyperplasia (Figure 3). In the second step, density of IPCL increases in mild dysplasia (Figure 4). In the third step, microvascular irregularities increase with branching of IPCLs (Figure 5). Moreover, marked IPCL proliferation with irregular branching is observed throughout the intraepithelial layer. Thickening of intraepithelial lesions becomes apparent (Figure 6). Finally, in the fourth step, dilated and branched IPCL increases in intraepithelial lesions, accompanied by increasing MVD, resulting in subepithelial invasion (Figure 7).

The results of the present study suggest that IESCC thickening by increased MVD plays an important role in the early stage of subepithelial invasion (Figures 8 and 9). On the other hand, the average value of MVD of dysplasia (including mild and severe) was 2.9, which was lower than the average value of MVD of IESCC (4.0). The value of MVD of single basal cell hyperplasia was 2.1. This suggests that MVD increases step by step from basal cell hyperplasia through dysplasia to

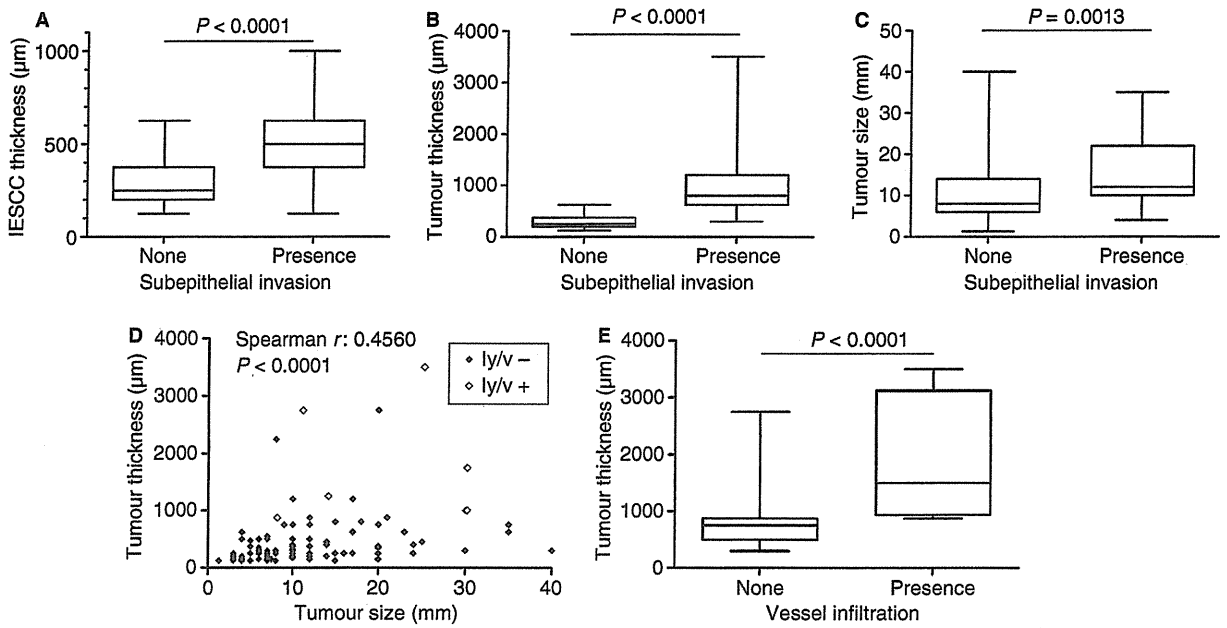

Figure 9. The relationships between intraepithelial squamous cell carcinoma (IESCC) thickness, tumour thickness, tumour size, subepithelial invasion, and vessel infiltration. **A**, The relationship between IESCC thickness and subepithelial invasion. **B**, The relationship between tumour thickness and subepithelial invasion. **C**, The relationship between tumour size and subepithelial invasion. **D**, The relationship between tumour thickness and tumour size. Black and white diamond symbols represent a case without lymphatic and venous infiltration and a case with lymphatic or venous infiltration, respectively. **E**, The relationship between tumour thickness and vessel infiltration in superficial-type pharyngeal squamous cell carcinomas (STPSCCs) with subepithelial invasion (n = 29). A P-value <0.05 was taken to indicate a statistically significant difference.

carcinoma. An interesting previous report examined MVD in epithelial proliferative lesions of breast.[12] In florid ductal hyperplasia usual type, atypical ductal hyperplasia and atypical lobular hyperplasia, MVD values were intermediate between normal glandular structures and neoplastic lesions. Moreover, the MVD of higher-grade-type intraductal carcinoma was higher than that of lower-grade-type intraductal carcinoma. These findings suggest that an increase in blood supply is necessary for any type of epithelial proliferation, and higher MVD values are associated with a higher risk of progression to invasiveness in pre-invasive breast cancer as in intraepithelial squamous lesions in the present study. However, because the number of precursor lesions in the present study was small, further detailed analysis is needed in the future.

There were a few tumours showing remarkable intraepithelial spreading without either subepithelial invasion of the STPSCC or thickening of IESCC. The five tumours without subepithelial invasion were > 20 mm (20 mm <), and these tumours were categorized as T2 in the International Union Against Cancer classification.[13] The cases were assumed to be good candidates for endoscopic resection of STPSCC. As shown in Figure 9E, there was significant correlation between tumour thickness and vessel infiltration in STPSCC. However, as shown in Figure 9E, the greater the tumour thickness, the higher rate of vessel infiltration that was observed. These results suggest that the T factor incorporating tumour thickness may be reasonable, as well as being a qualified risk factor, for clinicians when contemplating additional treatments such as lymph node dissection or selecting intense follow-up after EMR/ESD.

Multiple STPSCCs at oropharyngeal and hypopharyngeal mucosal sites were identified in 30 patients (43.5%) in the current study. This means that the same patient would need additional EMR/ESD. Numerous EMR/ESD can cause future stenosis of the pharynx. From the view of a preventive strategy, it is important to identify the abnormal proliferative factor influencing IPCL in STPSCC, and further future analyses are necessary.

The chance of diagnosing an early squamous epithelial lesion with microvascular irregularities detected by NBI endoscopy has improved and facilitates an understanding of the pathological features necessary for appropriate diagnosis. The present study has

shown microvascular irregularities to be associated with carcinogenesis of SCC. A better understanding of the pathogenesis of squamous epithelial lesions and early invasion of SCC should facilitate appropriate diagnosis and treatment of STPSCC.

Acknowledgements

We thank Ms Mai Okumoto for her excellent technical assistance in immunohistochemistry. This study was supported in part by a Grant-in-Aid for Cancer Research (19-10) from the Ministry of Health, Labour, and Welfare of Japan and a Grant-in-Aid for the Third Term Comprehensive 10-year Strategy for Cancer Control from the Ministry of Health, Labour, and Welfare of Japan.

References

1. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. *Int. J. Cancer* 1999; **80**; 827–841.
2. Nonaka S, Saito Y. Endoscopic diagnosis of pharyngeal carcinoma by NBI. *Endoscopy* 2008; **40**; 347–351.
3. Gono T, Yamazaki K, Doguchi N *et al.* Endoscopic observation of tissue by narrow band illumination. *Opt. Rev.* 2003; **10**; 211–215.
4. Inoue H, Honda T, Nagai K *et al.* Ultra-high maginification endoscopic observation of carcinoma *in situ* of the oesophagus. *Dig. Endosc.* 1997; **9**; 16–18.
5. Kumagai Y, Inoue H, Nagai K *et al.* Magnifying endoscopy, stereoscopic microscopy, and the microvascular architecture of superficial esophageal carcinoma. *Endoscopy* 2002; **34**; 369–375.
6. Barnes L, Eveson JW, Reichart P *et al.* eds. *Pathology and genetics of head and neck tumours: the World Health Organization classification of tumors*. Lyons: International Agency for Research on Cancer (IARC) Press, 2005.
7. Kodama M, Kakegawa T. Treatment of superficial cancer of the esophagus: a summary of responses to a questionnaire on superficial cancer of the esophagus in Japan. *Surgery* 1998; **123**; 432–439.
8. Asakage T, Yokose T, Mukai K *et al.* Tumor thickness predicts cervical metastasis in patients with stage I/II carcinoma of the tongue. *Cancer* 1998; **82**; 1443–1448.
9. Lim SC, Zhang S, Ishii G *et al.* Predictive markers for late cervical metastasis in stage I and II invasive squamous cell carcinoma of the oral tongue. *Clin. Cancer Res.* 2004; **10**; 166–172.
10. Muto M, Nakane M, Katada C *et al.* Squamous cell carcinoma *in situ* at oropharyngeal and hypopharyngeal mucosal sites. *Cancer* 2004; **101**; 1375–1381.
11. Barnes L. *Surgical pathology of the head and neck*, 2nd edn. New York, Basel: Informa Healthcare, Marcel Dekker, 2001.
12. Viacava P, Naccarato AG, Bocci G *et al.* Angiogenesis and VEGF expression in pre-invasive lesions of the human breast. *J. Pathol.* 2004; **204**; 140–146.
13. Sobin LH, Wittekind C eds. *International Union against Cancer TNM classification of malignant tumours*, 6th edn. New York: Wiley-Liss, 2002.

