Studying relationship between increased 5-methylcytosine oxidation and expression of cancer stem cell markers in glioblastoma cell lines

Submitted by

Vithusha Wijeyagumar BSc

University of Nottingham

School of Medicine

Division of cancer and stem cells

MRes Thesis

September 2017

Contents

1. Acknowledgements 3
2. Abstract 4
3. Introduction 5
   - 3.1 Epigenetics 5
   - 3.2 DNA Methylation 5
   - 3.3 DNA Demethylation 7
     - 3.3.1 TET Family enzymes 9
     - 3.3.2 Oxidised forms of 5mC 14
   - 3.4 Active Demethylation of DNA in Cancer 17
   - 3.5 GBM 18
     - 3.5.1 Epidemiology 18
     - 3.5.2 Pathophysiology 19
       - 3.5.2.1 Molecular Subtypes of GBM 20
   - 3.6 DNA Methylation Patterns in GBM 21
     - 3.6.1 *IDH* Mutations & glioma-CIMP G-CIMP 22
   - 3.7 Demethylation in GBM 24
   - 3.8 Cancer Stem cells (CSCs) 25
   - 3.9 Glioblastoma Stem cells (GSCs) 25
     - 3.9.1 Glioblastoma Stem Cell markers 26
       - 3.9.1.1 CD133 27
       - 3.9.1.2 Sox2 (Sex-determining region Y-Box) 27
       - 3.9.1.3 Nestin 28
       - 3.9.1.4 Olig2 30
       - 3.9.1.5 Notch1 31
       - 3.9.1.6 EGFR 33
4. Materials 36
   - 4.1 Primer 37
5. Methods 38
   - 5.1 Cell Culture 38
   - 5.2 Primer Design 38

5.3 RNA Extraction 38
5.4 cDNA Synthesis 39
5.5 qPCR 40
5.6 Immunostaining 40
5.7 Confocal Microscopy 40
5.7.1 Generation of Confocal Images 41
5.7.2 Intensity Profiling 41
5.7.3 2.5D Intensity Plots 42
5.8 Co-localization Analysis 42
5.8.1 ZEN Black 42
5.8.2 Fiji ImageJ 42
5.8.3 Pearson's Correlation Coefficient – EndMemo 43
5.9 Statistical Analysis 43
6. Results 44
6.1 5caC staining is heterogenous in glioblastoma cell lines but not in induced pluripotent stem cells (iPSCs) 44
6.2 Studying the expression levels of stem cell markers in GBM and medulloblastoma cell lines by qPCR 45
6.3 Validation of the antibodies for these markers 46
6.3.1 The spatial distribution of the markers' immunostaining in the cells lines 49
6.3.2 Studying potential correlation between the signal intensities of 5caC and cancer stem cell markers' patterns in GBM and medulloblastoma cell lines 53
7. Discussion 60
7.1 The possible correlation between increased levels of 5caC and increased expression of SOX2 and Olig2 63
8. REFERENCES 66
9. Supplementary information 91
10. Abbreviations 120

1. ACKNOWLEDGEMENTS

I would like to thank Ashley Ramsawhook for designing and guiding me to understand the topic and complete my MRes project; Dr Lara C Lewis for training me in all experimental techniques conducted in this project and providing exceptional support in understanding procedures and answering any query I went to her with; Maria Eleftheriou and Abdulkadir Abakir for providing me all the relevant cell lines needed to carry out my experiments; Dr Alexey Ruzov for supervising me throughout my whole journey and being a supporting backbone to every decision I took and guided me in the right direction.

Finally I would like to thank my colleagues Cho Kiu Lo (Peggy), Jordan Thorpe, Jamie Bhagwan, Lazaros Fotopoulos, Dr Gary Duncan and Zubair Ahmed for their support and making my MRes year an enjoyable experience.

2. ABSTRACT

DNA methylation (5-methylcytosine, 5mC) is an epigenetic modification generated by the addition of a methyl group to the 5 position on the cytosine ring of the DNA and it is important for a range of biological processes such as development and chromosome stability. 5mC can be enzymatically oxidised to 5-hydroxymethylcytosine, 5-fCformylcytosine and 5-carboxylcytosine (5caC) by TET proteins. These oxidised forms of 5mC may both serve as intermediates in the process of active DNA demethylation and/or contribute to the regulation of transcription. According to our results, 5caC exhibits heterogeneous distribution between different cells in glioblastomas (GBM), aggressive gliomas classified as grade IV astrocytoma which is always rapidly growing and highly malignant. Interestingly, GBMs are highly heterogeneous cancers containing a cancer stem cell (CSCs) subpopulation which is responsible for the reoccurrence of the tumour after treatment.

In the present study we examine the spatial distribution of 5caC in brain tumour cell lines and show that the levels of this modification positively correlate with the staining intensity of such CSCs markers as SOX2 and Olig2 in GBM cells. Our results imply that GBM CSCs may be enriched in 5caC and therefore, this oxidised form of 5mC may be employed for identification of stem cell-like populations in GBM with potential future implications for diagnostics and therapy of the brain tumours.

3. INTRODUCTION

3.1 Epigenetics

The earliest definition used for Epigenetics was by Aristotle who termed it as the sequence of steps for development which then advanced into the idea of genes having the capability to turn on and off by Spemenn and Mangold (1924)(Lester, Conradt and Marsit, 2016). Nowadays epigenetics is defined by “an addition to changes in the genetic sequence” (Weinhold, 2006). It studies the “mitotically heritable and reversible molecular information outside of the DNA sequence” (Ladd-acosta and Fallin, 2016). Epigenetics links genotype and the phenotype resulting in a vital component in genome stability, chromatin modification and DNA methylation. Waddington introduced the term epigenetics (Waddington, 1942) to define his term of the “epigenetic landscape” (Tronick and Hunter, 2016). He explained that epigenetics is part of a large system which defines the molecular mechanism involved in the process of a disease, as it is not solely due to gene and environmental exposures alone because they are also components of this large molecular system.

3.2 DNA Methylation

The most commonly studied epigenetic modification is DNA methylation. DNA methylation, originally discovered in the 1970s, is the addition of a methyl group to the 5 position on the cytosine ring of the DNA producing 5-methylcytosine (5mC) (Jones, 2012). DNA methylation is presumably a signalling tool in cells which is used to turn the genes “off” and it is a vital component in several cellular processes such as development, genomic imprinting, X-chromosome inactivation and chromosome stability (Phillips, 2008). Methylation is

crucial for cell differentiation and embryonic development by repressing gene expression. Silencing gene expression is essential for maintaining genome stability during DNA replication, otherwise this can cause aberrant gene activation which can result in transcriptional deregulation of certain genes leading to the development of diseases and abnormalities (Robertson, 2005).

DNA methylation is driven and regulated by DNA methyltransferase (DNMT) enzymes. There are three major types of DNMTs in mammals: DNMT1, DNMT3A and DNMT3B (Robertson *et al.*, 1999). DNMT1, known as the maintenance methyltransferase enzyme, is responsible to reproduce the methylation pattern after each cycle of DNA replication (Robertson *et al.*, 1999). The methyltransferase reads the DNA sequence with the hemimethylated CpG sites and methylates the new daughter strand (Jones and Liang, 2009). This maintains the DNA structure and controls gene expressions.

DNMT3A and DNMT3B are known to be *de novo* methyltransferase enzymes and show equal affinity for hemi- and unmethylated DNA (Okano *et al.*, 1999). They are present during early embryonic development and methylate unmethylated CpG dinucleotides in the DNA (Okano *et al.*, 2017). *De novo* methylation during early embryonic development is crucial for mammalian development and tissue differentiation by organising and compartmentalising the genome leading to tissue specific gene expression (Santos and Dean, 2004). Mutations in DNMT3 can cause chromosome instability leading to abnormalities such as acute myeloid leukaemia (ALL) (Ley *et al.*, 2010). As embryonic stem cells (ESCs) differentiate into adult cells, DNMT3 levels downregulate whereas maintenance methlytransferase, DNMT1, is upregulated (Okano *et al.*, 2017).

During early embryonic development, DNMT3 is responsible for the methylation of regions which are CpG rich (Hackett and Surani, 2012).

CpG sites (non-CpG islands) are regions of DNA which have a linear sequence of cytosine nucleotide followed by guanine nucleotide along the 5' to 3' direction and elevated levels of CpG sites in a specific region are known as CpG islands (Illingworth and Bird, 2009). CpG islands are areas of the genome containing over 200 CpG nucleotides linked consecutively and are commonly identified in the promoters of housekeeping genes, regulating their expression (Vinson and Chatterjee, 2012). CpG islands located within promoter regions are usually unmethylated allowing them to be transcriptionally active (Skalnik, 2001), facilitating different development processes and allowing tissue specific expression (Meng *et al.*, 2015). However, CpG islands associated with promoters of the genes involved in X-chromosome inactivation become methylated to control dosage compensation of inactivated chromosomes (Meng *et al.*, 2015). Moreover, specific non CpG islands (CpG sites) are commonly methylated to silence expression of certain genes such as oncogenes (Han *et al.*, 2011). However, there is a number of CG sites which remain unmethylated which vary between different tissues creating methylation patterns that are tissue-specific (Cheishvili, Boureau and Szyf, 2015).

3.3 DNA Demethylation

Equally important as DNA methylation is DNA demethylation which is the process of the removal of the methyl groups from DNA. It is involved in epigenetic reprogramming of genes and also in mechanisms underlying different diseases such as tumour progression (Szyf, Pakneshan and Rabbani, 2004). Demethylation is important in

primordial germ cells to allow new imprinting to be established according to the sex of the germ line (Feng, Jacobsen and Reik, 2010) and also important during early embryonic stages to create pluripotent states (Wu and Zhang, 2014). It is characterised by two types, passive demethylation and active demethylation (Moore, Le and Fan, 2013).

During cell replication, the DNMT1 enzyme maintains the methylation of hemi-methylated DNA. However, the loss of DNMT1 results in diluting out 5mC from DNA after several successful rounds of cell replication, consequently causing passive demethylation (Niehrs, 2009). Passive demethylation can occur during early embryonic development as DNMT1 is low during these stages (Robertson *et al.*, 1999). Any kind of loss or gradual decrease in the level of DNMTs may cause passive demethylation (Niehrs, 2009; Bhutani, Burns and Blau, 2011b).

On the other hand, active demethylation occurs in non-replicating cells or cells where replication is inhibited (Bhutani, Burns and Blau, 2011a). It occurs in animals and plants, during differentiation and early development phases (Niehrs, 2009). One of the most well characterised pathways of active DNA demethylation is via ten-eleven translocator (TET) family of enzymes and TDG/base excision repair (BER) pathway (Bhutani, Burns and Blau, 2011a). The TET enzymes drive the process of oxidation of 5mC to 5-hydroxylmethylcytosine (5hmC) and also the further oxidation of 5hmC to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Iyer *et al.*, 2009). 5fC and 5caC are excised from DNA by thymine DNA glycosylase (TDG) via the base excision repair (BER) pathway (He *et al.*, 2011).

TDG is an enzyme responsible for DNA mismatch repair by binding and excising mismatches in pyrimidines in GU and GT pairs (Lindahl and Wood, 1999). TDG has been shown to excise 5fC and 5caC but not 5hmC from DNA (Maiti and Drohat, 2011). TDG removes the methyl group which would be identified as the “incorrect” base from the sugar-phosphate backbone of the DNA leaving an abasic site and this is replaced with an unmodified cytosine by other enzymes during base excision repair (BER) pathway causing active DNA demethylation (Rasmussen and Helin, 2016), (Gehring, Reik and Henikoff, 2009).

DNA demethylation is important during embryonic development to specialise embryonic stem cells into tissue-specific cells and to activate silenced genes (Teperek-Tkacz *et al.*, 2011).

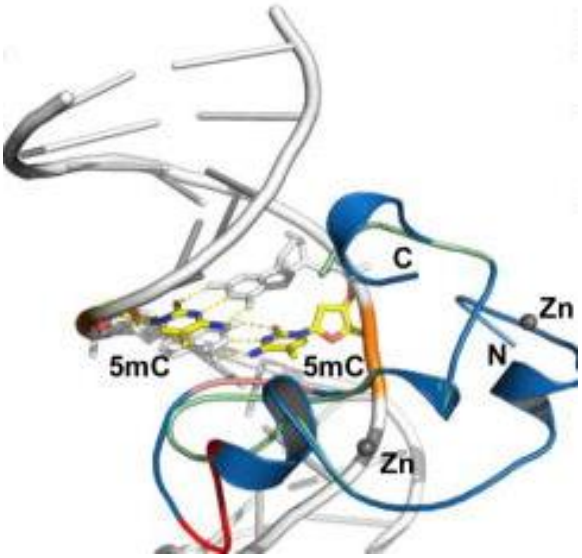


Figure 1: DNA containing 5mC- and 5hmC-modified CpG with TET1 CXXC domain bound away from the major groove (Xu *et al.*, 2011).

3.3.1 TET FAMILY ENZYMES

There are three main ten-eleven translocator (TET) enzymes, which are TET1, TET2 and TET3. All the TET enzymes contain a C-terminal

catalytic domain consisting of the double stranded b-helix (DSBH) fold characteristic of the 2-oxogluterate (2OG)- and Fe(II)-dependant dioxygenase (2OGFeDO) and cytosine rich region. They also belong to the 2OGFeDO superfamily (Tahiliani *et al.*, 2009). TET1 and TET3 also have a CXXC zinc finger domain (Iyer *et al.*, 2009). The TET1 with the N-terminal CXXC domain binds to both methylated and unmethylated CpG sites. Xu et al. 2011 performed a study to create a model of the structure of the TET1 CXXC and noticed that the domain binds to the DNA via a major groove in the CpG sites; the shortened loop in the domain moves a little from the major groove and this allows the binding of 5mC or 5hmC (Fig 1) which predicts the ability of the TET1 CXXC domain to bind to methylated and/or unmethylated CpG containing DNA (Xu *et al.*, 2011). The expression of these enzymes vary in different cell types. Ito et al. 2010 conducted a study to examine the expression of these enzymes in various mouse cell types via RT-qPCR and showed that TET1 and TET2 are highly expressed in embryonic stem cells whereas TET3 is most highly expressed during germ line specification.

Primarily distinguished as a fusion partner of histone methyltransferase MLL (mixed lineage leukaemia) in acute myeloid leukaemia (Ono *et al.*, 2002), TET1 has later been found to portray enzymatic capabilities for hydroxylating 5mC to generate 5hmC (Tahiliani *et al.*, 2009). Studies on human TET1 have indicated that it can to convert 5mC to 5hmC in an Fe(II)- and α-KG-dependant manner *in-vitro* (Tahiliani *et al.*, 2009). A study performed by Tahiliani et al. 2009 involved RNA interference-mediated depletion of TET1, which decreased the levels of 5hmC within the mouse embryonic stem cells

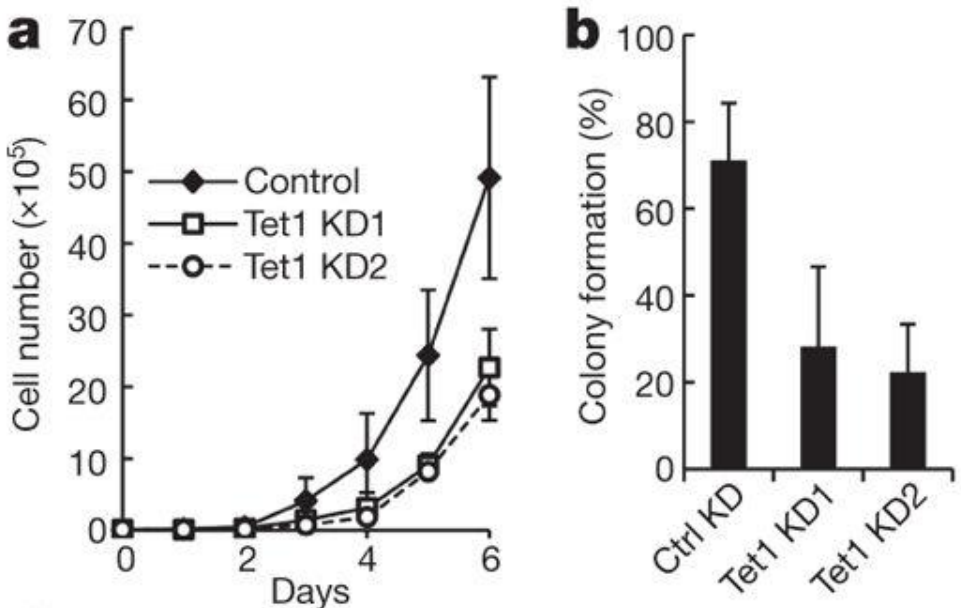


Figure 2: **a** TET1 knockdown showing reduction in ES cells proliferation. Growth curves represent control and TET1 knockdown in KD1 and KD2 cells by counting number of cells every day. Average number of cells and s.d. are shown. **b** study conducted included a plate with single knockdown or control cell and its ability to form colonies within a timescale of 6 days. Results indicated no difference between colony sizes however a decrease in number of colonies in TET1 knockdown in comparison to that of control cells (Ito *et al.*, 2010).