厚生労働科学研究費補助金

第３次対がん総合戦略研究事業

「QOLの向上をめざしたがん治療法の開発研究」

に関する研究

平成22年度～平成25年度　総合研究報告書

研究代表者　　江角　浩安

平成26（2014）年　5月

平成23年度

研究成果の刊行物・別刷

HUGO J (2010) 4:35–48
DOI 10.1007/s11568-011-9150-9

RESEARCH ARTICLE

Genome-wide identification and annotation of HIF-1α binding sites in two cell lines using massively parallel sequencing

Kousuke Tanimoto · Katsuya Tsuchihara · Akinori Kanai · Takako Arauchi · Hiroyasu Esumi · Yutaka Suzuki · Sumio Sugano

Received: 15 November 2010 / Revised: 27 January 2011 / Accepted: 1 February 2011 / Published online: 19 February 2011

Abstract We identified 531 and 616 putative HIF-1α target sites by ChIP-Seq in the cancerous cell line DLD-1 and the non-cancerous cell line TIG-3, respectively. We also examined the positions and expression levels of transcriptional start sites (TSSs) in these cell lines using our TSS-Seq method. We observed that 121 and 48 genes in DLD-1 and TIG-3 cells, respectively, had HIF-1α binding sites in proximal regions of the previously reported TSSs that were up-regulated at the transcriptional level. In addition, 193 and 123 of the HIF-1α target sites, respectively, were located in proximal regions of previously uncharacterized TSSs, namely, TSSs of putative alternative promoters of protein-coding genes or promoters of putative non-protein-coding transcripts. The hypoxic response of DLD-1 cells was more significant than that of TIG-3 cells with respect to both the number of target sites and the degree of induced changes in transcript expression. The Nucleosome-Seq and ChIP-Seq analyses of histone modifications revealed that the chromatin formed an open structure in regions surrounding the HIF-1α binding sites, but this event occurred prior to the actual binding of HIF-1α. Different cellular histories may be encoded by chromatin structures and determine the activation of specific genes in response to hypoxic shock.

Keywords HIF-1 alpha · Hypoxia · ChIP-Seq · Transcriptome

Introduction

Cancer cells in solid tumors are frequently exposed to limited oxygen environments because aberrantly growing cancer cells can cause a shortage of blood flow (Pouyssegur et al. 2006; Harris 2002; Keith and Simon 2007). Hypoxia is a term that is used to describe such cellular environments, which have an oxygen concentration between 0.02 and 3% (Thomas and Johannes 2007). It is important to understand the behavior of cancer cells under hypoxia because hypoxic conditions can be utilized to develop a therapeutic method to treat cancers (Fyles et al. 1998; Nordsmark et al. 1996). If cancer cells were vulnerable to hypoxia, they would lose their ability to survive. In addition, controlling the hypoxic responses of cells may lead to new clinical therapeutics not only for solid tumors (Semenza 2003; Semenza 2010) but also for tissue ischemia and several inflammatory diseases (Fraisl et al. 2009; Semenza 2000).

Tumor cells adapt to hypoxic conditions through modifications of their gene expression patterns. For example, ATP metabolism changes from oxidative phosphorylation to anaerobic glycolysis. During this process, glucose uptake is enhanced, angiogenesis is induced to obtain a sufficient blood supply, and cell migration is promoted (Harris 2002; Pouyssegur et al. 2006; Gatenby and Gillies 2004; Hirota and Semenza 2006; Keith and Simon 2007). The central regulators of these changes in gene expression, which collectively lead to changes in cellular responses,

Electronic supplementary material The online version of this article (doi:10.1007/s11568-011-9150-9) contains supplementary material, which is available to authorized users.

K. Tanimoto · A. Kanai · T. Arauchi · Y. Suzuki (✉) · S. Sugano
Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
e-mail: ysuzuki@k.u-tokyo.ac.jp

K. Tsuchihara · H. Esumi
Cancer Physiology Project, Research Center for Innovative Oncology, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa, Chiba 277-8577, Japan

are the transcription factors hypoxia-inducible factors 1 and 2 (HIF-1/2) (Wang and Semenza 1995; Tian et al. 1997). Both HIF-1 and HIF-2 consist of two subunits, HIF-α and HIF-β (or ARNT). The α-subunit (HIF-1α, HIF-2α) is the oxygen-sensitive subunit. Under hypoxic conditions, HIF-1α (or HIF-2α) translocates from the cytoplasm to the nucleus and dimerizes with HIF-β. The HIF heterodimer complex binds to a sequence motif called the hypoxia response element (HRE) and activates the expression of genes that are necessary for cellular survival under hypoxic conditions (Maxwell 2005; Poon et al. 2009). Although HIF-1α demonstrates extensive sequence similarity to HIF-2α and is regulated via the same mechanism, HIF-1 and HIF-2 are assumed to possess distinct biological roles. HIF-2α expression is restricted to particular cell types, whereas HIF-1α is expressed ubiquitously (Wiesener et al. 2003). Additionally, genes encoding glycolytic enzymes such as PGK1 and GLUT1 are unique targets of HIF-1 (Wenger et al. 2005; Raval et al. 2005).

Because of the inferred important roles of HIF-1α, there have been several attempts to comprehensively identify HIF-1α targets (Ortiz-Barahona et al. 2010). Although computational methods alone may still require further innovation due to the ambiguity of the HRE motif, the combination of computational methods and expression analyses using microarrays has proven to be a powerful approach (Benita et al. 2009). More recently, several papers have reported the identification of HIF-1α binding sites based on chromatin immunoprecipitation and microarray (ChIP-chip) technology (Mole et al. 2009; Xia et al. 2009; Xia and Kung 2009). These studies have reported that about 400–500 of the HIF-1α binding sites in the human genome and approximately 70% of the HIF-1α binding sequences contain HREs.

Recent microarray-based approaches have produced far more information regarding HIF-1 targets than have conventional methods. However, these microarray methods have several limitations; namely, they cannot be used to analyze previously uncharacterized genes, for example, intergenic transcripts or transcripts of alternative promoters (Davuluri et al. 2008; Landry et al. 2003). In addition, the resolution of the data does not exceed the intervals of the DNA probes. The recently developed massively parallel sequencing technologies have provided a potential means to improve upon these methods. For example, the Illumina GA sequencer can sequence 200–250 million sequences per run. By utilizing Illumina GA and ChIP technologies (ChIP-Seq), the binding sites of several transcription factors have been reported (Valouev et al. 2008; Robertson et al. 2007; Kwon et al. 2009; Welboren et al. 2009; Park 2009). On the other hand, we also developed a method for sequencing the 5′-ends of mRNAs by combining our oligo-capping technique with the Illumina GA sequencer (Tsuchihara et al. 2009). This method, which we named TSS-Seq, enables us to obtain precise information regarding the positions of the transcriptional start sites (TSSs) together with their expression levels in a high-throughput manner. In our previous report, we applied TSS-seq to observe gene expression changes in response to hypoxia in several cell lines (Tsuchihara et al. 2009).

To identify TSSs with hypoxic responses elicited directly by HIF-1α binding, we performed HIF-1α ChIP analysis and combined the resultant data with the TSS-Seq data. Furthermore, we carried out Illumina GA sequencing analysis using micrococcal nuclease-digested genomic DNA (Nucleosome-Seq) (Albert et al. 2007; Jiang and Pugh 2009; Schones et al. 2008). We examined the changes in nucleosome structure caused by HIF-1α binding and the subsequent transcriptional activation. We also examined the chromatin statuses surrounding the HIF-1α binding sites by ChIP-Seq analysis of RNA polymerase II (pol II) and a series of histone modifications, namely, H3K4 trimethylation (H3K4me3), H3 acetylation (H3Ac) and H3K27 trimethylation (H3K27me3) (Barski et al. 2007). H3K4me3 and H3Ac are known to be characteristic of an open chromatin structure, and H3K27me3 is characteristic of a closed chromatin structure. A colorectal cancer epithelial cell line (DLD-1) and a normal embryonic lung fibroblast cell line (TIG-3) were used for these analyses. Here, we report the genome-wide identification and characterization of HIF-1 targets using several types of massively parallel sequencing analyses. We further compared the observed targets between DLD-1 cells and TIG-3 cells to examine the cellular responses characteristic of these cells.

Materials and methods

Accession numbers

The short-read tag sequences used in the present work have been deposited in DDBJ under the following accession numbers: SRA003625 (TSS-Seq data), DRA000007-000008 and DRA000190-000191 (pol II ChIP-Seq data), DRA000073-000075 and DRA000003 (Nucleosome-Seq data), DRA000178-000185 (HIF-1α ChIP-Seq data), DRA000277-000284 (H3K4me3 ChIP-Seq data), DRA000285-000288 and DRA000293-000296 (H3Ac ChIP-Seq data) and DRA000289-000292 and DRA000297-000300 (H3K27me3 ChIP-Seq data).

The data generated in this study are accessible via following URL in BED file format: http://www.dbtss.hgc.jp/cgi-bin/downloader2.cgi/chip.tar.gz (ChIP-Seq data), http://www.dbtss.hgc.jp/cgi-bin/downloader2.cgi/nucleosome_bed.tar.gz (Nucleosome-Seq data), http://www.dbtss.hgc.jp/

cgi-bin/downloader2.cgi/rnaseq.tar.gz (RNA-Seq data) and ftp://ftp.hgc.jp/pub/hgc/db/dbtss/dbtss_ver7/bed_files/ (TSS-Seq data).

Cell culture

The human colorectal cancer cell line DLD-1 (Japan Health Science Research Resources Bank cell number JCRB9094) was maintained in Dulbecco's Modified Eagle's Medium (DMEM, Nissui) Supplemented with 10% fetal calf serum, 4.5 g/l glucose, and antibiotics. The human normal embryonic lung fibroblast cell line TIG-3 (Japan Health Science Research Resources Bank cell number JCRB0506) was maintained in MEM (GIBCO) Supplemented with 10% fetal calf serum and antibiotics. The cells were cultured under standard conditions and subjected to hypoxic shock (1% O_2) in a similar manner.

Construction of the ChIP-seq library and analysis of ChIP-seq tags

DLD-1 and TIG-3 cells (1×10^8) were fixed using a final concentration of 1% formaldehyde at room temperature for 10 min. To quench the reaction, 150 mM (final concentration) glycine was added to the samples followed by a 5-min incubation. The cells were washed twice with PBS, harvested and then lysed in 5 ml of Lysis Buffer 1 (50 mM HEPES–KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100). The lysates were incubated at 4°C for 10 min and centrifuged at 1,500 rpm for 5 min at 4°C. The pellets were then resuspended in 5 ml of Lysis Buffer 2 (10 mM Tris–HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), incubated at room temperature for 10 min, and centrifuged at 1,500 rpm for 5 min at 4°C. The pellets were resuspended in 1 ml of Lysis Buffer 3 (10 mM Tris–HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-deoxycholate, 0.5% *N*-lauroylsarcosine) and sonicated with 18 cycles of 30 s each on ice using a sonicator (TOMY SEIKO). One hundred microliters of 10% Triton-X 100 was added to the samples, and the cell lysates were centrifuged for 10 min. Fifty microliters of the supernatant was saved for the controls (named whole cell extract (WCE) DNA). Washed magnetic beads bound to 10 μg of rabbit polyclonal anti-HIF-1α antibody (Novus Biologicals, Inc., NB100-134), monoclonal anti-RNA polymerase II antibody (Abcam, ab817), monoclonal anti-H3K4me3 antibody (ab1012), polyclonal anti-H3Ac antibody (Millipore, 06-599) or monoclonal anti-H3K27me3 antibody (ab6002) was added to the supernatant. The samples were rotated at 4°C overnight and washed 8 times with 1 ml of wash buffer (50 mM HEPES–KOH, pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-deoxycholate) and once with TE buffer containing 50 mM NaCl. The sample was then eluted with 200 μl of elution buffer (1 M Tris–HCl, pH 8.0, 0.5 M EDTA, pH 8.0, 1% SDS) at 65°C for 15 min. The supernatant was transferred to a new tube and incubated at 65°C overnight. Concurrently, 150 μl of elution buffer was added to the saved WCE-DNA, and the samples were incubated at 65°C overnight. Next, 200 μl of TE and 8 μl of 10 mg/ml RNase A (Funakosi) were added to the IP and WCE DNA samples. The samples were then incubated at 37°C for 2 h. Subsequently, 4 μl of 20 mg/ml proteinase K (Takara) was added to the samples, which were then incubated at 55°C for 2 h. The DNA samples were purified using TRIzol LS (Invitrogen) and ethanol precipitation. Samples destined for ChIP-seq by Illumina GA were prepared according to the manufacturer's instructions.

Computational procedures

The generated ChIP-Seq tags were mapped to the reference human genome sequences (hg18), and the tags that were uniquely mapped allowing two-base mismatches were used. The binding sites of HIF-1α, RNA polymerase II and histone modifications, based on the short-read tag information, were identified as follows: the regions encompassed by the mapped tag sequences were extended to 120 bp, which reflected the insertion sites of the sample DNA fragments. For each genomic position, the number of overlapping extended tags was counted. Based on the calculated tag information, the sum of the included tags was evaluated to determine whether more than a tenfold difference between the immunoprecipitation (IP) and the whole cell extract (WCE) was present. Genomic regions in which positive enrichment of the tags continued for more than 121 bp were then selected. Among these regions, those containing more than 5 different sequence tags were identified as potential HIF-1α and RNA polymerase II binding sites. We normalized the tag counts for each region to tags per million (ppm) and divided the results by the length of the region (ppm/bp). These values were determined as the peak strength in this study. In the HIF-1α ChIP-Seq analysis, if the identified regions demonstrated an overlap between 1% O_2 (hypoxia) and 21% O_2 (normoxia), regions with more than a twofold (hypoxia/normoxia) change in peak strength were identified as "bindings enriched in hypoxia." The statistical significance determined for this selection procedure compared to the background rate was evaluated using Poisson probabilities as previously described (Robertson et al. 2007):$P(x, \lambda) = 1 - \sum_{t=0}^{x} \frac{e^{-\lambda}\lambda^t}{t!}$,where $P(x,\lambda)$ is the probability of enrichment, λ is the expected tag number in the 121-bp

window calculated for the WCE sample, and x is the observed tag number in the 121-bp window.

To analyze the gene expression changes in the transcription factors in DLD-1 and TIG-3 cells, 140 transcription factors were selected in TRANSFAC (Rel. 2010.1). The consensus sequences of the transcription factors were also evaluated using MATCH with cutoff values determined using minFP, which minimizes false-positive results. The enrichment of the detected putative binding sites in the HIF-1α target genes in DLD-1 cells, TIG-3 cells and both cell lines were evaluated by calculating the hypergeometric distributions. Putative transcription factor binding sites that were statistically significantly enriched ($P < 0.05$) in the respective groups were selected. For details, see Supplementary Table 3.

TSS-Seq analysis and RNA-Seq analysis of polysome fractions

TSS-Seq tags were prepared and analyzed as described in the reference (Tsuchihara et al. 2009). After clustering the TSS-tags using 500-bp bins, representative TSSs were selected as the position from which the largest number of TSS tags was mapped. The selected representative TSSs were used for further analyses.

For the RNA-Seq analysis of polysome-incorporated RNA, the cytoplasmic fraction was isolated from DLD-1 cells using the PARIS Kit (Ambion, AM1921). Polysome-incorporated fractions were further fractionated from cytoplasmic fraction using the sucrose-density-gradient method described in the reference (Koritzinsky and Wouters 2007). Thirty-six-base-pair single-end-read RNA-Seq tags were generated using an Illumina GA sequencer according to the standard protocol. RNA-Seq tags that were mapped to the human reference genome sequences without any mismatches were used. RNA-Seq tags corresponded to RefSeq transcripts when their genomic coordinates overlapped. Details of the characterization and the RNA-Seq libraries are presented in Supplementary Fig. 6.