(mESCs) verifies the potential role of TET1 in modifying 5mC to 5hmC. To speculate further into the function of TET1 in the developmental aspect within embryonic stem cells, further studies were conducted, concluding in the likelihood of its role in ES cell self-renewal and maintenance (Ito *et al.*, 2010). This supposition was proposed since

when TET1 is knocked down, there is a decrease in the growth rate of ES cells (Fig 2), suggesting that it may be significantly down to affect self-renewal abilities.

Moreover, TET1 knockdown has also been shown to have an impact on the expression of Nanog and a minor impact on the expression of SOX2 and OCT4 (Ito *et al.*, 2010). This is also a contributing factor to the role of TET1 in ES maintenance (Ito *et al.*, 2010) as TET1 was shown to uphold the hypomethylated state of the promoters allowing the genes to be transcriptionally active for ES cell maintenance. However, Wu et al. 2011 identified the interplay between TET1 and Polycomb group proteins in silencing of developmental regulators during ES cell maintenance. The study proposed that TET1 is involved in the *PRC2* (polycomb repressive complex 2) recruitment as the depletion of TET1 causes a decrease in the *Ezh2* (enhancer of zeste homolog), a histone methyltransferase (Simon and Lange, 2008), at PRC2 targets (Wu *et al.*, 2011). This corroborates a probable dual functionality for TET1 in the regulation of transcription in ES cells (Wu *et al.*, 2011).

TET gene expression declines during differentiation from early hematopoietic progenitors and another TET homolog, TET2 is coincident with the expansion of monocyte-macrophage lineage cells *in-vitro* (Holmfeldt and Mullighan, 2011), highlighting the potential importance of TET2 in epigenetic regulation of haematopoiesis. As identified in a previous study, TET2 mutations are also concomitant with low levels of 5hmC and global hypomethylation (S. Ito *et al.*, 2010). TET2 loss of function mutations throughout the gene, including missense, nonsense and frameshift mutations, have been reported for hematologic malignancies such as myelodysplastic syndrome and

myeloproliferative neoplasms (Holmfeldt and Mullighan, 2011). There has been suggestive evidence to claim that TET2 mutations may play a vital role in leukemogenesis. In a previous study, it was shown that in AML there are exclusive mutations of TET2 and isocitrate dehydrogenase genes *IDH1/IDH2* and 2-hydroxygluterate. The by-product of the neomorphic IDH1/2 mutant proteins, inhibits the activities of TET2 (Fig ueroa *et al.*, 2010).

There is experimental evidence to suggest that all the TET proteins play a functional role in the specification of ES cells during development and differentiation (Xu *et al.*, 2012). It has been shown that only the member of the TET family, TET3, is highly expressed in oocytes and zygotes and within the paternal pronucleus, where it is

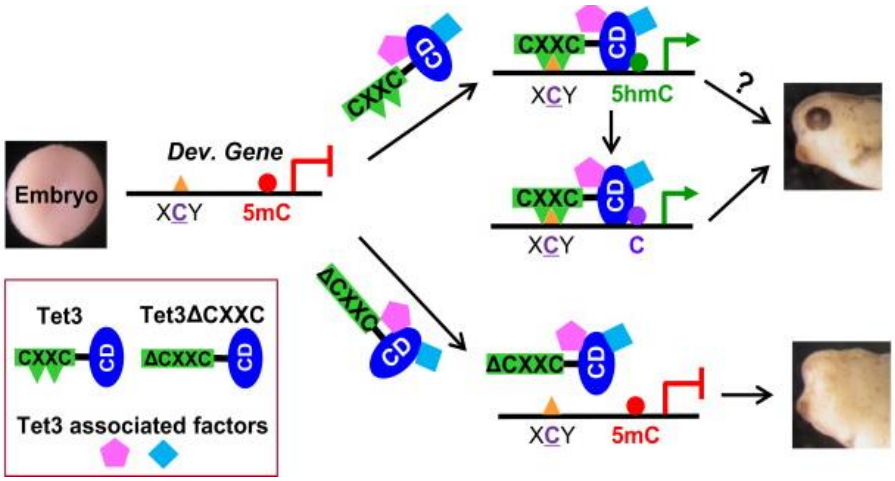


Figure 3: A model representing TET3 dependant gene regulation in Xenopus Laevis. TET3 with intact CXXC domain allows specific binding to target gene promoter allowing hydroxylase activity. TET3 with mutant CXXC (ΔCXXC) loses the specificity of TET3's ability to bind to its target gene promoters therefore prohibiting embryonic development activities. CD = catalytic domain (Xu *et al.*, 2012)

responsible for the conversion of 5mC to 5hmC (Gu *et al.*, 2011; Iqbal *et al.*, 2011). Although more research is required to confirm the functional role of TET3, there has been a study which identified TET3

to be a transcriptional regulator important during early eye and neural development in the frog, Xenopus laevis, as TET3 directly triggers the vital genes required for the development process (Xu *et al.*, 2012).

The findings generated by Xu et al. 2012 resolved the structure of TET3 and its CXXC domain, suggesting that the CXXC domain within TET3 allows specific DNA binding for target gene regulation by binding to target gene promoters and allowing the conversion of 5mC to 5hmC by hydroxylase activity of the catalytic domain (Fig 3).

3.3.2 OXIDISED FORMS OF 5MC

5hmC was initially detected in mammalian DNA in 1972 (Penn *et al.*, 1972). However, 5hmC gained significant interest in 2009 as two studies carried in mouse brain and embryonic stem cells demonstrated substantial levels of 5hmC in these systems (Kriaucionis and Heintz, 2009), (Tahiliani *et al.*, 2009).

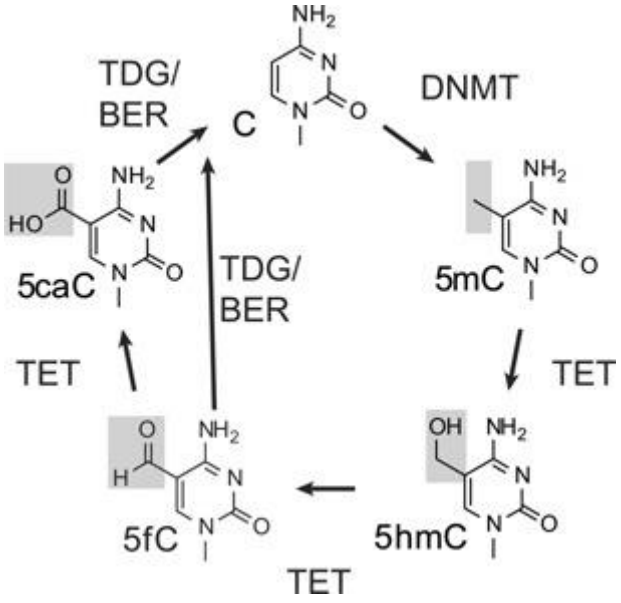


Figure4: DNA demethylation pathway displaying the oxidative forms of 5mC. Taken from Szulwach & Jin 2014

In the brain, there is an increased level of 5hmC with the highest content of this modification in hypothalamus (Wen and Tang, 2014). As well as by being an intermediate of the demethylation pathway, 5hmC has also been shown to be an epigenetic modifier directly regulating gene expression (Stroud *et al.*, 2011) producing an open chromatin structure on the DNA (Mendonca *et al.*, 2014). There are high levels of 5hmC in pluripotent cells (Choi *et al.*, 2014). However glioblastoma studies have shown that TET1 and 5hmC are contributing factors for tumorigenesis as there are elevated levels of both in proneural glioblastomas (Takai *et al.*, 2014). The study undertaken by Takai et al. 2014 showed that knockdown of TET1 in GBM leads to decreased proliferation, sphere formation and tumorigenicity; however by adding a wild-type TET1 with no mutation, proliferation, sphere formation and tumorigenicity was restored as well as the expression of 5hmC in the same glioblastoma cells. This suggests that TET1 and 5hmC may be important for the pathogenesis of the tumour.

The further oxidation of 5hmC by the TET family of enzymes causes the production of 5-formylcytosine (5fC) (Ito *et al.*, 2011). Studies conducted by Pfaffeneder et al. have provided evidence of the increased accumulation of 5fC in mouse embryonic stem cells however decreases after stem cell differentiation, in which it was hypothesized that this marker may be a contributing factor for development and germ cell programming (Pfaffeneder *et al.*, 2011). 5fC also has been shown to play an important role in establishing cell-specific development (Sadakierska-Chudy, Kostrzewa and Filip, 2015). A study confirmed this by providing evidence of the increased presence of 5fC but not 5hmC or 5mC on CGI-containing promoters,

which was in direct correlation with gene expression (Raiber *et al.*, 2012).

5-carboxylcytosine (5caC) is another product of the oxidation of 5hmC mediated by TET enzymes (Gu *et al.*, 2011). Studies have shown the deficiency in TET3 has an impact on the demethylation of OCT4 and Nanog paternal genes and insufficient levels of TET3 in oocytes has a reduced effect on the ability to reprogram injected somatic nuclei (Gu *et al.*, 2011), suggesting the importance of 5hmC, 5fC and 5caC marks (DNA demethylation) in development and reprogramming. Moreover, a genome-wide distribution analysis of 5fC/5caC in mESCs conducted by Shen et al. 2013 via TDG knockdown, indicates the extensive involvement of TET/TDG mediated active demethylation occurring within mammalian genomes. Further studies involving the knockout of TDG indicated the importance of its role in embryonic development and epigenetic aberrations during cell lineage commitments (Cortázar *et al.*, 2011). These studies correlated with results of Wheldon et al. 2014, when there was an accumulation of 5caC-positive cells during differentiation of neural stem cells (NSCs). These results suggest that the accumulation of 5caC can be an indication of active demethylation required in epigenetic reprogramming for determining lineage specification. 5caC distribution within cancer has become an emerging topic, as studies in the past have shown that there is a potential loss of 5hmC within human cancer tissues (Haffner *et al.*, 2011), (Lian *et al.*, 2012). Therefore, it would be ideal to study the oxidised forms such as 5fC and 5caC in a cancer context. There was a recent study performed by Eleftheriou et al. 2015 showing the increase of 5caC within breast cancers and glioma and therefore, indicating the possibility for the involvement of DNA demethylation

within cancer (Eleftheriou *et al.*, 2015). However further study is still warranted to determine the role of 5caC and active DNA demethylation in pathogenesis of cancer.

3.4 Active Demethylation of DNA in Cancer

Defects in DNA methylation and demethylation are crucial and have significant effect on the development of malignancies. As cancer involves the activation and inactivation of certain genes for the progression of tumorigenesis, studies have indicated the methylation and demethylation pattern of these certain types of genes can be a contributing factor to cancer pathogenesis (Szyf, Pakneshan and Rabbani, 2004).

In cancer, hypermethylation occurs in CpG rich regions of the genome and, most usually, in promoters (Wajed, Laird and DeMeester, 2001). When methylation increases in promoter regions of certain genes, such as tumour suppressor genes, their expression is silenced therefore causing uncontrollable cell proliferation as the tumour suppressor gene can no longer suppress tumorigenic genes from being actively expressed (Wajed, Laird and DeMeester, 2001). Hypermethylation causes a permanent inactivation in the expression of certain genes causing long-term silencing (Kalari and Pfeifer, 2010). Studies showed the hypermethylation of genes involved in normal development such as cyclin-dependant kinase inhibitor 2A (CDKN2A) which are responsible for stem cell numbers and cell cycle functions; genes involved in proper epithelial development such as GATA binding protein 4 and 5 (GATA-4 and -5); also, the anti-apoptotic gene, death-associated protein kinase (DAPK). The identification of the areas within the genome that are hypermethylated is crucial as it can be a useful tool for the verification of the tumour cells and may

possibly aid to research targeted to therapy of genes involved in tumour initiation and progression (Fukushige and Horii, 2013).

The increased loss of DNA methylation, also known as hypomethylation, has been shown to activate genes which are not required, possibly leading to the progression of tumour (Herceg and Ushijima, 2010). Hypomethylation was one of the first epigenetic alterations identified in human tumours in 1983. Global hypomethylation is frequently observed in cancer development (Ehrlich, 2009). It can be observed in very early stages of tumorigenesis becoming more pronounced during tumour progression (Ehrlich, 2009). There have been many studies on different types of cancers such as glioblastoma (GBM), ovarian epithelial carcinoma, cervical cancer, etc. which has shown global hypomethylation (Ehrlich, 2009). Nonetheless the concept of hypomethylation, such as active demethylation, involved with the progression of tumours is still being studied in-depth to get a better understanding of its biological significance (Kushwaha *et al.*, 2016).

3.5 GBM

Glioblastoma is an aggressive form of brain tumours, designated by WHO (World Health organisation) as IV astrocytoma accounting for up to 15% of all brain tumours (Zhang *et al.*, 2012). It is known to be a highly heterogeneous tumour with scattered variation of genetic loss across the genome affecting most of its chromosomes (Friedmann-Morvinski, 2014).

3.5.1 EPIDEMIOLOGY

GBM has been ascertained as sporadic, like many other types of tumours. Mostly diagnosed within elder patients with an average age

ranging from 62-64 years however it does occur in children and it is very erratic and accounts to approximately 3% of all brain tumours (Thakkar *et al.*, 2014). Moreover, it has been indicated that there are higher frequencies of GBM within men compared with women. GBM is mainly developed within the supratentorial region of the brain which are the frontal, temporal, parietal and frontal lobes, and very rarely observed in the cerebellum and spinal region (Thakkar *et al.*, 2014). GBM patient's survival rates are very low, with an average of 12-14 months, depending on the progression and metastasis of the tumour (Thakkar *et al.*, 2014).

3.5.2 PATHOPHYSIOLOGY

There are two clinical classes of GBM which are primary and secondary GBM. Primary GBM are *de novo*, therefore occurring spontaneously with no pre-existing lesions (Ohgaki and Kleihues, 2013). It is a common type of GBM occurring mostly in elder patients typically over the age of 50 (Evska, Evar and Komel, 2013), and its genetic characteristics are frequently identified as EGFR amplification, PTEN mutation and loss of chromosome 10 (Ohgaki and Kleihues, 2013). Secondary GBM is commonly developed by slow progression from low grade astrocytoma tumours (Ohgaki and Kleihues, 2007) and anaplastic astrocytoma and usually accounts for approximately 40% of glioblastomas. Secondary glioblastoma is a less commonly occurring type of glioblastoma and mostly affects younger patients (Ohgaki and Kleihues, 2007). Although both are histologically indistinguishable, they present different genetic alterations and molecular mutations (Brem and Abdullah, 2016). Mutation in tumour suppressor protein (*p53*), isocitrate dehydrogenase 1 (*IDH1*), mesenchymal epidermal transition (*MET*) amplification and overexpression of platelet-derived

growth factor receptor A (*PDGFRA*) can be used to identify secondary glioblastomas (Alifieris and Trafalis, 2015). Regardless of these mutations, both clinical forms, primary and secondary have similar end results as the signalling pathways are altered and response to standard treatment methods (Alifieris and Trafalis, 2015).

3.5.2.1 MOLECULAR SUBTYPES OF GBM

The Cancer Genome Atlas (TCGA) has subdivided this tumour into four molecular subclasses due to the expanding research findings in the molecular heterogeneity. These are classical, mesenchymal, neural and proneural. Each molecular subtype harbours different genetic alteration and mutations. A classical tumour is characterised by the abnormal overexpression of epidermal growth factor (EGFR) receptor which are proteins which send signals to keep growing, as well as EGFR mutation and chromosome 10 loss (Aum *et al.*, 2014). The mesenchymal subtype is characterised by low level expression of NF1, an NF1 gene deletion or point mutation (Bhat *et al.*, 2013). Along with the NF1 alteration, there is increased expression of tumour necrosis factor (TNF) family genes, which explains the reason behind increased levels of necrosis in the mesenchymal tumour also high levels of expression in the genes involved in the NF-kB pathway (Bhat *et al.*, 2013). The neuronal subtype shows intermediate expression characteristics between proneural and mesenchymal tumours. However, there is unique expression of neuron marker genes, which are associated with neuronal differentiation and shows overexpression of EGFR like the classical tumours (Verhaak *et al.*, 2010). The proneural subtype is mostly associated with secondary glioblastoma and it is characterised by point mutations in the isocitrate dehydrogenase genes (*IDH1* and *IDH2*) and mutations in the

tumour suppressor gene, *p53* (Goodenberger and Jenkins, 2012). There is also high amplification of platelet derived growth factor receptor A (*PDGFRA*); however in tumours that do not show *PDGFRA* abnormalities, it has been found *P13K* mutations are commonly seen (Verhaak *et al.*, 2010). Furthermore, IDH mutated glioblastomas have been studied to show unique DNA methylation pattern called CIMP (CpG-island methylator phenotype) and usually CIMP positive tumours are classified as proneural; however not all proneural tumours have CIMP (Abel, T., K. Aldape, S. Clark, C. Vnencak-Jones, 2015). Some alterations are in multiple subtypes with different expression levels (Verhaak *et al.*, 2010).