Construction of the nucleosome-Seq library and analysis of nucleosome tags

Cells were cultured to 80% confluence and treated with micrococcal nuclease to generate mononucleosomes using the ChIP-IT Express Enzymatic Kit (active motif). The cells were formaldehyde-cross-linked prior to the nucleosome isolation. Cross-linking was performed with fixation solution (DMEM, 1% formaldehyde) for 10 min at room temperature. The cells were then washed with PBS, and the cross-linking reaction was stopped by incubating the cells with a glycine quenching solution (PBS glycine buffer) for 5 min at room temperature. After the dishes were washed with PBS, an ice-cold cell scraping solution was added. The cells were then harvested and resuspended in ice-cold lysis buffer, incubated on ice for 30 min, and homogenized by 15–20 strokes with a Dounce homogenizer. The nuclei were suspended in digestion buffer and pre-warmed in this solution for 5 min at 37°C. An enzymatic cocktail (200 U/ml) was added to the sample, and the digestion reaction was incubated at 37°C for 15 min. After the reaction was stopped with EDTA, the nuclei were pelleted by centrifugation, and the supernatants were collected. Formaldehyde cross-linking was reversed by the addition of 5 M NaCl and RNase followed by an incubation at 65°C for more than 4 h. Proteinase K was added to the samples, and the mixtures were incubated at 42°C for 1.5 h. DNA was purified by phenol/chloroform extraction and ethanol precipitation. The extracted DNA was used for the sample preparation for Illumina GA according to the manufacturer's instructions.

Real-time PCR analysis

Real-time PCR primers were designed within the identified HIF-1α binding sites. Primer3 (http://www.frodo.wi.mit.edu/primer3/) was used to design the primers. The fold enrichment at each locus was calculated according to the ΔΔCt method. The primer sequences are presented in Supplementary Table 1.

Results and discussion

Identification of HIF-1α binding sites

Using 22,315,752 and 15,389,864 36-bp single-end-read sequence tags generated for the ChIP-Seq analysis of DLD-1 and TIG-3 cells (Supplementary Fig. 1), we searched for HIF-1α binding sites by quantifying the local enrichment of tags (Robertson et al. 2007) (see "Materials and methods"). Next, we classified the putatively identified binding sites as "hypoxia-responsive" or "non-hypoxia-responsive." "Hypoxia-responsive" binding sites were the regions identified only under hypoxia or the regions in which the associated ChIP-Seq tags were enriched more than twofold under hypoxia. "Non-hypoxia-responsive" binding sites were those that did not show more than a twofold change in tag concentrations between normoxia and hypoxia. A typical example for each category is shown in Fig. 1a. As summarized in Table 1, there were 531 and 616 "hypoxia-responsive" HIF-1α binding sites in DLD-1 and in TIG-3 cells, respectively. As shown in Fig. 1b, we selected and validated 29 of the indentified HIF-1α binding sites by real-time PCR. We confirmed that 28 cases (97%) showed at least a fourfold signal enrichment in the immunoprecipitated

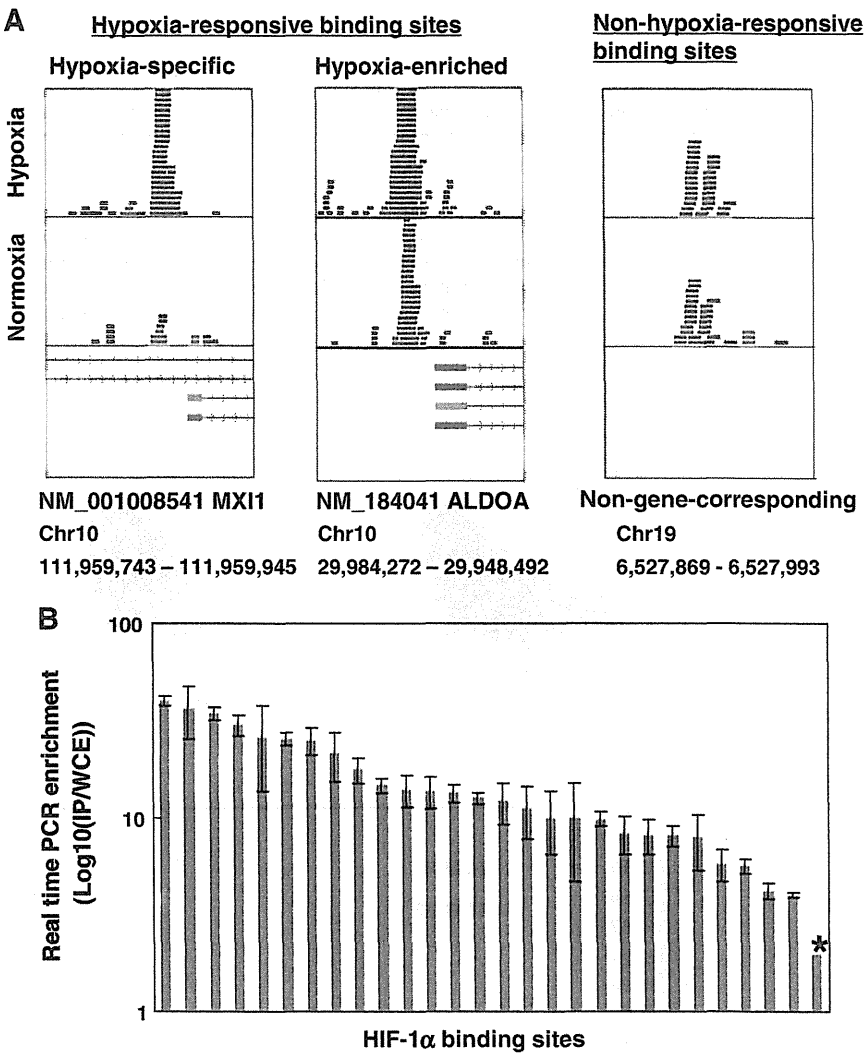

Fig. 1 Identification of HIF-1α binding sites. **a** Representative sites identified as "hypoxia-responsive", "hypoxia-enriched" and "non-hypoxia-responsive" HIF-1α binding sites. The definitions for each group are described in the text. The former two categories were considered for identifying HIF-1α target genes. *Blue* and *red* horizontal bars indicate 36-bp sequence tags that were mapped to the forward and reverse strands, respectively. The genomic coordinates and RefSeq IDs of the indicated transcript models are shown in the margin. **b** Validation of the identified binding sites by ChIP-Seq analysis using real-time PCR. The fold enrichment of the signals determined for immunoprecipitated samples compared with the background noise is shown on a log scale for 29 selected cases. Genomic regions as specified in Supplementary Table 1 were used for the validation analysis. For details, see "Materials and methods" and Supplementary Fig. 2. Error bars represent the standard deviations of triplicate experiments. (Color figure online)

samples compared with the background noise (see also Supplementary Fig. 2 for agarose gel images of the PCR bands).

Using the identified "hypoxia-responsive" HIF-1α binding sites, we searched for the HRE motif. We found that 441 (83%) and 413 (67%) regions contained the HRE motif in DLD-1 and in TIG-3 cells, respectively (fourth column, Table 1). These frequencies were similar to those reported in previous HIF-1α ChIP-chip studies (Mole et al. 2009; Xia et al. 2009), although the detected genomic regions containing the HIF-1α binding sites were shorter in length due to the improved resolution of the analysis (the lengths of the detected binding sites were more than threefold shorter than those detected in previous studies using ChIP-chip, which reflected the intervals of the designed DNA probes (Mole et al. 2009); see also Supplementary Fig. 3a). Occasionally, two or three adjacent core HRE motifs were detected in a single gene, such as the genes encoding transferrin and the glucose transporter (Wenger et al. 2005). We examined the number of HRE motifs and the ChIP-Seq tag concentrations and found no correlation between them (Supplementary Fig. 3b). The HRE motif plays a role in defining the binding positions but may not contribute directly to the binding strength.

Transcripts in the proximal regions of the identified HIF-1α binding sites

In DLD-1 and TIG-3 cells, 193 (37%) and 161 (26%) of the "hypoxia-responsive" HIF-1α binding sites, respectively, were located from 10 kb upstream to 1 kb downstream of the 5′-ends of the RefSeq transcript models (23771 transcript models in 18,001 genes in total). The locations of the binding sites relative to the RefSeq genes are summarized in Venn diagrams in Fig. 2a. For these binding sites, we analyzed the distributions of the relative distances from the

Table 1 Number of HIF-1α binding sites

	Hypoxia-responsive binding sites				Non-hypoxia-responsive binding sites
	Hypoxia-specific	Hypoxia-enriched	Total binding sites	Sites containing HRE	
DLD-1	499	32	531	441 (83%)	35
TIG-3	599	17	616	413 (67%)	54

The number of identified HIF-1α binding sites in DLD-1 and TIG-3 cells. "Hypoxic-responsive" HIF-1α binding sites are the sum of the "hypoxia-specific" binding sites, which were detected only under hypoxia, and the "hypoxia-induced" binding sites, which were detected under normoxia, but the ChIP-Seq tag concentration was induced more than twofold under hypoxia. Note that most of the identified "hypoxia-responsive" HIF-1α binding sites are "hypoxia-specific." *HRE* hypoxia response element, 5′-RCGTG-3′

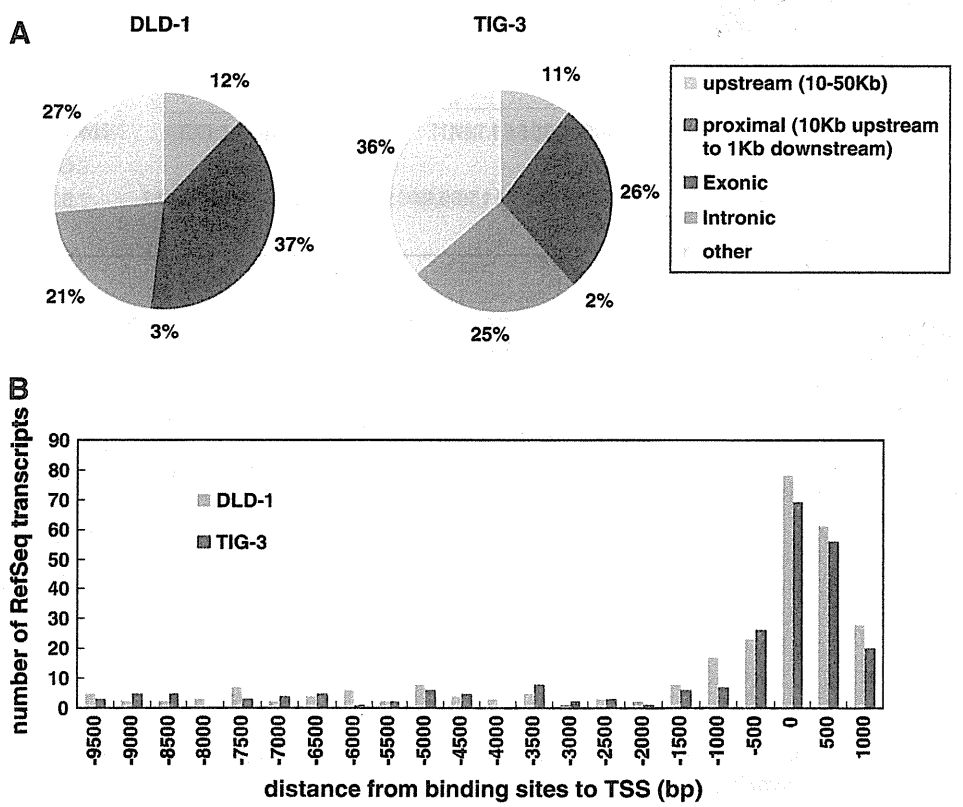

Fig. 2 Characterization of RefSeq transcripts. **a** Genomic locations of the identified HIF-1α binding sites in DLD-1 and TIG-3 cells compared to the RefSeq transcript models. Positions relative to the RefSeq genes are shown. When a RefSeq gene contained multiple transcript models, all of the transcript models were considered. **b** Distributions of the distances from the identified binding sites to the 5′-ends of the RefSeq transcript models. *Blue* and *red bars* represent DLD-1 and TIG-3 cells, respectively. When a RefSeq gene contained multiple transcript models, all of the transcript models were considered and counted redundantly. (Color figure online)

5′-ends of each RefSeq transcript model (Fig. 2b). When a RefSeq gene contained multiple transcript variant models, all of the variants were considered and counted redundantly. We found that approximately 70% of the binding sites were located within 1 kb of the 5′-ends of the RefSeq transcript models, which is consistent with previous estimations (Lin et al. 2009; Mole et al. 2009).

To directly examine the positions of the TSSs and their expression changes in response to HIF-1α binding in DLD-1 and TIG-3 cells, we used TSS-Seq data. This dataset contains a total of 141,590,200 TSS tags, each of which corresponds to the cap site of a single mRNA (Tsuchihara et al. 2009). We clustered the TSS tags into TSS clusters (TSCs) to separate individual putative promoter units (see the reference (Tsuchihara et al. 2009) for details). The numbers of TSS tags were normalized to parts per million mapped TSS tags, and the sum of the TSS tag concentrations associated with the same TSC was defined as the expression level of the TSC. The fold changes of each TSC in response to hypoxia were evaluated as the changes in

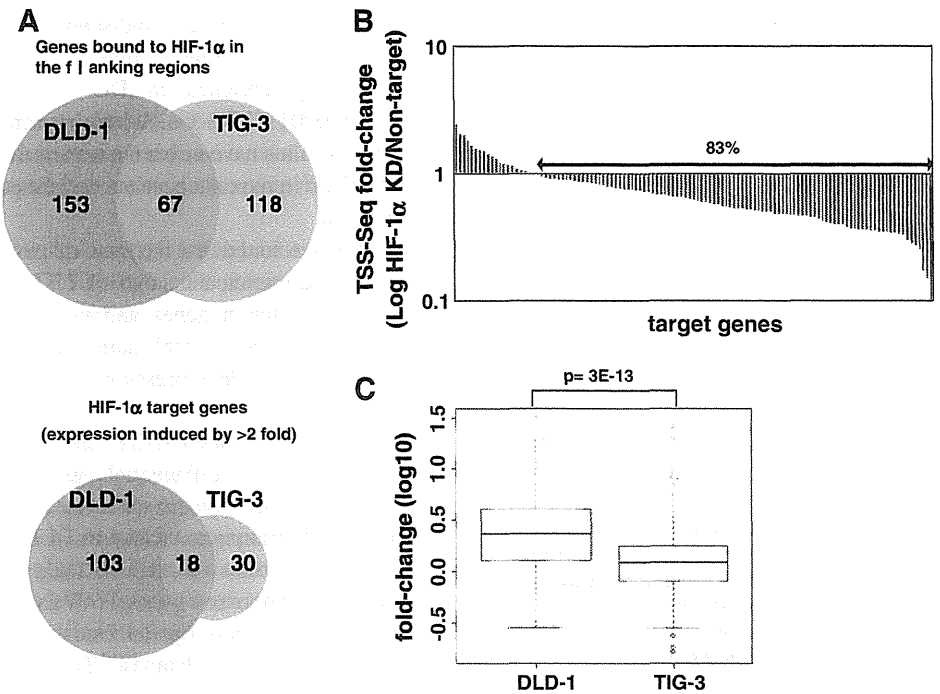

Fig. 3 Correlations of HIF-1α target genes and their expression levels. **a** Venn diagram of the number of identified HIF-1α target genes in DLD-1 and TIG-3 cells. The *upper panel* shows the total number of genes for which HIF-1α binding sites were identified in the flanking regions. The *lower panel* shows the identified target genes with expression levels that were induced more than twofold under hypoxia, as determined on the basis of digital TSS tag counts. **b** Fold changes in the expression levels of the identified target genes in HIF-1α knock-down cells. The fold changes were calculated as the TSS tag concentration in knock-down cells/TSS tag concentration in normal cells. The *bars* below the axis represent cases in which gene expression was repressed in HIF-1α knockdown cells in comparison to normal cells (83% of all cases). See Supplementary Fig. 4 for details of the validation analysis. **c** Distributions of the fold changes in expression levels of the genes bound by HIF-1α in DLD-1 and TIG-3 cells, as determined by the digital TSS tag counts. The statistical significance of the difference in distributions was evaluated using the Wilcoxon rank sum test ($P = 1e{-}13$)

TSS tag concentrations under hypoxia versus normoxia. [The statistics obtained for the TSS tags are shown in Supplementary Fig. 4. Additional details can be also found in the reference (Tsuchihara et al. 2009).]