3.6 DNA Methylation Patterns in GBM

Methylation in GBM has been and is still being extensively studied. The heterogeneity and the epigenetic characteristics of glioblastoma prompted many studies which discovered that the outcome of gene expression and patterns possibly gives rise to the different glioma subgroups (Zheng *et al.*, 2011). Numerous studies have identified that hypermethylation mostly occurs in secondary GBMs which led to the hypothesis that this information could be used to distinguish between secondary and primary GBMs. GBM are characterised by genome wide hypomethylation along with specific regions of hypermethylation (Thon, Kreth and Kreth, 2013). Due to this pattern of methylation, there has been activation of oncogenes and silencing of tumour suppressor genes such as *TP53* and *PTEN* in conjunction with increased levels of genomic instability (Thon, Kreth and Kreth, 2013).

One of the characteristics of methylation in GBM is seen in O(6) – methylguanine-DNA-transferase (*MGMT*), which is a DNA-repair gene (McNamara, Sahebjam and Mason, 2013). *MGMT* is responsible for the alkylation of guanine which induces cell death and apoptosis by double-strands breaking and base mispairing, which protects normal cells from carcinogenesis; however, the methylation on the promoter of MGMT causes silencing impairing DNA repair (McNamara, Sahebjam and Mason, 2013). Nonetheless, this factor has been taken advantage of for treating patients with chemotherapy drugs consisting of alkylating agents such as temozolomide, therefore the silencing of the *MGMT* gene can allow the drug to damage the DNA of tumour cells (McNamara, Sahebjam and Mason, 2013).

However, there are forms of methylation which do not have a prognostic outcome. Moreover, it has been shown that there is more promoter methylation occurring in secondary GBM than primary GBM (Nagarajan *et al.*, 2014). The Cancer Genome Atlas (TCGA) performed a pilot study on the epigenetic alterations in GBM and identified promoter methylation of DNA in groups of patients with repetitive methylation pattern called glioma-CpG island methylator phenotype (G-CIMP phenotype) (Roszkowski *et al.*, no date). This hypermethylation of G-CIMP is associated with a proneural subtype and IDH mutations (Roszkowski *et al.*, no date).

3.6.1 *IDH* MUTATIONS & GLIOMA-CIMP G-CIMP

Mutations in IDH1 and IDH2 genes commonly occur in GBM. The functions of these genes are to produce NADPH and alpha-ketoglutarate. (Zhang *et al.*, 2013) NAPDH is essential for cellular

functions and defence against oxidative damage and alpha-ketoglutarate plays a vital role as a cofactor for a family of oxygenases such as ten-eleven translocation enzymes (TETs) (Zhang *et al.*, 2013). In GBM, it is reported that there is mutation in the IDH1, when the arginine in R132, within the enzyme's substrate binding site, is replaced by histidine (R132H) causing the mutation. Very rarely in IDH2, when the arginine in R172, also in the substrate binding site, is replaced by lysine, methionine, glycine or tryptophan by point mutation, causing the enzyme to lose its function. This loss in function of the IDH1/2 causes reduced production of NAPDH and alpha-ketoglutarate compared to the wild-type IDH1/2. Therefore, the loss in alpha-ketoglutarate results in the TET enzyme not being provided with oxygen which prevents demethylation by oxidation of 5mC to 5hmC (Ito *et al.*, 2011).

IDH1 mutations are mostly harboured in secondary glioblastoma occurring in younger patients and had improved prognosis as opposed to older patients carrying wild-type IDH1 (Parsons *et al.*, 2008). The mutations in IDH1 is closely associated with G-CIMP positive tumours (Fig ueroa *et al.*, 2010). CIMP was first identified and is still being studied extensively in colorectal cancer however it is now being used to describe the increased methylation activity in CpG promoters for other tumours such as GBM (Hughes *et al.*, 2013). A glioma-CIMP (G-CIMP) is a classification of glioma based on distinct clinical and genetic features (Hughes *et al.*, 2013). The mutations in IDH1 and the production of 2-hydroxyglutarate (2-HG) has been hypothesised to cause hypermethylation since 2-HG inhibits the TET enzyme activity therefore not allowing demethylation to occur hence large regions of genes stay methylated resulting in G-CIMP (Kondo *et al.*, 2014).

However, this still needs to be studied further to conclude that TET enzyme activity inhibition is the major cause for aberrant DNA hypermethylation, although studies do show a promising link.

3.7 Demethylation in GBM

Very recently, there have been studies to show the involvement of active DNA demethylation in GBM causing the activation of oncogenes and resulting in genomic instability (Tabu *et al.*, 2008). An example of an oncogene which undergoes hypomethylation in GBMs is melanoma antigen gene, *MAGEA1* (melanoma-associated antigen 1) which belongs in the *MAGE* family (Cadieux *et al.*, 2006). Moreover, GBM also shows global hypomethylation of the tandem repeat satellite (*Sat2*) and of the insulin-like growth factor 2 (*IGF-2*) gene causing it to activate gene activity therefore inducing tumour development (Natsume *et al.*, 2010). Activation of certain genes cause increased proliferation of tumour cells resulting in rapidly increased tumour mass, however the exact molecular pattern of hypomethylation still needs extensive studies to understand the process and alterations involved in the cause of activation of demethylation (Nagarajan *et al.*, 2014). Moreover, there have been studies to show the effect of global hypomethylation on the activation of certain stem cell markers such as CD133, also known as prominin-1, as the hypomethylation pattern may be causing the development or maintenance of brain tumour-initiating cells, BTIC (Tabu *et al.*, 2008). The expression of CD133 may serve as a marker for stem cell populations such as cancer stem cells.

3.8 Cancer Stem cells (CSCs)

Stem cells have the potential to self-renew and differentiate into specialised cells (Yu *et al.*, 2012). CSCs are known to be a subpopulation of cells which obtain the stem-like properties to self-renew and differentiate however also possess cancer cell characteristics. CSC have been studied extensively for over 10 years however their definition and importance are controversial (Rosen and Jordan, 2009). CSCs can be distinguished by the different genetic structure and alterations in the gene expression. The role of stem cells within cancer began in 1994 in a study on acute myeloid leukaemia, when AML initiating cells from AML patients were transplanted in severe combined immune-deficient (SCID) mice (Lapidot *et al.*, 1994). Cancer stem cells can cause the re-occurrence of tumours preventing traditional treatment methods from working effectively, as the traditional treatments only have the capabilities to target the cells within the tumours which are rapidly proliferating but not for the slowly progressing cancer stem cells (Shukla *et al.*, 2017). There have been recent studies to show the presence of cancer stem cells in solid tumours such as breast, brain, pancreas, liver, etc. using the cell surface markers expressed which includes CD133, CD44, etc. (Munoz, Iliou and Esteller, 2012). Also, these studies showed heterogeneous cell populations within cancer stem cells which have the capabilities of phenotypic changes within the CSCs as the tumour progresses.

3.9 Glioblastoma Stem cells (GSCs)

Glioblastoma has been one of the many tumours which have been studied for the presence of CSCs suggesting that these cells are the cause of the tumour presenting itself again after resurrection and

chemo- and radio- therapy treatment (Bar, 2011). GBM stem cells have been identified to display similar characteristics to normal neural progenitors also expressing certain neural stem cell markers (Schonberg *et al.*, 2014). Along with the ability to self-renew and prolonged proliferation rate, they also have the proficiency to form neurospheres (Schonberg *et al.*, 2014). Epigenetics plays a vital role in the regulation of adult and embryonic stem cells by controlling transcription to maintain differentiation patterns and self-renewal (Herceg and Ushijima, 2010).

3.9.1 GLIOBLASTOMA STEM CELL MARKERS

One of the first markers proposed to be associated with GSCs is the neural stem cell marker, CD133 (Lathia *et al.*, 2015). Previous studies have shown that the CD133 has the potential to cause the occurrence of tumour growth when transplanted *in-vivo* into immunocompromised mice (Singh *et al.*, 2004). Nonetheless, there has been emerging evidence of the presence of additional markers which are expressed to enhance GBM, but CD133 remains to be the quite popularly studied marker (Schonberg *et al.*, 2014). Another marker used to characterise GBMs is CD15, also known as stage-specific embryonic antigen-1 (SSEA-1), which is initially expressed on embryonic and adult CNS stem cells but is also expressed in GBM. There are many other markers which are also expressed such as transcription factors, cytoskeletal proteins, post-transcriptional factors and Polycomb transcriptional suppressors (Jin *et al.*, 2013). Nonetheless, some of the other interesting markers i.e. SOX2, Nestin, Olig2, Notch1 and EGFR, which will also be studied in this project are elaborated further.

3.9.1.1 CD133

As a surface marker expressed on hematopoietic stem cells, CD133 has been studied extensively to be established as one of the commonly identified CSC markers in many tumours, including GBM (Sibin *et al.*, 2015). It can be used to isolate neural stem cells from human fetal brain (Essential, Stem and Maintenance, 2013). CD133 has been recently discovered to be able to change its subcellular location between the cytoplasm and the plasma membrane of neurosphere cells (Sibin *et al.*, 2015). It was initially thought that in GBM, only CD133+ positive cells can cause the reoccurrence of the tumour when the tumour biopsy material was inserted into mice (Tabu *et al.*, 2008). However, recent studies have shown that CD133- cells also have the capability to initiate the reformation of brain cancer but the proliferation rate is not as rapid as with CD133+ cells (Wang *et al.*, 2008).

3.9.1.2 SOX2 (SEX-DETERMINING REGION Y-BOX)

SOX2 is the commonly studied marker from the SOX family and it belongs to the SoxB1 family (Feng, 2015). It is responsible for the regulation of self-renewal of embryonic stem cells and maintaining stems cells from the neural system and skin. SOX2 expression is upregulated during early stages of embryonic development however as the neural cells differentiate, the expression becomes downregulated (Feng, 2015). It is also an important factor employed for the generation of induced pluripotent stem (iPS) cells (Takahashi and Yamanaka, 2006). The expression of SOX2 is controlled by extracellular signals and intracellular cofactors and occurs in the inner cell mass, initially beginning with expression in both the inner cell

mass and the trophectoderm in the modular stage (Sarkar and Hochedlinger, 2013). Previous studies have shown that epiblast formation does not occur within the zygote after deletion of SOX2 therefore leading to early embryonic death, which suggests the importance of SOX2 in embryonic development (Avilion *et al.*, 2003). During early stages of neural development, studies have shown the expression of SOX2 is essential for the commitment of neural progenitor cell (NPC) fate and this is regulated by the STAT3 signalling pathway that subsequently regulates Nestin expression (Foshay and Gallicano, 2008).

However, it has been reported that the expression of SOX2 is elevated in cancer in relation to normal tissues. In GBM for instance there was an increase in expression of over 85% (Alonso *et al.*, 2011). Immunohistochemistry studies showed the presence of SOX2 in the nuclei of GBM cells (Annovazzi *et al.*, 2011). A study performed on breast cancer patients also showed these results when immunohistochemistry was carried out on breast cancer cells; there was strong SOX2 expression within the nuclei in comparison to weak signalling in non-tumorigenic mammary epithelial tissues (Rodriguez-Pinilla *et al.*, 2007). Moreover, the expression of SOX2 in GBM neurospheres correlates with the tumour grade (Berezovsky *et al.*, 2014). SOX2 is expressed in the proneural subclass of GBM and known to have high levels of promoter hypomethylation (Berezovsky *et al.*, 2014).

3.9.1.3 NESTIN

A neural stem cell marker, Nestin, is classed as a grade 5 intermediate filament protein, expressed in many cell types during development

(Matsuda *et al.*, 2013). The expression of Nestin is downregulated and is replaced by neurofilament once differentiation occurs (Matsuda, Hagio and Ishiwata, 2013). It is expressed mostly in migrating and proliferating cells during early embryogenesis, however in adult tissue, Nestin expression is mainly localised to specific tissues/organs to support them to regenerate (Wiese *et al.*, 2004). It has also been acknowledged that Nestin expression can be used to identify mesenchymal stem cells derived from the bone marrow and they possess progenitor activities of mesenchymal stem cells such as fibroblastic colony-forming units and CFU-Fs and also the proficiency to self-renew and trilineage differentiation (Mendez-Ferrer *et al.*, 2010). Nestin has been seen to be re-expressed during many repair processes such as infarcted myocardium (El-Helou *et al.*, 2005).

Previous studies have also reported increased expression of Nestin in various cancer cells, with a possible correlation with poor prognosis (Matsuda *et al.*, 2013). It is also a commonly expressed marker in GBM (Lu *et al.*, 2011). The increased expression in GBM has been shown to result in lower survival rate due to higher grade glioma occurrences. Therefore Nestin is now being used commonly as a cancer stem cell marker due to its expression levels with the progression of GBM (Lu *et al.*, 2011; Bradshaw *et al.*, 2016). Moreover, studies have indicated that Nestin-positive cells migrate into glioblastoma cells giving rise to vasculature of the tumour therefore possibly characterising Nestin as a vital component within the tumour environment (Najbauer *et al.*, 2012).

3.9.1.4 OLIG2

Olig2 is one of the members of the basic helix-loop-helix (bHLH) transcription factors (Takebayashi *et al.*, 2000). It is important for the development of oligodendrocytes and a number of studies identified that expression of Olig2 is observed within immature neuronal and multipotential neuron/glia progenitors (Takebayashi *et al.*, 2000). These are usually identified during CNS development to give rise to the motor neurons and oligodendrocytes (Malatesta *et al.*, 2017). The regulation and phosphorylation of Olig2 at a site within the bHLH domain allows the ability to abruptly switch from maintaining immature progenitors of motor neurons (pMN) to developing motor neurons (Li *et al.*, 2011).

Olig2 has a unique factor regarding its functions within development. There are two classifications for bHLH transcriptional factors which are anti-neurogenic (pro-mitotic), essential for the prevention of cell cycle exit during early stages of development, and neurogenic (anti-mitotic), essential for cell cycle arrest, specification and differentiation (Kageyama and Nakanishi, 1997). However, Olig2 incorporates both functional characteristics of these classifications. It behaves as a neurogenic factor during differentiation of motor neurons and oligodendrocytes, but functions as an anti-neurogenic factor when maintaining the state of pMN progenitors to accomplish replication of gliogenesis (Lee *et al.*, 2005). This presence of anti-neurogenic factor functions within Olig2 suggested the possible relationship between Olig2 expression and the presence of “tumour-initiating” stem cells in glioma (Meijer *et al.*, 2012). Olig2 is not only expressed within stem cell progenitors in the normal brain, but also in subpopulations of gliomas (Verhaak *et al.*, 2010). Furthermore it has been suggested

that Olig2 plays a role in the defining molecular subclasses of glioblastoma, for example, the removal Olig2 supposedly causes change of proneural phenotype, where PDGFRA is enriched, to classical phenotype, where EGFR is enriched, however further studies are mandatory to understand this definitively (Leelatian and Ihrie, 2016).

3.9.1.5 NOTCH1

The *Notch* gene was originally discovered in the *Drosophila Melanogaster* in 1913, and it was identified as a sex-linked mutation showing a phenotype for serrated wing margin (Dexter, 1914; Morgan and Bridges, 1916). Since then, there has been an expansion in the field of *Notch* and its role within biological evolution (Artavanis-Tsakonas and Muskavitch, 2010).