We combined the ChIP-Seq and TSS-Seq data to examine the transcriptional consequences of HIF-1α binding. Among the above-mentioned binding sites, which were associated with the RefSeq genes, the TSS-Seq data revealed the presence of TSCs that overlapped the 5'-ends of the RefSeq transcript models in 220 and 185 genes in DLD-1 and TIG-3 cells, respectively (note that a particular binding site sometimes corresponds to multiple TSCs). Particularly, in 121 and 48 genes, the expression levels of the TSCs were induced more than twofold under hypoxia, respectively, (Fig. 3a; referred to as "HIF-1α target genes/TSCs" hereafter). As shown in Supplementary Fig. 4c, the expression levels of representative target genes were validated by real-time PCR. We also performed similar TSS-Seq analyses in HIF-1α knock-down cells. Of the 121 putative HIF-1α target genes in DLD-1 cells, the hypoxia-responsive induction of gene expression was repressed in 100 cases in the TSS-Seq analysis in which HIF-1α was knocked down (83%; Fig. 3b). Supplementary File contains the top 25 HIF-1α target genes, which are ranked according to their fold change at the transcription level in DLD-1 and TIG-3 cells (see also Supplementary Fig. 5 for an example of HIF-1α target genes identified by ChIP-Seq and TSS-Seq analyses).

We further analyzed whether the identified HIF-1α targets were translated into proteins. To achieve this goal, we analyzed the RNAs that were incorporated into the translating ribosome (polysome) fractions. RNA was isolated from DLD-1 cells that were cultured under normal and hypoxic conditions. Polysome-incorporated RNA was isolated from each cell culture by sucrose density gradient purification of the cytoplasmic RNA (see "Materials and methods"). A total of 46,170,807 36-bp single-end-read RNA-Seq tags were generated, and the abundance of the transcripts was evaluated by digital counts of the RNA-Seq tags (Supplementary Fig. 6). We expected that the changes in the RNA-Seq tags could be used as an indicator to measure gene expression changes at the translational level. As shown in Fig. 4, the levels of RNA-Seq tags for the putative HIF-1α target genes were significantly induced in

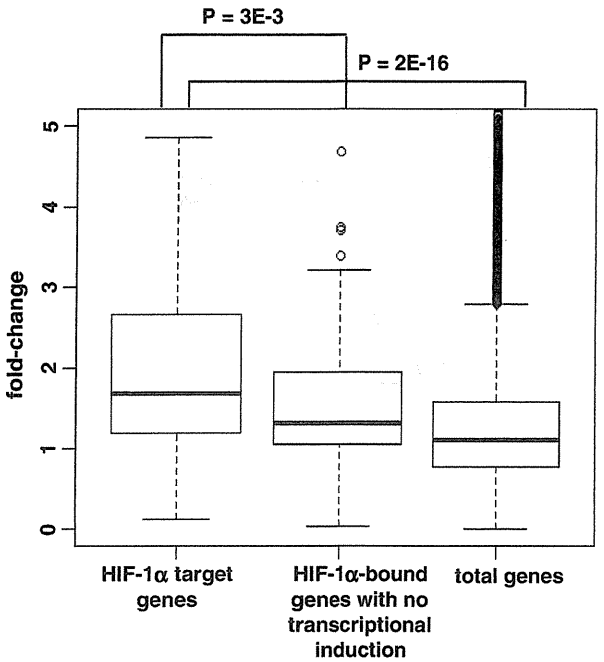

Fig. 4 Fold change in the amount of RNA present in polysome fractions in response to hypoxic stimuli. Distributions of the fold changes of transcripts in the polysome fractions in response to hypoxia, as determined by digital RNA-Seq tag counts. Fold changes in HIF-1α-bound target genes for which transcriptional induction was observed (*left*), HIF-1α-bound genes for which no transcriptional induction was observed (*middle*) and all RefSeq genes (*right*) are shown. Statistical significance in the differences in the distributions were evaluated using the Wilcoxon rank sum test and are shown in the margin

the polysome fractions in response to hypoxia, which suggested that a majority of the gene expression was induced at the protein level, too. Taking these results together, we concluded that we had identified direct targets of HIF-1α that were bound by HIF-1α and that the expression of these targets was induced at transcriptional and translational levels.

Target genes of HIF-1α in DLD-1 and TIG-3 cells

We compared the HIF-1α target genes in DLD-1 cells and TIG-3 cells. We found that 18 genes overlapped, and the remaining genes were unique to each cell type. We performed Gene Ontology (GO) term (Ashburner et al. 2000) analysis to examine whether the gene groups of particular functional categories were enriched in the respective cell types. Seven and four GO terms were enriched (hypergeometric distribution; $P < 0.01$) in DLD-1 and TIG-3 cells, respectively (see also Supplementary Tables 2a and 2b for the full list). Of these, GO term "glycolysis" (GO: 0006096) was commonly enriched. "Response to hypoxia" (GO: 0001666) was enriched only in DLD-1 cells. It was intriguing that the genes belonging to this GO term, which are considered to have indispensable roles in the maintenance of homeostasis under hypoxic conditions, were selectively enriched in DLD-1 cells (proportion test; $P = 0.0001$). TIG-3 cells, which represent a normal cell type and thus have never been exposed to hypoxia, may not be robust in hypoxia because several essential genes are not induced.

We also found that the hypoxic response of DLD-1 cells was more significant than that of TIG-3 cells, both in the number of target genes and in the degree of expression induced (Wilcoxon rank sum test; $P = 3e{-}13$). Figure 3c illustrates the fold expression changes in DLD-1 and TIG-3 cells.

To further characterize the molecular mechanisms underlying the differential gene expression observed between the cell types, we assessed the global expression patterns of transcription factors in DLD-1 cells and TIG-3 cells on the basis of digital TSS tag counts for the genes encoding transcription factors. We used a total of 140 well-characterized transcription factors that were registered in the representative database TRANSFAC. As shown in Supplementary Fig. 7, transcriptional induction of the transcription factor genes was more significant in DLD-1 cells than in TIG-3 cells ($P = 0.02$) (see also Supplementary Fig. 8 for an example). To specify the transcription factors that might be responsible for the activation of particular genes, we examined the enrichment of consensus binding sites for specific transcription factors in the surrounding regions of HIF-1α binding sites. We examined the genes which were identified as HIF-1α target genes in DLD-1 cells, TIG-3 cells or in both cell types. We found that fourteen consensus binding sequences were enriched around the HIF-1α binding sites in total. Of these sequences, four transcription factors were transcriptionally induced more than twofold in the TSS-Seq analysis in the corresponding cell types (see Supplementary Table 3 for the full list). These transcription factors may regulate the transcription of HIF-1α target genes in cooperation with HIF-1α in DLD-1 cells or TIG-3 cells. It is also interesting to speculate that the coordinated expression of the transcription factors might have been acquired throughout cellular history. DLD-1 cells originated from a solid tumor that experienced severe hypoxic conditions (Rajaganeshan et al. 2009; Schmitz et al. 2009). Similar analyses will be necessary in the future to assess a greater variety of cell types in order to validate this hypothesis.

Epigenetic regulation in regions surrounding HIF-1α binding sites

We also analyzed the nucleosome structure in the surrounding regions of the HIF-1α binding sites in DLD-1 and TIG-3 cells. It has been reported that nucleosomes acquire

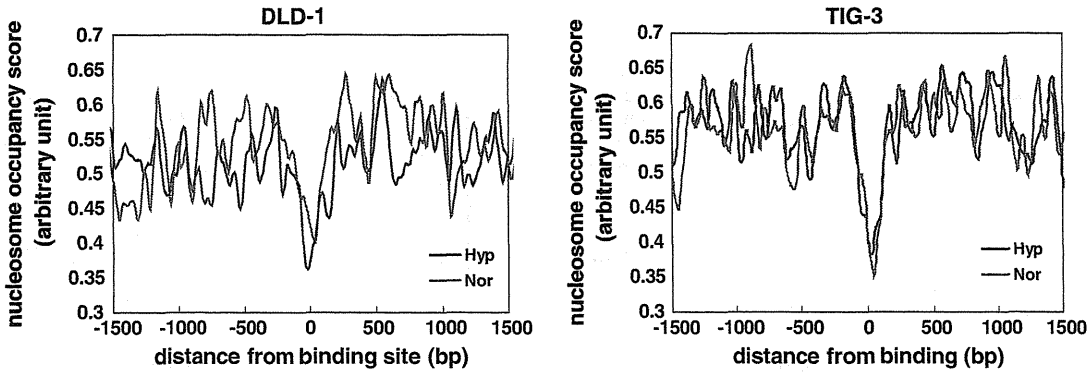

Fig. 5 Nucleosome structures in regions surrounding HIF-1α binding sites. Nucleosome occupancy in regions surrounding the identified HIF-1α binding sites. The calculated nucleosome occupancy scores (y-axis) are plotted against every genomic coordinate (x-axis). The computational procedure used to calculate the nucleosome occupancy score is described in "Materials and methods". The putative binding site of HIF-1α, represented by HRE, was defined as position 0 (x-axis). *Blue* and *red* lines indicate the nucleosome occupancy scores under hypoxia and normoxia, respectively. Only the HIF-1α binding sites for which the associated TSCs were induced by more than twofold were used in the plot for each cell type. For details regarding the Nucleosome-Seq datasets, see Supplementary Fig. 9. (Color figure online)

an open chromatin structure following the binding of transcription factors (Zhang et al. 2009). To analyze the nucleosome structure, Nucleosome-Seq (Albert et al. 2007; Jiang and Pugh 2009; Schones et al. 2008) was used in DLD-1 and TIG-3 cells. We generated a total of 130,088,634 36-bp sequence tags from micrococcal nuclease-digested genomic DNA and calculated the occupancy of the nucleosome according to a previously reported procedure (Albert et al. 2007) (Supplementary Fig. 9; see "Materials and methods"). As shown in Fig. 5, the occupancy score for DLD-1 cells dropped sharply in the regions surrounding hypoxia-responsive HIF-1α binding sites under hypoxia (Fig. 5, blue line in the left panel), which suggested that the nucleosome structures formed an open structure. On the other hand, the occupancy scores for HRE motifs that were not bound by HIF-1α did not indicate such an open structure (Supplementary Fig. 10). The same results were obtained for TIG-3 cells (Fig. 5, blue line in the right panel). These results indicated that only a part of HRE motifs had an open nucleosome structure suitable for subsequent transcriptional activation, although HRE motifs were present throughout the human genomic sequences. Previous HIF-1α ChIP-chip studies have reported that less than 1% of the HRE motifs within the genomic sequences that are located in regions used for the microarrays are bound by HIF-1α (Mole et al. 2009; Wenger et al. 1998, 2005). Epigenetic regulation of the human genome plays an essential role in specifying the binding sites of HIF-1α in addition to the consensus binding sequences.

Interestingly, for the HIF-1α target HRE, we observed that the open chromatin structure was already formed under normoxia (Fig. 5, red line in the left panel for DLD-1 cells and in the right panel for TIG-3 cells). On the other hand, the TSS tags induced by hypoxia for the approximately 1-kb downstream regions were first detected after the binding of HIF-1α under hypoxia. To further examine the chromatin status surrounding HIF-1α binding regions, we analyzed the histone modification patterns (H3K4me3, H3Ac and H3K27me3) and binding status of pol II by ChIP-Seq analyses. The numbers of sequence tags generated for each analysis are shown in Supplementary Fig. 11 and 12. Figure 6 shows the observed histone modification and pol II binding profiles in the regions surrounding the HIF-1α binding sites. Consistent with the TSS-Seq profiles, pol II binding first increased under hypoxia. On the other hand, the histone markers for active chromatin, H3K4me3 and H3Ac, were already observed under normoxia, whereas the marker for silent chromatin, H3K27me3, was absent throughout the hypoxic shock period (Fig. 6b–d, f–h). These results suggested that open and active chromatin structures originally formed under normoxia, irrespective of the pol II binding status and downstream transcriptional induction. The cells seemed to stay in a somewhat prepared status for a hypoxic response. The cels might memorize their histories with respect to the chromatin status. Such a prepared status of the cells via epigenetic regulation has been discussed in previous reports (Heintzman et al. 2007; Kim et al. 2008; Mole et al. 2009); however, this is the first study to provide direct evidence supporting this possibility.

HIF-1α regulation of alternative promoter variants

One of the greatest advantages of using ChIP-Seq and TSS-Seq is that the HIF-1α binding sites and subsequent transcriptional induction can be detected separately for alternative promoters. Of the identified 121 and 48 HIF-1α

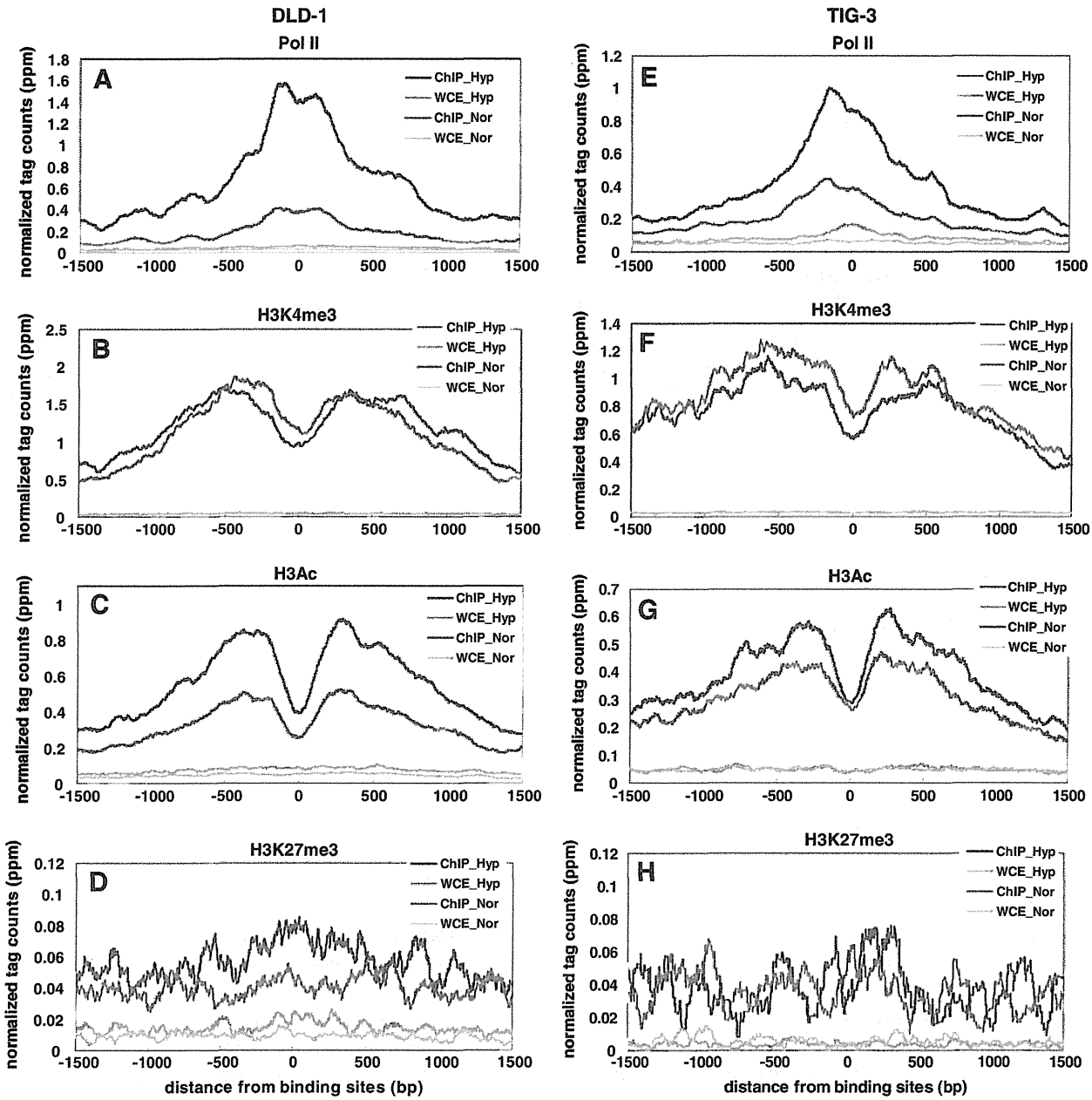

Fig. 6 Profiles of RNA polymerase II binding and histone modification in the regions surrounding HIF-1α binding sites. Average tag concentrations (y-axis) obtained in the ChIP-Seq analyses of pol II (**a**, **e**), H3K4me3 (**b**, **f**), H3Ac (**c**, **g**) and H3K27me3 (**d**, **h**) are plotted for each genomic coordinate (x-axis). The observed profiles are shown for DLD-1 (**a**–**d**) and TIG-3 cells (**e**–**h**). In each panel, *red* and *blue lines* represent the tag counts under hypoxia and normoxia for the IP samples, respectively. *Green* and *sky blue lines* represent the results obtained for the background control (whole cell extract samples). The putative binding site of HIF-1α, represented by HRE, was defined as position 0 (x-axis). For each panel, the tag counts for 290 and 221 binding sites, which have HRE among the HIF-1α target TSCs (either geneic or intergenic TSCs), were averaged in DLD-1 and TIG-3 cells, respectively. (Color figure online)

target genes, 15 and 5 target genes had more than one alternative promoter variant in the RefSeq transcript models in DLD-1 and TIG-3 cells, respectively (see Supplementary Fig. 13 for one of these examples). It should also be noted that the majority of HIF-1α target alternative promoters are novel promoters. One hundred and ninety-three and 123 TSCs were identified as previously uncharacterized putative alternative promoters in DLD-1 and