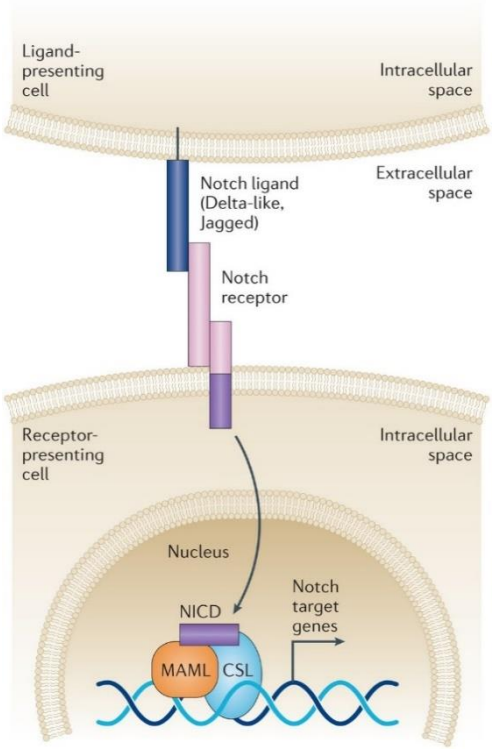


Figure 5: The Notch signalling pathway showing the simple mechanism involving a cell with a ligand inducing proteolytic cleavage events in Notch receptor upon contacting another cell. The Notch intracellular domain (NCID) is released; translocates to the nucleus and binds to transcription factor, CBF1, suppressor of Hairless, Lag-1 (CSL) and Mastermind-like (MAML) transcriptional co-activator genes. Together, transcription of Notch targets genes is activated. *http://www.nature.com/nrd/journal/v13/n5/full/nrd4252.html* (Andersson and Lendahl, 2014) *(14-Sept-2017)*

Notch1 is one of the four (*Notch1-4*) *Notch* family of transmembrane proteins which serve as receptors for membrane-bound proteins, *Notch* ligands (Yamamoto, Schulze and Bellen, 2014). There are two classes of *Notch* ligands which are Jagged proteins and Delta-like proteins (Andersson and Lendahl, 2014). These are involved in the *Notch* signalling pathway which is responsible for development and maintenance of self-renewing adult tissues (Kofler *et al.*, 2011). *Notch* signalling occurs during close cell-cell interaction due to the fact that the cell membrane consists of the relevant ligands and receptors, in the form of transmembrane protein, needed to activate the pathway (Yamamoto, Schulze and Bellen, 2014).

There are three proteolytic steps involved in mechanism of the canonical *Notch* signalling pathway, the first one being the S1 cleavage step mediated by *Furin* which results in a heterodimer form of *Notch* on the cell surface consisting of the ligand binding *Notch* extracellular domain (NECD) and the signalling domain, *Notch* intracellular domain (NICD) (Lake *et al.*, 2009). Following this is the S2 cleavage phase where the NECD undergoes lysosomal degradation when they are internalised via endocytosis by the ligand-binding cell (Bray, 2006). The S1 cleavage causes the S3 cleavage, where the NICD is cleaved from the *Notch1* receptor by y-secretase (Guo *et al.*, 2015), to translocate to the nucleus where it binds to the CSL and MAML and initiates the transcription of the *Notch* target genes (D'souza, Meloty-kapella and Weiaster, 2010).

Notch signalling's involvement in cancer was initially shown in T-cell Acute Lymphoblastic leukaemia (T-ALL) (Weng *et al.*, 2004). Mutations in Notch1 affected the C-terminal domain in Notch1 which is

responsible for the proteasome mediated degradation of Notch1 (Thompson *et al.*, 2007). Following this finding, involvement of Notch1 in the pathogenesis of other solid tumours (Lobry, Oh and Aifantis, 2011) such as GBM was discovered (Kanamori *et al.*, 2007). Notch signalling is important for the normal maintenance of brain development such as neural stem cell renewal, differentiation, memory, learning and gliogenesis (Lasky and Wu, 2005). Kanamori et al. (2007) have proposed the idea that contribution of RAS-induced transformation of glial cells to glioma growth and survival, may be activated by Notch signalling pathway. It is also reported that Notch1 and its ligands, Jagged-1 and Delta-like 1 are being overexpressed in GBM and vital for the growth of the tumour (Purow *et al.*, 2005). These results can demonstrate that Notch1 may represent a potential oncogene (Teodorczyk and Schmidt, 2015).

3.9.1.6 EGFR

Epidermal growth factor receptor (EGFR) is part of the receptor tyrosine kinase superfamily and detectable in cell membranes (Wieduwilt and Moasser, 2008). The signalling pathway activated by EGFR contributes in development and in maintaining mammalian cells such as during growth, survival, proliferation and differentiation (Wiley, Shvartsman and Lauffenburger, 2003). This growth factor receptor located at the cell surface activates a tyrosine kinase when a ligand is bound to it, which causes a number of phosphorylation activities to occur leading to cell growth, DNA synthesis and expression of certain oncogenes (Voldborg *et al.*, 1997). There are a number of ligands which can bind to EGFR such as EGF and transforming growth factor-α (Marquardt *et al.*, 1983).

It has also been proposed that an amplification of EGFR signalling can cause uncontrolled cell growth resulting in malignant phenotypes (Voldborg *et al.*, 1997). Studies have suggested EGFR and EGF-like proteins are overexpressed in human cancer which possibly causes cell transformation (Normanno *et al.*, 2006). It has also been implied that there are mutations within EGFR that causes the receptor to be permanently activated in cancers such as glioma (Wong *et al.*, 1992). The mutation occurring within the gliomas are commonly in relation to in-frame deletions which cause structural rearrangements in the extracellular domain of the receptor (Ekstrand *et al.*, 1992).

In GBM, EGFR overexpression has been shown to promote tumour growth, migration, angiogenesis and metastasis (Tortora *et al.*, 2007). The overexpression of EGFR is commonly found in primary GBM (Watanabe *et al.*, 1996) and correlates with poor survival in GBM patients (Shinojima *et al.*, 2003). Studies have shown that glial tumours which overexpress EGFR also overexpress its active mutant variant, ΔEGFR, which contributes to induced tumour formation, increased proliferation and reduces apoptosis (Nagane *et al.*, 1996). This possibly suggests that ΔEGFR can be an oncogene common in GBM and can perhaps be a target for therapeutic treatment (Nishikawa *et al.*, 1994).

Our preliminary data demonstrated heterogeneous distribution of 5caC in different cells of GBM culture, therefore we questioned if this heterogeneity correlates with any of the GBM markers. In this study we aim to identify the possible functions of the oxidative form of 5mC, 5caC, produced during active DNA demethylation and its involvement in the tumorigenesis in GBM and investigate any potential correlation

between the levels of 5caC and cancer stem cell markers expression in GBM.

4. **MATERIALS**

Immunostaining	
PBS Tablets	Calbiochem 524650-IEA
Anti-Olig2 Antibody, clone 211F1.1, Primary	Mouse mAb, Millipore MABN50
Anti-Nestin Antibody, clone 10C2, Primary	Mouse mAb, Millipore MAB5326
NOTCH1 Monoclonal Antibody (A6), Primary	Mouse mAb, Invitrogen MA5-11961
EGFR Monoclonal Antibody (H11), Primary	Mouse mAb, Invitrogen MA5-13070
SOX2 Monoclonal Antibody (20G5), Primary	Mouse mAb, Invitrogen MA1-014
AlexaFluor 555 Donkey Anti-Mouse, Secondary	Invitrogen, T30954
Alexa Fluor(TM) 633 Goat Anti-Mouse IgG (H+L) highly cross-adsorbed, Secondary	Invitrogen, A-21052
Donkey anti-Rabbit IgG (H+L), HRP, Secondary	Invitrogen, A-16035
TSA Plus Fluorescein System	Perkin Elmer, NEL741001KT
Chamber Slide, Lab-Tek, Glass, 8-well	Thermo Scientific Nunc, 177402
SLS Coverslip No 0, 22x64mm	Scientific Laboratory Supplies, MIC3208
RNA Extraction	
RNeasy Mini Kit - mini spin columns, collection tubes (1.5/2ml), RNase-free reagents & Buffers	QIAgen, 74104
Nanodrop	
Centrifuge	Eppendorf Centrifuge, 5424R
PB Buffer	Qiagen 19066
cDNA Synthesis	
Block Heater	Stuart Blockheater, SBH200D
5x first strand	Invitrogen, 1862765
GoTaq - qPCR master mix	Promega, A6001
SuperScript® III Reverse Transcriptase	Invitrogen, 18080044
Deoxynucleotide (dNTP) Solution Mix	NEB, N0447S

Random Hexamers	Thermo Fischer, N8080127
qPCR	
MicroAMP Fast Optical 96-well Reaction plate	Applied Biosystems (Life Tech), 4346907
Optical Adhesive Covers	Applied Biosystems (Life Tech), 4360954
PCR Machine	7500 Real Time PCR System & Software v2.0.6
Cell Culture	

All chemicals and reagents were purchased from Sigma-Aldrich, unless otherwise stated.

4.1 Primer

Name of Primer	Sequence
Beta-III Tubulin F	CCTCAAGATGTCCTCCACCTTCAT
Beta-III Tubulin R	GTTCATGTTGCTCTCGGCCTC
CD15_SSEA1_FUT4 F	GACGACTTCCCAAGTGCCTC
CD15_SSEA1_FUT4 R	CAAGTTCCGTATGCTCTTGGGC
CD133_Prominin1 F	GAACTCCTTTTCAGGAGGGCAG
CD133_Prominin1 R	GTATCTTCTGGGAAATCACGCGG
GFAP F	GCCAGCTACATCGAGAAGGTTC
GFAP R	CTGTGCCAGATTGTCCCTCTC
GLAST_SLC1A3 F	GGCCAACGAAACGCTTGTG
GLAST_SLC1A3 R	CAAAGAACTCTCTCAGGGCCTG
Hes1 F	GACCCAGATCAATGCCATGACC
Hes1 R	CCCGTTGGGAATGAGGAAAGC
Nestin F	GGAAGAGGAAGAGAACCTGGGAAA
Nestin R	CTCAGATTCAGCTCTGCCTCATC
NF1 F	CAGTGCAACAGGTGGCTTG
NF1 R	GACAGCATCAGCATGTAGCG
Notch1 F	CTTTGAGACTGGCCCACCTC
Notch1 R	CCGCAGAGGGTTGTATTGG
Olig2 F	GATGACCTTTTTCTGCCGGCC
Olig2 R	GTGCTGGACGAGGATGACTTG
PDGFRA F	GAGATGCTTTGGGGAGAGTGAAG

PDGFRA R	GCCTGCCTTCAAGCTCATTCTC
SOX2 F	CACATGAACGGCTGGAGCAAC
SOX2 R	GTAGGACATGCTGTAGGTGGG

5. METHODS

5.1 Cell Culture

Four glioblastoma cell lines; U251, LN18, LN229 and U87MG; one medulloblastoma cell line, UW228-3; one fetal brain cell line, FB83, provided by City Hospital, Nottingham and one human induced pluripotent stem cell (hiPSCs), Rebl-Pat were used. LN18 and LN229 were cultured in DMEM supplemented with 10% fetal bovine serum (FBS). U87MG and U251 were cultured in DMEM supplemented with 10% FBS. HiPSCs were cultured in E8 media. Cells were incubated overnight at 37^{o}C and 5% CO_2. Media was replenished 2-3 times per week for cancer cells until they reach 70-80% confluence at which point they are further passaged and split into sub cultures. For Rebl-Pats, media is changed every day and split every 2 days using TrypIE

5.2 Primer Design

Gene sequences were obtained from Ensembl and inputted into SnapGene to create primers with 50-60% GC content, 50-60^{o}C melting temperature and approximately 200 bp long. All primers were checked using OligoCalc, NCSI Blast and UCSC genome browser.

5.3 RNA Extraction

RNA extraction was performed using an RNeasy® Mini Kit, Qiagen® as per the manufacturers protocol. The cell pellet was lysed in 350 ul RLT buffer to which 1 volume of 70% ethanol was added and mixed by

pipetting. This mixture was then added to a Mini spin column and centrifuged for 15s at ≥8000 x *g*. After discarding the supernatant, 350ul Buffer RW1 was added to the column and was centrifuged again for 15s at ≥8000 x *g.* The supernatant was discarded and 10ul of prepared DNase I stock solution combined with 70 ul Buffer RDD following gentle inversion, was added. The DNase mix was added to the column and left at room temperature for 15 mins. Following this, 350 ul Buffer RW1 was added to the column and centrifuged at the same settings. The supernatant was discarded and 500 ul Buffer RPE was added and centrifuged at the same setting. The supernatant was discarded and further 500 ul Buffer RPE was added and centrifuged for 2 mins at ≥8000 x *g*. The supernatant was discarded and the collection tube was replaced and centrifuged for further 1 min. The column was then replaced into a 1.5 ml collection tube and 50 ul RNase-free water was added to column membrane and centrifuged for 1 min. The RNA was collected in the 1.5 ml collection tube. RNA was quantified using Nanodrop™ spectrophotometry measurements at 260/280nm and 230/260nm and obtained samples with values >2 to determine purity and contamination.

5.4 cDNA Synthesis

Combined 2 ug RNA, 1 ul dNTPs, 2 ul Random hexamers and made up total of 14 ul with RNase/DNase free water and heated at 65^{o}C for 5 mins. Following this, the mixture was put directly onto ice and 4 ul of 5x first strand Buffer, 1 ul DTT (0.1M) and 1 ul RT (Superscript III) was added and mixed and the mixture was left at room temperature for 5 mins. It was placed in a block heater for 1 hour at 50^{o}C followed by 15 mins in a block heater at 70^{o}C. Then, it was further diluted cDNA by 10X by adding 180 ul RNase/DNase free water.

5.5 qPCR

Gene expression of LN229, LN18, U251, U87MG and Rebl-Pat was measured via qPCR, experiments were performed using an Applied Biosystems FAST 7500 Real Time PCR machine. A 10 ul reaction volume was prepared with 5ul GoTaq Mastermix, 0.25ul forward and reverse primers of a 10 uM stock (final concentration 0.25 uM), 2.5 ul RNase/DNase free water and 2 ul sample cDNA. Gene expression was measured using comparative CT standard 2hour run at 40 cycles with denaturation at 95^{o}C, annealing at 60^{o}C and extension at 72^{o}C. Gene expression was normalised relative to the levels of GAPDH expression.

5.6 Immunostaining

Immunostaining was performed according to a protocol detailed by Abakir, Wheldon & Ruzov (2015). Cells seeded onto 8 well chamber slides were fixed in 4% PFA. They were permeabilised with PBX and depurated in no HCl and 4N HCl for 1 hour. Between these steps, the slides are washed several times with PBS. Following washes with PBS, slides were then incubated in humidity chamber with blocking solution (10% FBS in PBS), primary antibodies and secondary antibodies for 1 hour with PBT washes in between each incubation. Slides are then incubated in tyramide amplification solution, 1:200 dilutions, for 2 mins, as opposed to 3 mins as per protocol, and immediately washed several times with PBT. Removing any excess, DAPI containing mounting medium was added drop-wise to the slides before being covered with cover slip and sealed with nail polish.

5.7 Confocal Microscopy

The generation of images and analysis of images was carried out as in the previous study (*Ramsawhook A H. et al, 2017)*.

5.7.1 GENERATION OF CONFOCAL IMAGES

Following immunostaining, mounted slides were viewed under the confocal microscope using a 63X oil lens. To compare intensity profiles between cells lines for each marker, laser power and gain remained the same for each channel within each set of analysis.

		Laser Power	**Gain ()**
DAPI (0N)		2.0	625
5caC (488) (4N)		2.5	720
Nestin (555)	*4N*	2.0	715
	0N	2.0	625
Notch1 (555)	*4N*	2.0	865
	0N	2.0	775
Olig2 (555)	*4N*	3.0	900
	0N	2.0	875
SOX2 (555)	*4N*	2.0	800
	0N	2.0	750
EGFR (555)	*4N*	2.0	900
	0N	2.0	795

5.7.2 INTENSITY PROFILING

In the ZEN black, images are analysed in the “Imaging Processing” section. On each image obtained, a line was drawn through several cells (2-3). This line populated a table with the intensity measurements for pixel intensity of the two or three different channels: red, green and blue fluorescence. This table was saved as a .txt file in the appropriate location by right click and save. This process was carried out in triplicate to obtain three profiles. An intensity profile was also generated for a cell of interest, by drawing a line from one end of cell to the other. The table populated was saved the same way. This data was imported and analysed via Excel and PRISM followed by statistical analysis (One/Two-Way ANOVA).

5.7.3 2.5D Intensity Plots

In ZEN black within Image Processing, an image was opened and a cell of interest was selected and a rectangle drawn around it, and that region was cut, producing a new image of the cell of interest. After saving this image, it was sent to ZEN Blue. Within this software, once this image had loaded, on the right, “2.5D” option was clicked, producing a visualization of a 2.5D image of the cell of interest. To save this image, a print screen was taken and pasted into Microsoft Paint where it was cropped and saved.