Table 2 Number of TSCs associated with binding sites located in genic and intergenic regions

	#Binding sites	#HIF-1α bound genes in which TSCs overlapped the 5′-end of RefSeq transcript models	#HIF-1α target genes
a			
DLD-1	193	220	121
TIG-3	161	185	48

	#Previously characterized AP TSCs (#genes)	#Newly identified AP TSCs (#genes)
b		
DLD-1	15 (15)	193 (89)
TIG-3	5 (4)	123 (80)

	#Binding sites	#Intergenic TSCs	#Intergenic TSCs located in repetitive regions
c			
DLD-1	207	262	21
TIG-3	289	105	21

a The number of genes that were bound to HIF-1α in the flanking regions (from 10 kb upstream to 1 kb downstream of the binding sites) of the 5′-ends in the RefSeq transcript models. "HIF-1α target genes" indicate genes with expression levels that were induced more than twofold by hypoxia

b The number of HIF-1α target TSCs corresponding to alternative promoters (APs) which had been characterized as RefSeq transcript variants (first column) and those which were newly identified in this study (second column). The numbers of corresponding genes are shown in parentheses

c The number of HIF-1α target TSCs located in intergenic regions

TIG-3 cells, respectively (Table 2b). The overlap of newly identified HIF-1α target alternative promoters with previously characterized transcript models that are present in representative databases is shown in Supplementary Table 4. It is possible there is an unexpectedly large population of the hitherto undiscovered alternative promoters involved in hypoxic response of cells.

HIF-1α binding sites in intergenic regions

The combination of ChIP-Seq and TSS-Seq allowed us to identify HIF-1α targets even in previously unannotated (so-called "intergenic") regions. Approximately one-third of the HIF-1α binding sites were located in regions that were >50 kb from the nearest RefSeq genes. For these intergenic binding sites, we searched for TSCs located from 1 kb upstream to 10 kb downstream of the HIF-1α binding sites and identified 262 and 105 putative HIF-1α target TSCs in DLD-1 and TIG-3 cells, respectively (Table 2). Twenty-one and 21 binding sites were located in repetitive regions (see also Supplementary Table 5), which suggested that some HIF-1α target sites might also have arose from repetitive regions, as previously reported for several transcription factors (Tomilin 2008; Polak and Domany 2006) (see Supplementary Fig. 14 for one of these examples). We also found that the distributions of the distances from HIF-1α binding sites to TSCs were similar to those determined for the HIF-1α target RefSeq genes (Supplementary Fig. 15).

We searched the completely sequenced full-length cDNAs that corresponded to these TSCs and found seven overlapping cDNAs from the FLJ (Ota et al. 2004) and MGC (Gerhard et al. 2004) cDNA collections in DLD-1 and TIG-3 cells (Supplementary Table 6a). All of the corresponding full-length cDNAs lacked clear open reading frames. We determined the complete sequences of the overlapping ESTs for an additional three cases and obtained similar results (Supplementary Table 6b). An analysis of the RNA-Seq tags in the polysome fractions also suggested that majority of these intergenic HIF-1α targets were not translated (Supplementary Fig. 6c). Therefore, the intergenic TSCs of HIF-1α targets should be used to transcribe hypoxia-responsive non-protein-coding transcripts (Wilusz et al. 2009).

To determine whether these TSCs were precursors of micro-RNAs (He and Hannon 2004), we searched for miRNAs registered in the miRBase (Griffiths-Jones et al. 2006) from 10 kb upstream to 50 kb downstream of the intergenic HIF-1α binding sites. No miRNAs were located within these regions. Furthermore, we could not detect any

overlap against the previously reported repertoire of putatively-functional long non-coding RNAs (Guttman et al. 2009). We considered that these intergenic TSCs were not derived from our experimental errors, especially those identified in DLD-1 cells (Barski et al. 2007; Welboren et al. 2009), since we detected clear binding signals for pol II around 61 (21%) and 2 (2%) intergenic TSC targets of HIF-1α. These results collectively suggested that we had identified putative non-protein-coding transcript targets of HIF-1α that were comparable to the canonical RefSeq targets with respect to the number of targets and their levels of induction. However, further analyses are needed to determine the biological relevance underlying the induction of these non-coding transcripts.

Conclusion

In the present paper, we describe the genome-wide identification of HIF-1α binding sites in the cancerous DLD-1 cell line and the normal lung fibroblast TIG-3 cell line using ChIP-Seq. We utilized several types of Illumina GA analyses to further characterize the identified binding sites. The greatest advantages of using massively parallel sequencers may be the ability to obtain versatile biological data. In fact, we were able to obtain extensive biological information that allowed us to interpret the binding events; such information could not have been obtained using ChIP-chip or microarray analysis alone. We unexpectedly discovered that HIF-1α binding sites regulate an unexpectedly large population of alternative promoters and putative non-coding transcripts. These categories of the transcripts may also contribute to hypoxic response of the cell, though they have been overlooked in previous studies. The analyses targeting these transcripts were first enabled by the single-base resolution of massively parallel sequencing technologies, for which no previous knowledge was necessary for the probe design.

An obvious limitation of the present study was the use of only two cell types. Thus, we could not identify all of the possible binding sites of HIF-1α. The identification of different targets in DLD-1 and TIG-3 cells indicates that similar analyses should be conducted using a wider variety of cell types. In addition, our stringent criteria may have precluded the identification of many genuine targets. Indeed, some genes that have been well characterized as HIF-1α target genes were not found in our study, including the telomerase (TERT) gene (Nishi et al. 2004), the vascular endothelial growth factor A (VEGFA) gene (Carmeliet et al. 1998; Forsythe et al. 1996) and the erythropoietin gene (Semenza and Wang 1992) [also note that this result is similar to that obtained in a previous ChIP-chip analysis (Mole et al. 2009)].

All of the data generated in the present study were essentially the sequence data; however, we believe that the applications of massive sequencing at various biological levels have yielded substantial biological indications. We could find that the epigenetic regulation of chromatin status and the differential repertoires of expressed transcription factors played essential roles in the cell-type-specific induction of target genes in DLD-1 and TIG-3 cells in response to hypoxia. Further extensive analyses of epigenetic changes in cancer and normal cells using Nucleosome-Seq and ChIP-Seq and of the expression patterns of transcriptions factors using TSS-Seq or RNA-Seq could reveal the mechanisms by which the differential responses of various cells to hypoxia are encoded. Such information could reveal how hypoxic cellular responses are mediated by HIF-1α in a wide variety of cancer cells.

References

Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF (2007) Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. Nature 446:572–576

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 25(1):25

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129(4):823–837

Benita Y, Kikuchi H, Smith AD, Zhang MQ, Chung DC, Xavier RJ (2009) An integrative genomics approach identifies hypoxia inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res 37(14):4587–4602

Carmeliet P, Dor Y, Herbert JM, Fukumura D, Brusselmans K, Dewerchin M, Neeman M, Bono F, Abramovitch R, Maxwell P, Koch CJ, Ratcliffe P, Moons L, Jain RK, Collen D, Keshert E (1998) Role of HIF-1alpha in hypoxia-mediated apoptosis, cell proliferation and tumour angiogenesis. Nature 394(6692):485–490

Davuluri RV, Suzuki Y, Sugano S, Plass C, Huang TH (2008) The functional consequences of alternative promoter use in mammalian genomes. Trends Genet 24(4):167–177

Forsythe JA, Jiang BH, Iyer NV, Agani F, Leung SW, Koos RD, Semenza GL (1996) Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. Mol Cell Biol 16(9):4604–4613

Fraisl P, Aragones J, Carmeliet P (2009) Inhibition of oxygen sensors as a therapeutic strategy for ischaemic and inflammatory disease. Nat Rev Drug Discov 8(2):139–152

Fyles AW, Milosevic M, Wong R, Kavanagh MC, Pintilie M, Sun A, Chapman W, Levin W, Manchul L, Keane TJ, Hill RP (1998) Oxygenation predicts radiation response and survival in patients with cervix cancer. Radiother Oncol 48(2):149–156

Gatenby RA, Gillies RJ (2004) Why do cancers have high aerobic glycolysis? Nat Rev Cancer 4(11):891–899

Gerhard DS, Wagner L, Feingold EA, Shenmen CM, Grouse LH, Schuler G, Klein SL, Old S, Rasooly R, Good P, Guyer M, Peck