5.8 Co-localization Analysis

5.8.1 ZEN Black

In ZEN Black within Image Processing, with the image of interest on the left side of the software, tab labelled “Coloc” was clicked. In the lower half of the screen with tools, there was an option for a tool called “Closed Bezier” which was used to encircle a cell of interest. Values for this cell appeared in the table below. A scatter plot was also generated for red versus green channel. This process was repeated, encircling subsequent cells until dataset completed. The table was exported via right click and save to appropriate location. The scatter plot was exported as a TIF file.

5.8.2 Fiji ImageJ

The image of interest was dragged and dropped into the ImageJ software. Under the “Image” section in the “Colour” tab, the channels within the confocal image are split into two images for each channel in black and white. Following this, under “Analyze”, in the “Colocalization” tab, the Coloc2 analysis programme is opened. Without altering the default settings, the analysis was clicked to run.

This generated a PDF file which was saved in the appropriate location. In this PDF file, there was 2D histogram and the value for Pearson's correlation coefficient value, which was used for analysis of the correlation.

5.8.3 Pearson's Correlation Coefficient – EndMemo

Each image of interest is used to generate intensity profiles for each individual nucleus as described in "Intensity Profiling". The tables generate profiles for each channel – the "raw data". In the programme online, there is a box for two variables, x and y. In each variable, the values for each channel were inputted. For x variable, the red channel data was inputted and for y variable the green channel data was inputted. Clicking generate, the Pearson's correlation coefficient was calculated along with a scatter plot graph which were used for analysis of correlation.

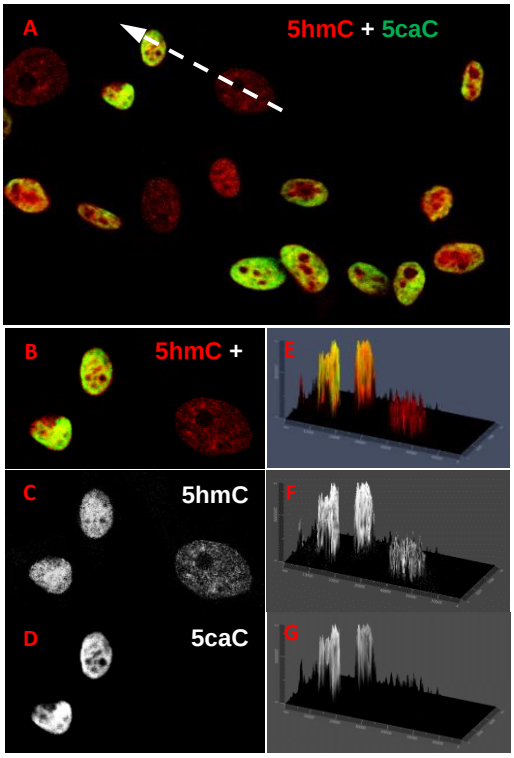
5.9 Statistical Analysis

All graphs were statistically analysed in GraphPad PRISM 7, with one-way or two-way ANOVA.

6. RESULTS

6.1 5caC staining is heterogenous in glioblastoma cell lines but not in induced pluripotent stem cells (iPSCs)

GBM is known to be a heterogeneous tumour, as indicated in its name "multiforme". Previous studies have reported positive 5caC levels within glioma tissues which may suggest that there is involvement of active DNA demethylation in the pathogenesis of brain tumours (Eleftheriou *et al.*, 2015). We revisited this data by performing immunostaining of 5caC on U87MG. To denature the DNA, U87MG was treated with 4N HCl, which allows the 5caC and 5hmC primary antibodies to bind onto the 5caC and 5hmC on DNA effectively, hence the positive staining within the nucleus and no staining present in the cytoplasm (Fig 5).



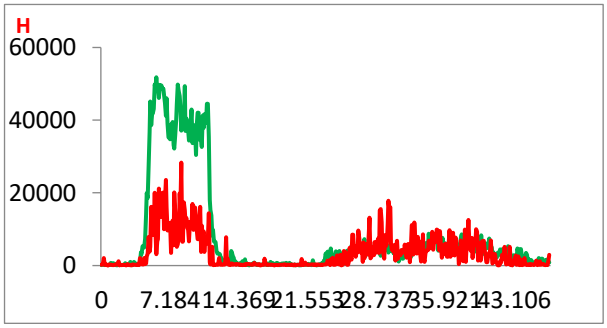


Figure 5: **A** Immunostaining of 5caC and 5hmC on U87MG **B**-**D** Representative images of cells with high and low signal intensity levels of 5caC (green channel) **E-G** 2.5D plot graphs of representative images of cells presenting location of signal within the nucleus **H** Profile intensity graph presenting pixel signal strengths of the different channels for 5hmC and 5caC

The varying signal strengths of 5caC staining intensity on U87MG presents subpopulations of cells with high and low signal strengths which can indicate the possible heterogeneity of 5caC staining in U87MG (Fig 5). These results are in line with results shown by Wheldon et al. 2014, which indicate the heterogeneous pattern of 5caC staining within mouse embryonic stem cells (mESCs) and embryonic brain in comparison to the uniform staining of 5caC shown in induced pluripotent stem cells (iPSCs) from human.

6.2 Studying the expression levels of stem cell markers in GBM and medulloblastoma cell lines by qPCR

As the aim of the study is to see any potential correlation between stem cell markers and 5caC levels. We chose to identify the

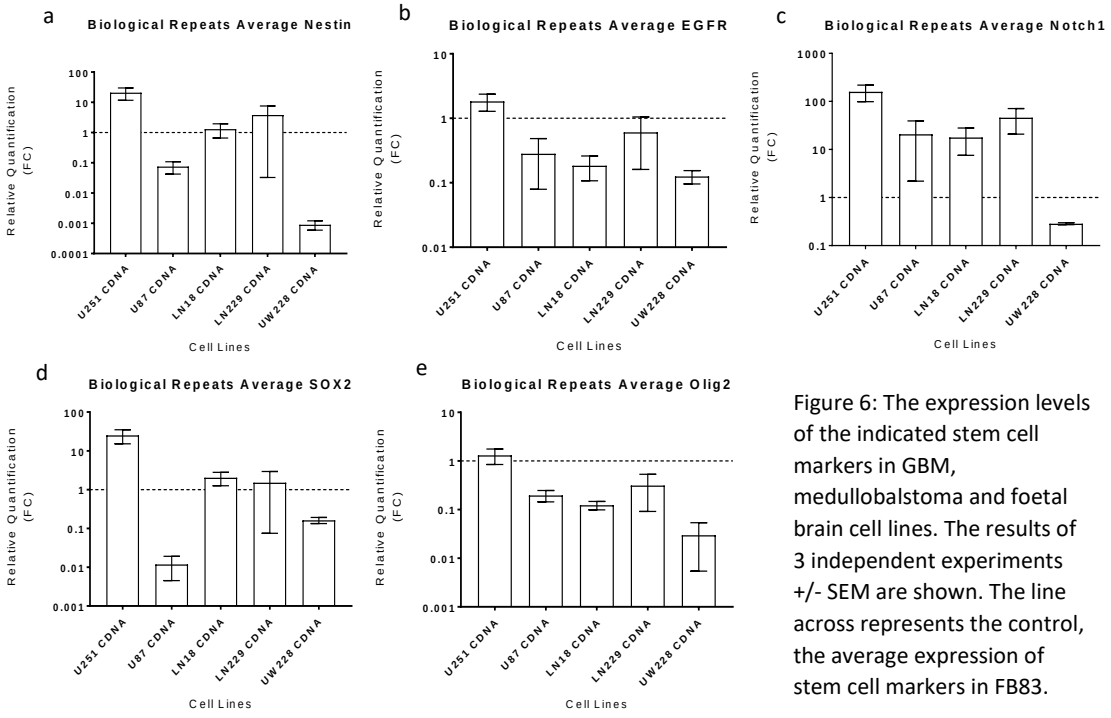


Figure 6: The expression levels of the indicated stem cell markers in GBM, medullobalstoma and foetal brain cell lines. The results of 3 independent experiments +/- SEM are shown. The line across represents the control, the average expression of stem cell markers in FB83.

expression of different stem cell markers in GBM cell lines. The qPCR revealed the expression levels of a selected choice of markers, B3T, PDGFRA, Rb1, MS1, NF1 (Supplementary information, Fig 1), Nestin, EGFR, Notch1, SOX2 and Olig2 (Fig 6) on GBM cell lines, LN18, LN229, U87MG and U251 and a medulloblastoma cell line, UW228-3; results obtained from the qPCR were normalised against the control, a human fetal brain cell line, FB83. We conducted the experiment in three biological repeats and obtained a representable average. The qPCR graphs exhibit high expression levels of all markers in U251 in comparison to FB83. However, UW228-3 displays low expressions of all markers in comparison to FB83. There is notable variation within the levels of expression of each marker in each cell line. In comparison to all the markers, Notch1 exhibits consistently increased expression on all GBM cell lines which possibly indicates its involvement in providing GBM cell lines with stem-like characteristics, reaching up to a relative quantification level of 100.

We selected five markers, EGFR, Nestin, Notch1, Olig2 and SOX2, for further analysis to understand their spatial expression and its suggestive relationship with 5caC distribution in GBM and medulloblastoma cells. The other markers, B3T, PDGFRA, Rb1, MS1 and NF1, all show high transcript levels in GBM and medulloblastoma cell lines, which may deem interesting to be further researched in future studies.

6.3 Validation of the antibodies for these markers

To further understand the spatial distribution of these markers, we performed an immunostaining of the markers on GBM and

medulloblastoma cell lines (Supplementary information, Fig 2-26). The immunostaining confirms the positive expression of the markers within each cell line as also indicated in the previous qPCR study (Fig 7).

The levels of expression of Nestin, Notch1, Olig2 and SOX2 is significantly higher in most of the cells line in comparison to EGFR as the average intensity strength is <16000.

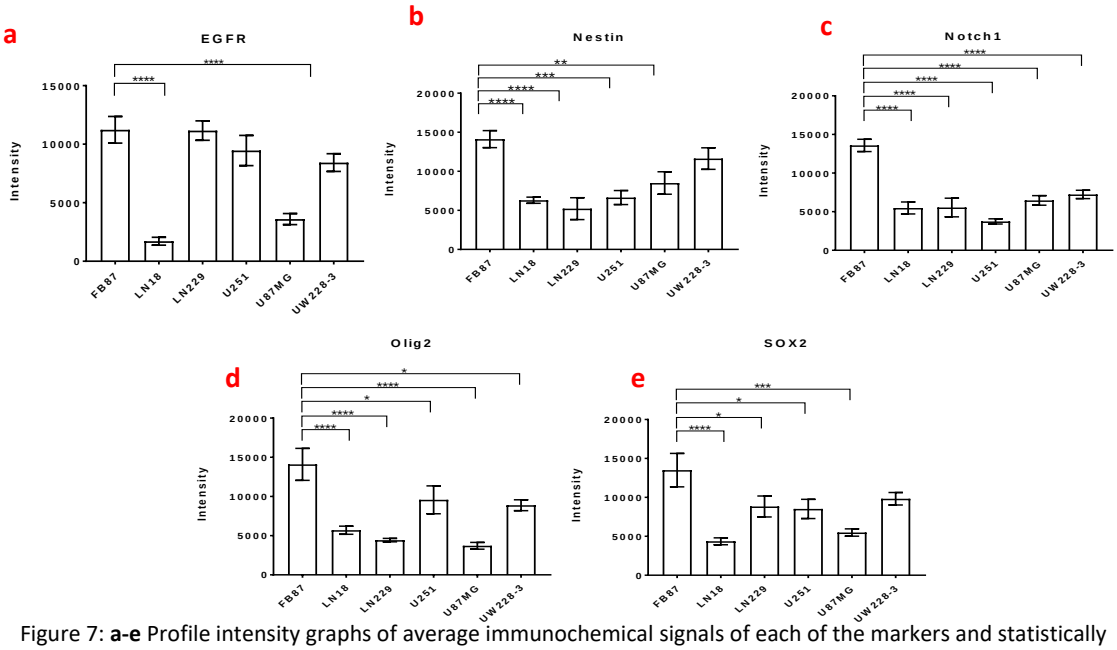


Figure 7: **a-e** Profile intensity graphs of average immunochemical signals of each of the markers and statistically analysed against FB83 as control. Representing the statistical significance of the marker expression in GBM stem cell lines and FB83. 1* being low statistically significant and increasing significance as stars increase.

As the aim of the study is to confirm any possible correlation between the expression of markers with the presence of 5caC, we co-stained the markers on each cell with 5caC. This was carried out by the cell lines being treated with HCl which allows the disulfide bonds to break within the DNA which allows specific binding of 5caC primary antibodies. The co-stain results depict the presence of expression of cell markers with a positive staining of 5caC (Fig 8). However, there is no clear representation of the possible correlation between the markers and 5caC. Further analyses are conducted using the results

obtained by immunostaining, such as 2.5D plots and fluorescence intensity graphs, to develop a better understanding. These will measure the strength of the signal intensity provided by the marker according to the level of expression and how this correlates with the signalling strength of the presence of 5caC, also possibly explaining the heterogeneity of GBM.

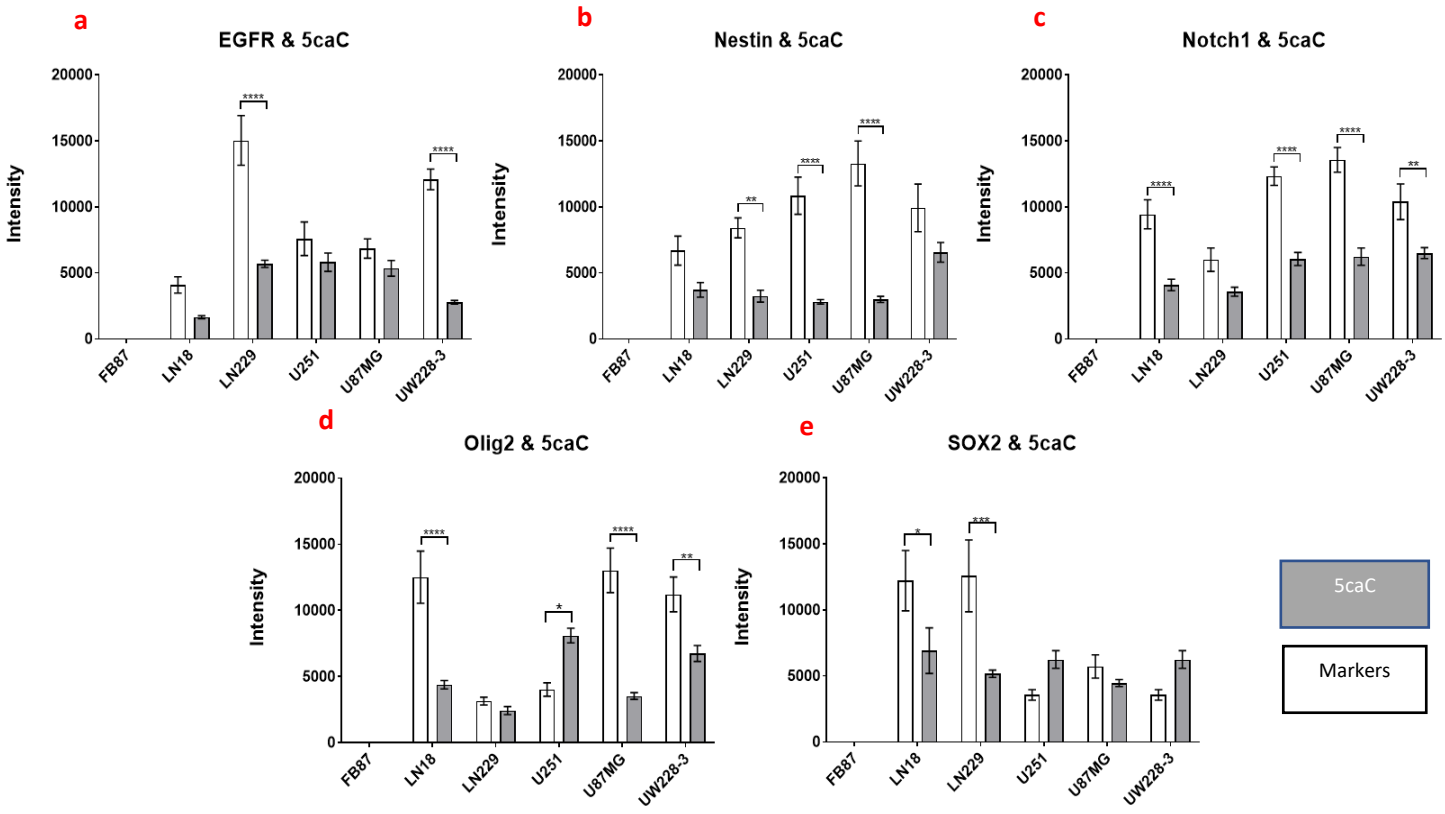


Figure 8: **a-e** Profile intensity graphs of average signals expressed by each marker and 5caC and statistically analysed against each other. Representing the statistical significance between the presence of 5caC and the expression of markers within each GBM stem cell line. Increased stars indicate high significance of results, therefore indicating possible reliable results.

6.3.1 The spatial distribution of the markers' immunostaining in the cells lines

All the markers (EGFR, Nestin, Notch1, Olig1 and SOX2) have shown high staining levels in GBM and medulloblastoma cell lines. Immunostaining each individual marker on the cell lines independently allowed us to understand the pattern of markers' distribution and whether this can possibly identify whether there is a pattern in signal strength when they are co-stained with 5caC.

As EGFR is a transmembrane protein, the results we had obtained depicted the staining pattern for EGFR to be present on the outer membrane of the nucleus. Although there was presence of EGFR staining in the cytoplasm, the signal strength within the outer membrane seemed much stronger (Fig 9).

Visually and in the signal intensity profiles, the signal strength of EGFR in U251 is stronger than that presented by the other cell lines with possibly a lower signal strength in LN18. Moreover, Olig2 also shows increased signal strength within U251 and lower signal strength within LN18. The presence of the signalling is observed strongly in the cytoplasm with fluctuating strengths within the nucleus (Fig 10). Notch1 also shows cytoplasmic positive staining, as the Notch signalling pathways consists of receptors and ligands which are transmembrane proteins (Bray, 2006) (Fig 11).

However, there are significantly different staining patterns for SOX2 and Nestin within the cell lines. SOX2 portrayed a clear nuclear staining with very little signal intensity present in the cytoplasm or the outer membrane (Fig 12), which is expected results for a transcription factor. Staining intensity for each cell line is weaker in comparison to the control, FB83. Additionally, LN229 also depicts low levels of

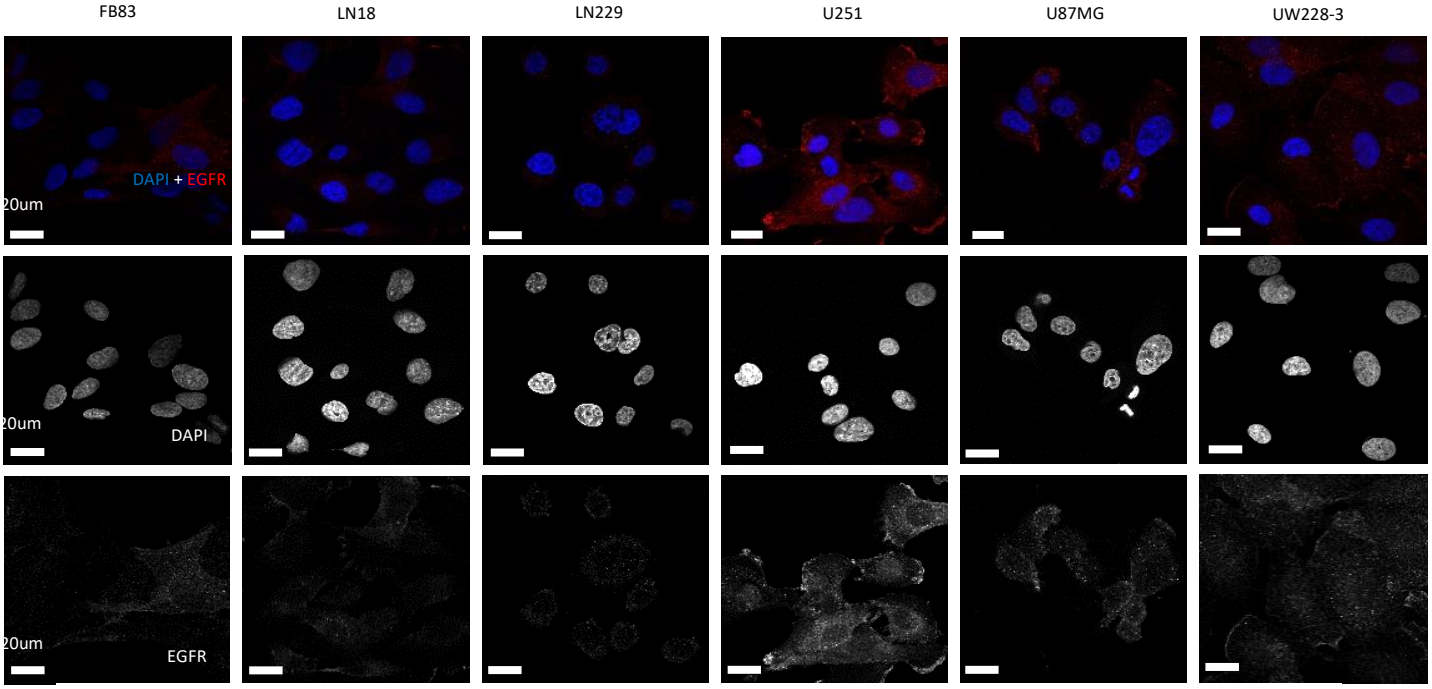


Figure 9: Immunostaining images of the expression of EGFR in GBM and medulloblastoma cell lines. Consists of merge of both DAPI and EGFR (blue and red channels) and split channels in grayscale.

signalling intensity within the cytoplasm as well as the nucleus (Supplementary information, Fig 22-26). Moreover, staining for Nestin positively showed the filament-like structure

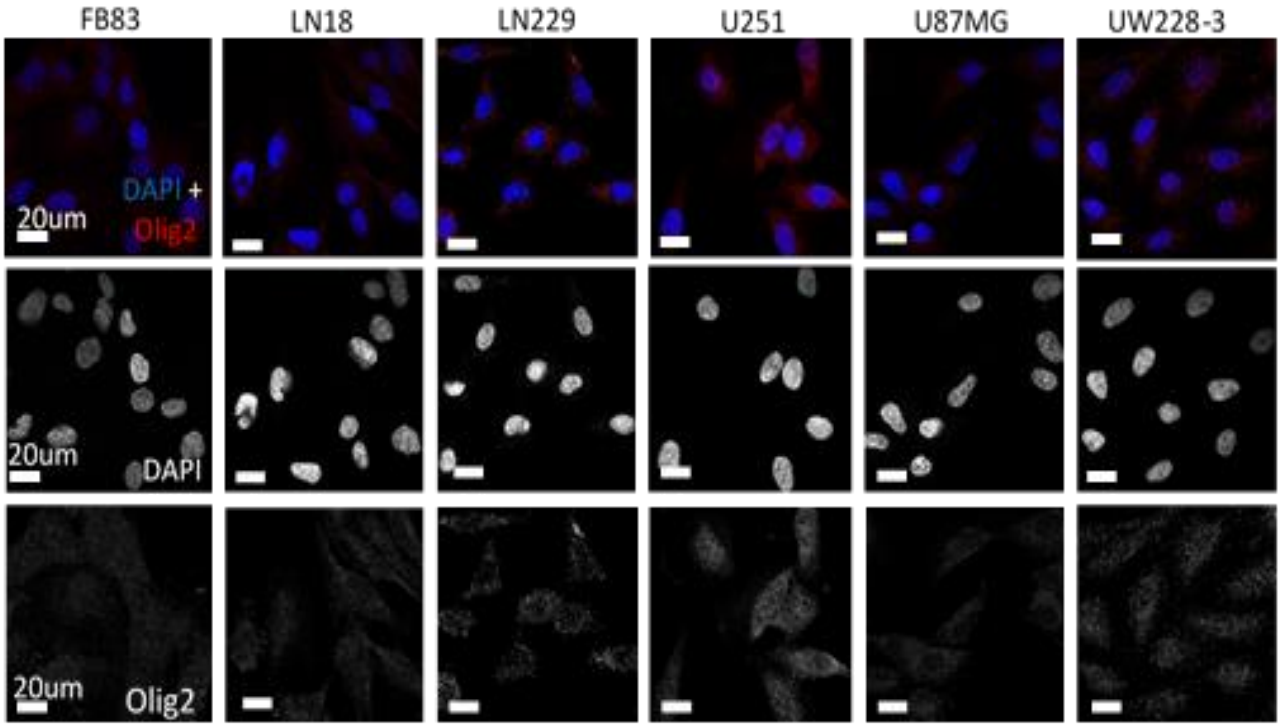


Figure 10: Immunostaining images of the expression of Olig2 in GBM and medulloblastoma cell lines. Consists of merge of both DAPI and Olig2 (blue and red channels) and split channels in grayscale.

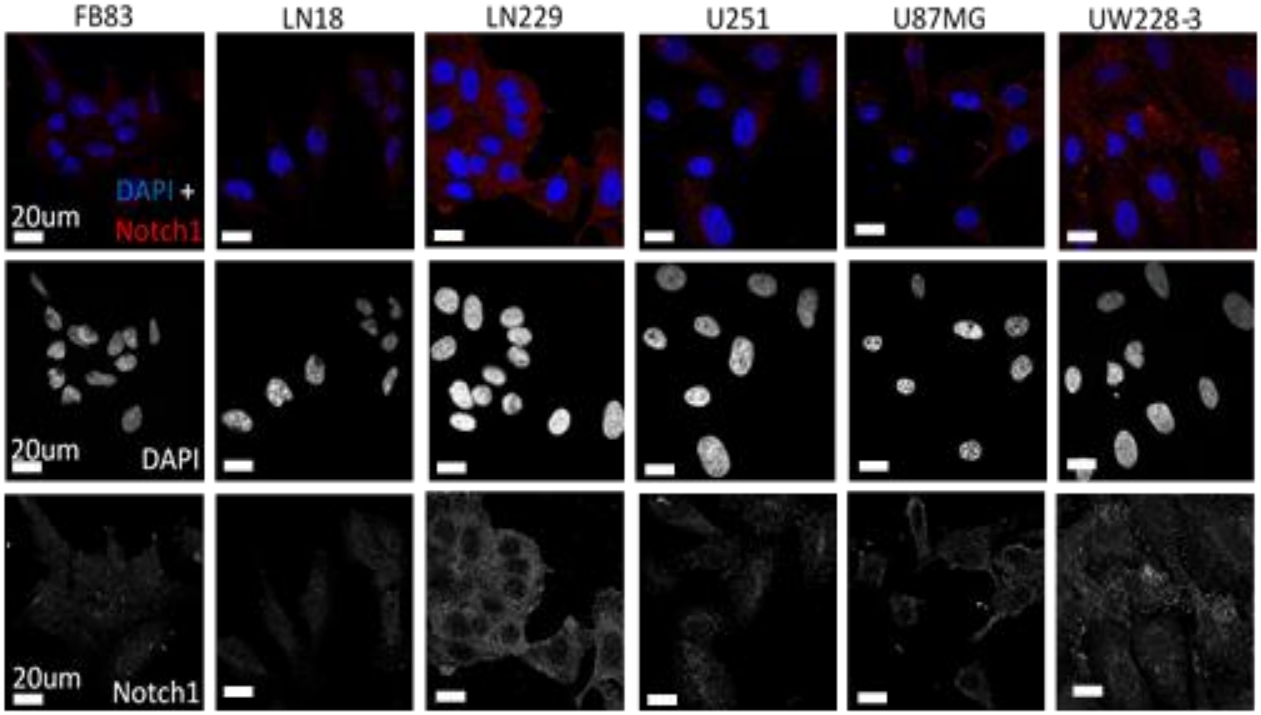


Figure 11: Immunostaining images of the expression of Notch1 in GBM and medulloblastoma cell lines. Consists of merge of both DAPI and Notch1 (blue and red channels) and split channels in grayscale.

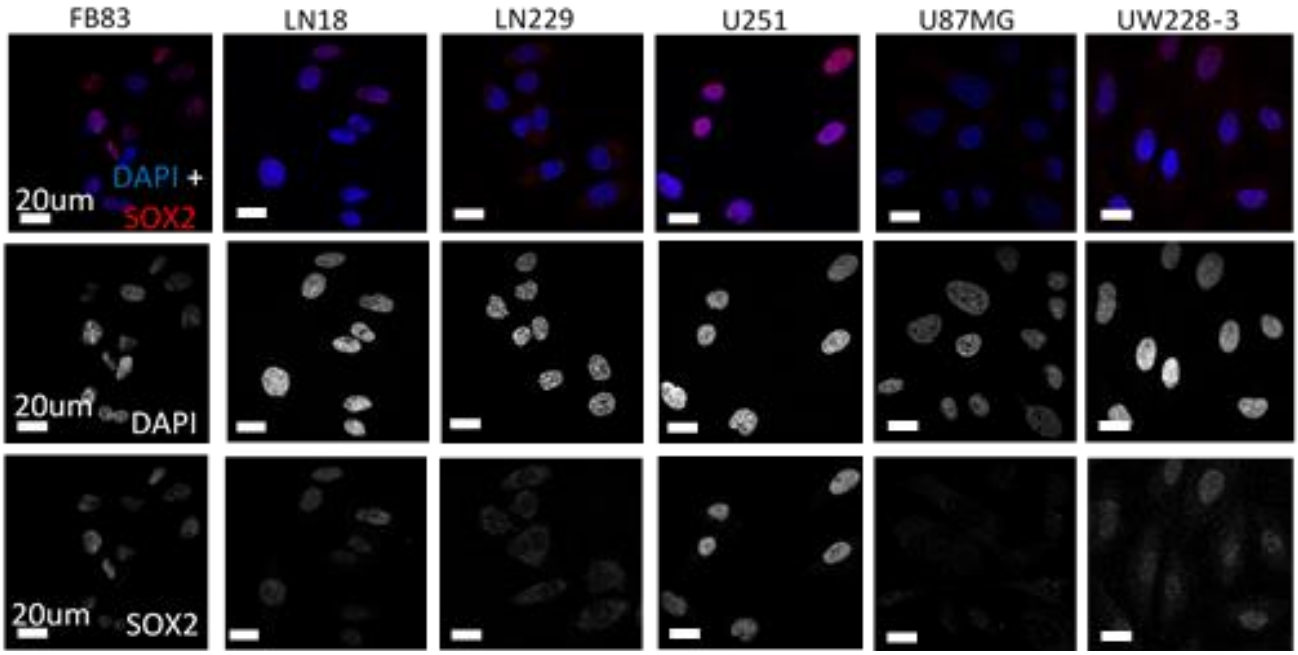


Figure 12: Immunostaining images of the expression of SOX2 in GBM and medulloblastoma cell lines. Consists of merge of both DAPI and SOX2 (blue and red channels) and split channels in grayscale.

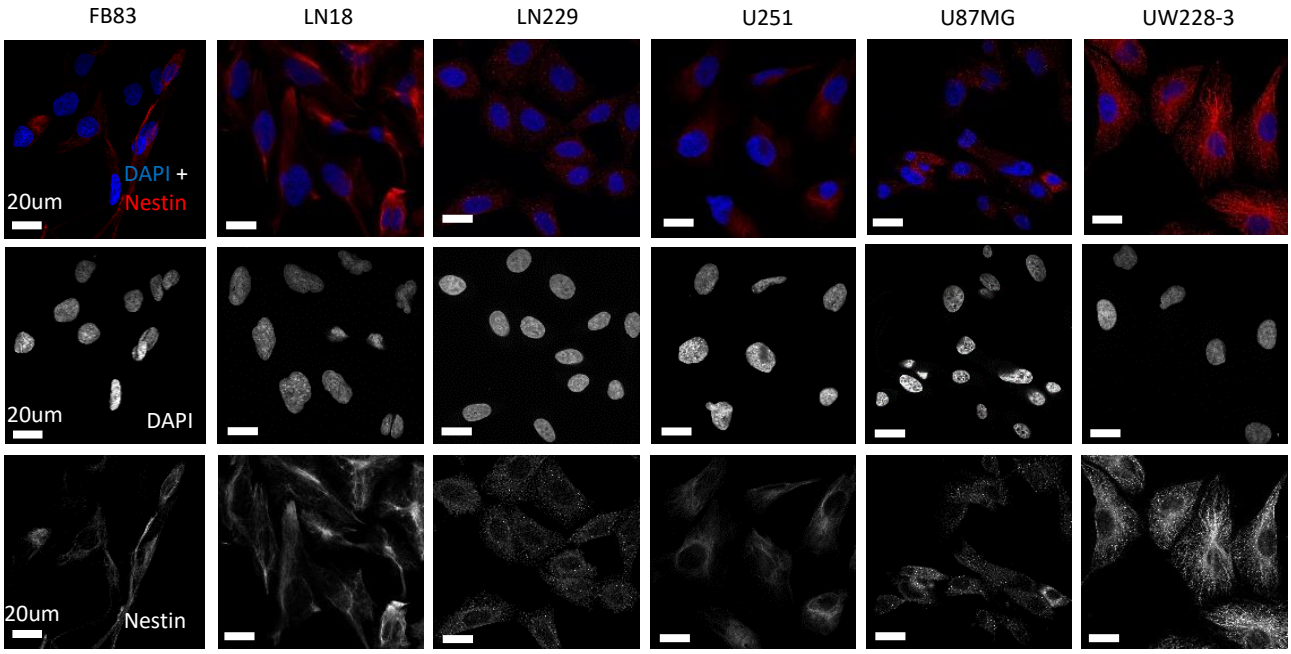


Figure 13: Immunostaining images of the expression of Nestin in GBM and medulloblastoma cell lines. Consists of merge of both DAPI and Nestin (blue and red channels) and split channels in grayscale.

as Nestin is an intermediate filament protein (Fig 13). The staining is present within the cytoplasm. The clarity of the signal intensity for Nestin staining is visible in all the GBM and medulloblastoma cell lines, with varying strengths between each cell line. The analysis showed a positive, statistically significant ($p<0.006$), expression of Nestin within each cell line in comparison to FB83 (Supplementary information, Fig 7-11).

6.3.2 Studying potential correlation between the signal intensities of 5caC and cancer stem cell markers' patterns in GBM and medulloblastoma cell lines

To understand whether the expression of each marker correlates with the presence and levels of 5caC we co-stained GBM and medullablastoma with antibody against 5caC, after treatment of the samples with 4N HCl.

EGFR, Notch1 and Nestin showed varying signal strengths between the cell lines (Supplementary information, Fig 12-14). EGFR generally displayed a statistically significant ($p<0.0001$) trend, as there was a negative correlation between 5caC and EGFR according to the fluorescence intensity graphs shown particularly in LN229 and UW228-3 (Supplementary information, Fig 12). Moreover, Nestin's filament staining intensity visually decreased post-HCl treatment. However, the signal intensity of Nestin in comparison to 5caC exhibited an increasing pattern, therefore as Nestin signal intensity increased, 5caC intensity also increased. This pattern is portrayed mostly in LN18 and U87MG. The significant difference in staining

pattern was observed in Notch1, as post-HCl treatment, there was presence of increased signal intensity in the nucleus but there is no identifiable trend present between the staining intensities of Notch1 and 5caC.

Nonetheless, SOX2 and Olig2 displayed interesting trends and patterns between the staining intensities of the markers and 5caC. Further analysis was conducted using colocalization of pixel intensity and this data was used to calculate the Pearson's correlation coefficient to analyse potential correlation. Pearson's correlation coefficient suggests that when the values are closer to -1 or +1, then the correlation is present, but it is either negative or positive. When the values are close to 0, then this concludes there is no correlation. Pearson's correlation coefficient measures the strength of any potential linear relationship between two variables (Sedgwick, 2012), in this case that would be the green channel's pixel intensity (5caC) and the red channel's pixel intensity (marker).

The signal intensity observed for Olig2 showed correlation with 5caC signal intensity in UW228-3 (Fig 14). There is presence of Olig2 within all the cell lines however it displays no distinctive trend or pattern. In UW228-3, the pattern observed shows Olig2 signal intensity increasing and the signal intensity of 5caC also increases. This pattern is also reciprocated when the signal intensity decreases for Olig2, it

also decreases for 5caC. Based on these results alone, we hypothesize that there may be a positive correlation between Olig2 and 5caC.

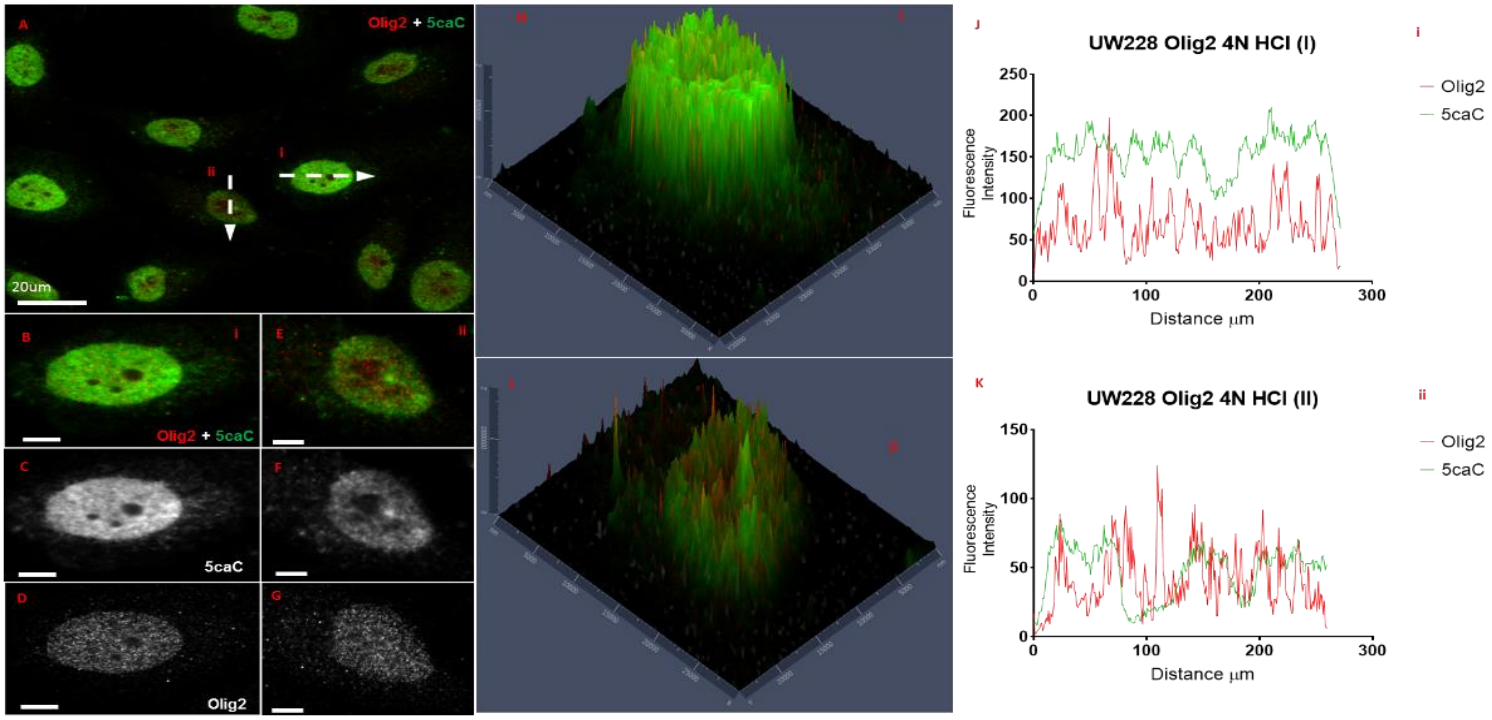


Figure 14: **A-G** Immunostaining images of UW228-3 expressing Olig2 marker and showing the presence of 5caC following treatment with 4N HCl. **i-ii** Single cells representing different levels of 5caC signals. Consists of merge of green and red channels and splits of both channels in grayscale. **H-I** 2.5D plots of **i** and **ii** showing the location of staining for both channels. **J-K** Fluorescence intensity graph showing signals from both channels for **i** and **ii** with high and low levels of 5caC and the difference in marker expression.

To further investigate the possible correlation between 5caC and cancer stem cell markers expression, we conducted studies on medulloblastoma cell line, UW228-3. Medulloblastoma is a paediatric brain tumour mostly occurring in the cerebellum and potentially spreads to the brain and spinal cord (Rutka 1997). Previous studies suggested high 5caC levels in medulloblastoma (Ramsawhook et al. 2017). Previous studies have indicated the possible presence of neuronal cells which suggests a link between neural stem cells and medulloblastoma development (Manoranjan et al. 2012). According to our study, there are high levels of Nestin present within UW228-3. There was also a positive indication of possible correlation between

5caC and Nestin as when 5caC levels are showing high intensity values, the Nestin signal within the cytoplasm shows a low intensity value. This can suggest that the expression of Nestin may be driven by the presence of high levels of 5caC.

Following the analysis of signalling intensities between SOX2 and 5caC shown in the average profile intensity graphs (Fig 15), we had concluded that there may be a potential correlation between the level of 5caC signal intensity being produced and the expression of SOX2. Thus, when 5caC levels decreased, SOX2 expression also decreases.

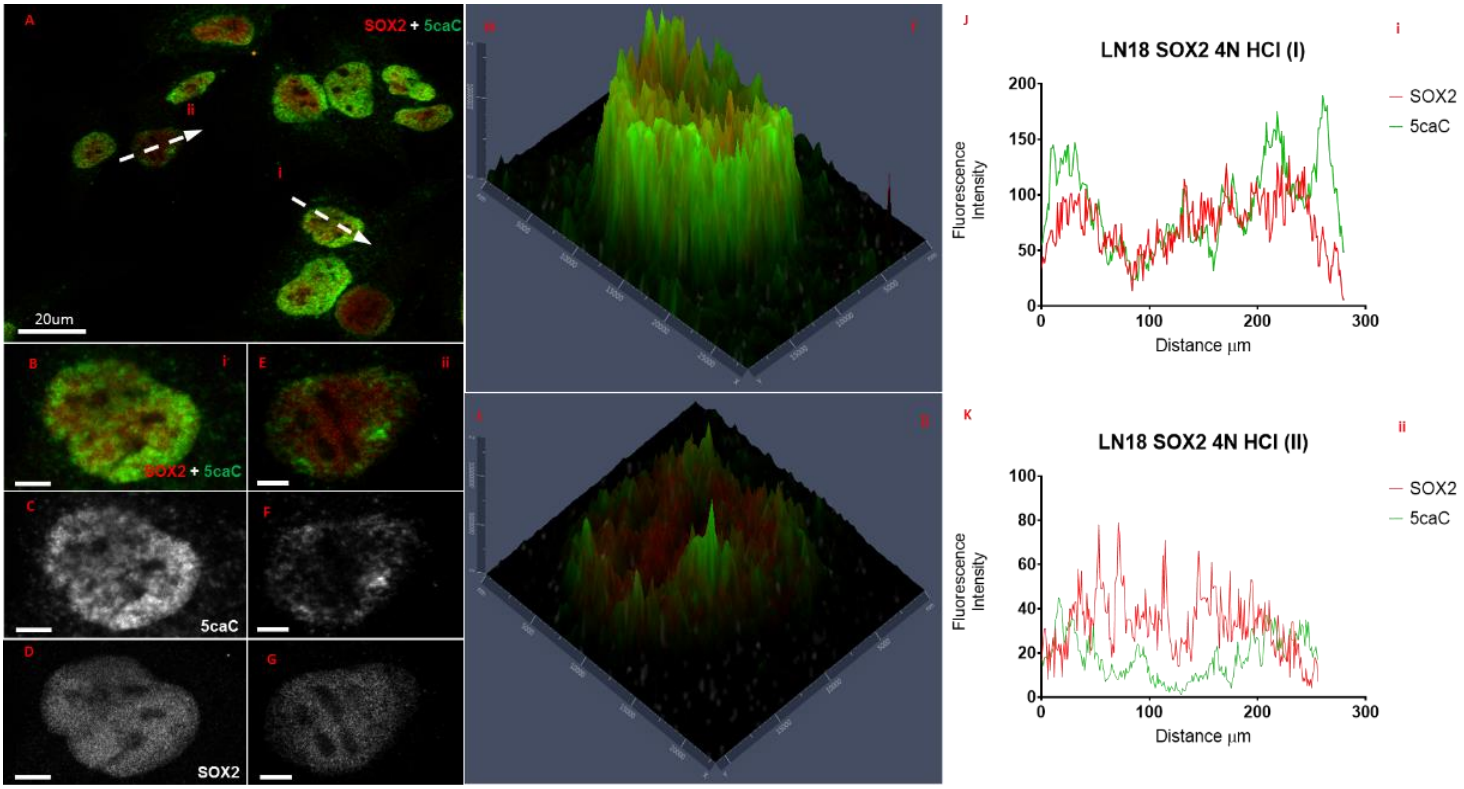


Figure 15: **A-G** Immunostaining images of LN18 expressing SOX2 marker and showing the presence of 5caC following treatment with 4N HCl. **i-ii** Single cells representing different levels of 5caC signals. Consists of merge of green and red channels and splits of both channels in grayscale. **H-I** 2.5D plots of **i** and **ii** showing the location of staining for both channels. **J-K** Fluorescence intensity graph showing signals from both channels for **i** and **ii** with high and low levels of 5caC and the difference in marker expression.

To gain a better understanding of the possible correlation pattern, we conducted a colocalization analysis of SOX2 and Olig2 using ZEN Black (Fig 16D). The histogram scatter plot graph generated for the cells showed no conclusive trend as the values are dispersed along the Green channel (bottom of the graph). Moreover, the table generated coefficient values for each cells pixel intensity and the average of the coefficient R is ~ 0.0009. According Ramsawhook et al. 2017, coefficient R denotes the Pearson's correlation coefficient (Ramsawhook and Lewis, 2017). Therefore, this result suggested the absence of correlation between 5caC and SOX2 levels.

Performing this analysis on Fiji ImageJ via Coloc2 (Fig 16C), the Pearson's correlation coefficient value was equal to 0.79 (Fig 16C), significantly different to that produced by ZEN Black software. Considering this value, we can conclude that there is a positive correlation between the level of 5caC and SOX2 expression. The histogram generated by Fiji ImageJ (Fig 16C) is similar to the one produced by ZEN Black, which can indicate that when the pixel intensity for 5caC increases (green channel), the pixel intensity for SOX2 (red channel) also increases.

As the coefficient values are significantly different for both sets of results (Fig 16C & 16D), we performed similar analysis on a different platform. Using the online Pearson's correlation coefficient calculator by EndMemo (http://www.endmemo.com/statistics/cc.php), we obtained a Pearson's coefficient R value of ~ 0.53 (Fig 16B). This value is relatively closer to that provided by Fiji ImageJ. It's a positive correlation however slightly weaker. The scatter plot graph generated by the calculator shows a trend of increasing 5caC level of intensity and an increasing SOX2 intensity, therefore increasing expression.

A

B

RAW Data Evaluation	SOX2 & 5caC	Olig2 & 5caC
Pearson's R Value	0.534772	0.759266

C

ImageJ-Coloc2 Evaluation	SOX2 & 5caC	Olig2 & 5caC
Pearson's R Value	0.79	0.81

D

ZEN Black Evaluation	SOX2 & 5caC	Olig2 & 5caC
Pearson's R Value	0.000909091	0.066363636

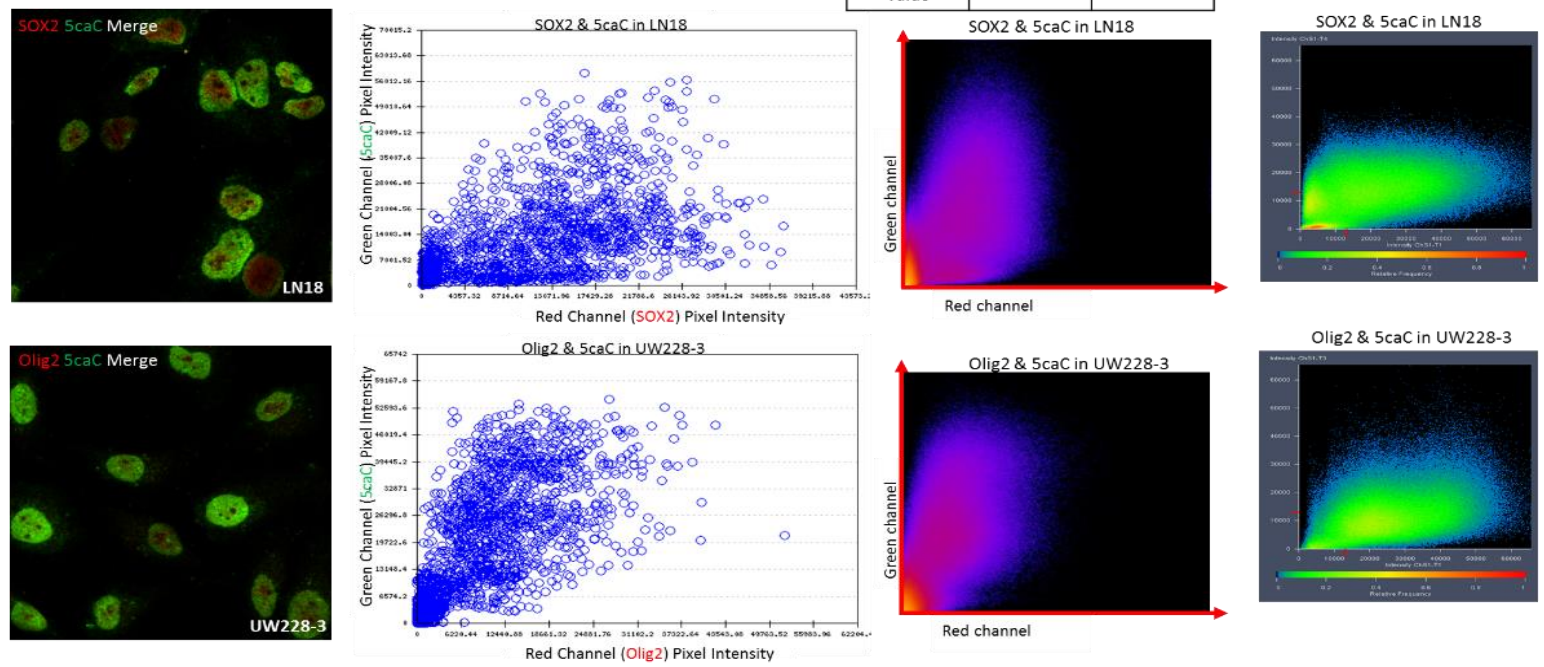


Figure 16: **A** Merged channel images of SOX2 in LN18 with 5caC and Olig2 in UW228-3 with 5caC used to generate correlation data. **B** Correlation coefficient generated by inputting raw data into Pearson's Correlation coefficient calculator on EndMemo (online software); Scatter graph produced by EndMemo showing SOX2 and Olig2's correlation pattern. **C** Correlation coefficient generated by Fiji ImageJ Coloc2 software; 2D histograms produced by Fiji ImageJ Coloc2 showing SOX2 and Olig2's correlation patterns. **D** Correlation coefficient generated by ZEN Black software; Histograms produced by ZEN Black showing correlation pattern of SOX2 and Olig2

Using the ZEN Black software, we conducted an analogous analysis on UW228-3 for 5caC and Olig2 and obtained R of ~ 0.066 (Fig 16D). Taking into consideration the correlation R value denotes the Pearson's correlation coefficient, then this value is close to 0 therefore there is no potential correlation. However, performing this analysis on Fiji ImageJ, the value for Pearson's correlation coefficient was 0.81 (Fig 16C) thus suggesting that there is a possibly strong positive correlation between the level of 5caC and the expression of Olig2. According to the 2D histogram produced by Fiji ImageJ (Fig 16C), it shows a positive trend with increasing pixel intensity of green

and red channel. Also, this histogram looks similar to that produced by ZEN Black (Fig 16D).

The EndMemo generated value for this was equal to ~ 0.76 (Fig 16B). This value is relatively close to the value generated by Fiji ImageJ. Moreover, the scatter plot graph (Fig 16B) reciprocates a similar trend shown in the histogram produced by the other two software.

The ZEN black colocalization analysis possibly produced alternative results because the formula used to generate the correlation coefficient value may vary from the original Pearon's correlation coefficient formula. Furthermore, as the histograms/scatter plot graphs produced by all the analysis are similar, it can also suggest the difference in the formula used via ZEN black.

7. DISCUSSION

Each marker has a specific involvement within GBM and its functionality. There are studies which have shown EGFR overexpressed in glioblastoma causing a number of downstream signalling effects such as impaired apoptosis, angiogenesis and necrosis (Taylor, Furnari and Cavenee, 2012). Moreover, formerly recognised as a neural stem cell marker, studies have shown the presence of expression Nestin in tumour stem cells from brain tumours contributing to the tumorigenicity and resistance to therapies in GBM (Chinnaiyan *et al.*, 2008). The Notch signalling pathway has also shown to contribute in the pathogenesis of cancer, it takes part in the control of cancer stem cell renewal, differentiation, apoptosis and angiogenesis. The Notch pathways may have a contributing factor to cancer stem cells activities such as its growth and invasiveness within GBM (Cenciarelli *et al.*, 2017). Furthermore, the presence of Olig2, a transcription factor, occurs mostly during linear specification of progenitor cells to neuronal subtypes, therefore during the early stem cell stage of central nervous system development (Ligon *et al.*, 2007). Studies have shown that Olig2 is suggested to be a specific marker for glioblastoma stem cells (GSCs) (Trépant *et al.*, 2015). Likewise, there is evidence of SOX2 suggestively being involved in the maintaining of the "stemness" capabilities of cancer stem cells which is potentially associated with GBM (Garros-Regulez *et al.*, 2016). A previous study confirmed the overexpression of SOX2 within GBM may be caused by promoter hypomethylation (Alonso *et al.*, 2011).

In our study, we investigated the spatial distribution of 5caC, a DNA modification that is usually associated with active DNA demethylation

in GBM and medulloblastoma cell lines. Active DNA demethylation is important in several biological processes; however, the exact biological role of 5caC intermediate is not clearly understood (Song & He 2013). Therefore, this study was conducted to understand the potential role of 5caC and its relationship with potential cancer stem cell populations in GBM.

Studies have indicated GBM consists of sub-population of cells which are highly proliferative and some slow proliferative which are believed to be the glioma stem cells (Faria et al. 2006) possibly giving rise to the different subgroups in GBM. It has been previously shown that 5caC is present in a subpopulation of cells of paediatric brain tumours such as medulloblastoma and ependymoma (Ramsawhook et al. 2017). We believed that active DNA demethylation pathway might be involved in tumorigenesis possibly regulating the activities of cancer stem cells. However, there are significant variations in the heterogenic staining of 5caC in the different cell lines, this may be due to the cell lines being derived from different subtypes of GBM.

The protein markers chosen for this study mark stem cell populations regulating either their proliferation potential and or their differentiation into neuronal or glial lineages. As shown in our study, there is expression of all markers within GBM. However, the variation of expression of the markers in each cell line can possibly determine the characteristics of the tumour subtypes. Relative expression of the genes in GBM was compared to their expression in FB83, as fetal tissue is known to contain stem cells and progenitor cells (Ishii & Eto 2014). As identified in qPCR and immunostaining, there is presence of all the marker expression in FB83. Previous studies have showed the presence of stem cells within GBM (glioma stem cells). The co-

expression of markers with 5caC may indicate GBM to possess stemness characteristics such as differentiation, proliferation and self-renewal.

Nonetheless, to visualize the level of the 5caC oxidative derivative of active DNA demethylation in cancer cells we performed immunocytochemistry experiments. We co-immunostain at the same conditions for protein markers; however due to high concentration of HCl the protein structures of some of our markers were affected. Possibly disrupting the protein structure to disrupt and affecting the results. This may have occurred with Nestin as the signal in the cytoplasm decreases following treatment with 4N HCl, due to the structure disruption.

These suggestions stated above can potentially coincide with the results depicted by Notch1 staining with 5caC as it is known that the Notch1 signalling pathway consists of extracellular and intracellular activities (Bray 2006). The increase of the Notch1 expression as well as the increase of 5caC levels in the nucleus may suggest the possibility Notch1 signalling pathway is interpreting active DNA demethylation possibly promoting tumorigenesis of GBM. As Notch1 maintains proliferation and differentiation, permanent activation may cause decreased differentiation hence leading to tumour formation as suggested by previous studies (Stockhausen et al. 2010). This may also suggest that the increased expression can be shown as a permanent activation of Notch1 which is possibly driven by the levels of 5caC, however further studies need to be carried out to confirm this hypothesis.

7.1 The possible correlation between increased levels of 5caC and increased expression of SOX2 and Olig2

SOX2 expression is a characteristic of early human development, however there have been studies that have detected the expression of SOX2 in adult and has seen its presence within neural stem cells throughout its lifespan (Brazel et al. 2005). SOX2 is expressed in many malignant tumours where it is involved in sustaining growth and self-renewal of tumour adult stem cells (Fong et al. 2008)(Suh et al. 2007). Several previous studies have detected the heterogeneity of cells in gliomas at different stages of differentiation correlating with the variable expression of SOX2, ranging from 6-80% in different glioma samples (Schmitz et al. 2007)(Rita et al. 2009).

A previous study conducted by Alonso *et al.*, (2011) to investigate the effect of SOX2 and its contribution towards the malignancy of GBM, shows the presence of hypomethylation occurring in the CpG islands of the promoter regions of SOX2 (Alonso et al. 2011). In our study, we had identified the possible positive correlation between increased expression of SOX2 and the presence of increased levels of 5caC. The correlation also corresponds with decreased levels of 5caC, as the expression of SOX2 also decreases. The correlation shows that as SOX2 is a cancer stem cell marker, 5caC may also be a marker of cancer stem cells. This potential correlation pattern can indicate that 5caC levels drives the expression of SOX2 therefore promoting tumour proliferation and self-renewal.

Additionally, Olig2 has also shown characteristics to possibly identifying diffuse glioma tumour cells and the non-diffuse glioma, pilocytic astrocytoma according to a previous study (Utt et al. 2004).

Olig2 is crucial for determining oligodendrocytes and their specification and it has been reported that overexpression of Olig2 can cause differentiation within oligodendrocyte progenitor cells (OPCs) (Wegener et al. 2015). The expression of Olig2 has been suggested to also identify the classes and subtypes of gliomas (Utt et al. 2004) this can possibly indicate the cause of heterogenic cell population within glioblastoma.

A study conducted by Utt *et al.*, 2004, had described medulloblastoma, derived from cerebellar neuronal progenitor cells, showed mostly Olig2-negative cells however there were focal areas of the tumour which consisted of rare Olig2-positive cells. Considering our finding, there is a potential positive correlation between the increased levels of 5caC and Olig2-positive cells which may suggest that the rare expression of Olig2 can be driven by the presence of active DNA demethylation. In our observation, there is a possible positive correlation between Olig2 expression and 5caC levels in UW228-3. When the level of 5caC increases, the levels of Olig2 also increases. As previous studies have indicated, the presence of Olig2 maintains the tumorigenic state of medulloblastoma as it is found in the progenitors of granule cells in the tumour cells which indicates Olig2 may influence formation and proliferation of cancer cells (Schüller et al. 2008). Therefore, 5caC is present at high levels in cancer stem cells (CSCs) and the following concludes that active DNA demethylation probably occurs in CSC.

To confirm this conclusion of positive correlation between increased levels of 5caC in association with increased expression of SOX2 and Olig2, neurosphere assays can be produced. Neurosphere assays are conducted to enrich for the cancer stem cells from brain tumour.

These stem-like cells possess characteristics of neural stem cells found in normal human brain (Ignatova et al. 2002). This assay will mimic and recapitulate the tumour environment of GBM, providing us more accurate information about tumour characteristics. Using these culture conditions, further experiments can be performed such as immunostaining for the stem cell markers to evaluate their presence in the neurospheres. Moreover, DNA immunoprecipitation can be conducted followed by qPCR to identify the presence of 5caC in the promoter regions of the cancer stem cell markers. Another possible experiment to conduct would be to transiently knockdown SOX2 and Olig2, to examine whether the levels of 5caC are altered and how active DNA demethylation is affected.

In conclusion, this study indicates the potential role of active DNA demethylation in GBM CSC. It is considered due to the heterogeneity of 5caC levels in GBM which shows possible correlating patterns with the expression of CSC markers. These protein markers, i.e. SOX2 and Olig2, are involved in the maintenance and regulation of stem cells and development. Therefore, indicating the stem-like state in GBM cancer cells may contain population of GSCs which are characterised by high levels of active DNA demethylation. With further experiments, these markers may provide support into comprehending the stemness characteristics within GBM and the functional role of 5caC in active DNA demethylation to define the characteristics portrayed by the stem-like cancer cells and potential future implications for diagnostics and therapy of the brain tumours.

8. REFERENCES

Abel, T., K. Aldape, S. Clark, C. Vnencak-Jones, B. M. (2015) *Molecular Profiling of Glioma*, *My Cancer Genome*.

Alifieris, C. and Trafalis, D. T. (2015) 'Glioblastoma multiforme: Pathogenesis and treatment', *Pharmacology & Therapeutics*, 152, pp. 63–82. doi: http://doi.org/10.1016/j.pharmthera.2015.05.005.

Alonso, M. M. *et al.* (2011) 'Genetic and Epigenetic Modifications of Sox2 Contribute to the Invasive Phenotype of Malignant Gliomas', *PLOS ONE*. Public Library of Science, 6(11), p. e26740. Available at: https://doi.org/10.1371/journal.pone.0026740.

Andersson, E. R. and Lendahl, U. (2014) 'Therapeutic modulation of Notch signalling [mdash] are we there yet?', *Nat Rev Drug Discov*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 13(5), pp. 357–378.

Annovazzi, L. *et al.* (2011) 'SOX2 expression and amplification in gliomas and glioma cell lines.', *Cancer genomics & proteomics*. Greece, 8(3), pp. 139–147.

Artavanis-Tsakonas, S. and Muskavitch, M. A. T. (2010) 'Chapter One - Notch: The Past, the Present, and the Future', in Kopan, R. B. T.-C. T. in D. B. (ed.) *Notch Signaling*. Academic Press, pp. 1–29. doi: https://doi.org/10.1016/S0070-2153(10)92001-2.

Aum, D. J. *et al.* (2014) 'Molecular and cellular heterogeneity: the hallmark of glioblastoma', *Neurosurgical Focus*. Journal of Neurosurgery Publishing Group, 37(6), p. E11. doi: 10.3171/2014.9.FOCUS14521.

Avilion, A. A. *et al.* (2003) 'Multipotent cell lineages in early mouse development depend on SOX2 function', *Genes & Development*. Cold Spring Harbor Laboratory Press, 17(1), pp. 126–140. doi: 10.1101/gad.224503.

Bar, E. E. (2011) 'Glioblastoma, Cancer Stem Cells and Hypoxia', *Brain Pathology*. Blackwell Publishing Ltd, 21(2), pp. 119–129. doi: 10.1111/j.1750-3639.2010.00460.x.

Berezovsky, A. D. *et al.* (2014) 'Sox2 Promotes Malignancy in Glioblastoma by Regulating Plasticity and Astrocytic Differentiation', *Neoplasia*, 16(3), p. 193–206.e25. doi: https://doi.org/10.1016/j.neo.2014.03.006.

Bhat, K. P. L. *et al.* (2013) 'Mesenchymal differentiation mediated by NF-kappaB promotes radiation resistance in glioblastoma.', *Cancer cell*. United States, 24(3), pp. 331–346. doi: 10.1016/j.ccr.2013.08.001.

Bhutani, N., Burns, D. M. and Blau, H. M. (2011a) 'DNA Demethylation Dynamics', *Cell*, 146(6), pp. 866–872. doi: 10.1016/j.cell.2011.08.042.

Bhutani, N., Burns, D. M. and Blau, H. M. (2011b) 'Perspective DNA Demethylation Dynamics', *Cell*. Elsevier Inc., 146(6), pp. 866–872. doi: 10.1016/j.cell.2011.08.042.

Bradshaw, A. *et al.* (2016) 'Cancer Stem Cell Hierarchy in Glioblastoma Multiforme', *Frontiers in Surgery*. Frontiers Media S.A., 3, p. 21. doi: 10.3389/fsurg.2016.00021.

Bray, S. J. (2006) 'Notch signalling: a simple pathway becomes complex', *Nat Rev Mol Cell Biol*. Nature Publishing Group, 7(9), pp. 678–689.

Brem, S. and Abdullah, K. G. (2016) *Glioblastoma E-Book*. Elsevier Health Sciences.

Cadieux, B. *et al.* (2006) 'Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetrahydrofolate reductase allele status, and increased proliferation', *Cancer Res*, 66. doi: 10.1158/0008-5472.CAN-06-1547.

Cenciarelli, C. *et al.* (2017) 'The Interference of Notch1 target Hes1 affects cell growth , differentiation and invasiveness of glioblastoma stem cells through modulation of multiple oncogenic targets'.

Cheishvili, D., Boureau, L. and Szyf, M. (2015) 'DNA demethylation and invasive cancer: implications for therapeutics', *British Journal of Pharmacology*. Oxford, UK: BlackWell Publishing Ltd, 172(11), pp. 2705–2715. doi: 10.1111/bph.12885.

Chinnaiyan, P. *et al.* (2008) 'The prognostic value of nestin expression in newly diagnosed glioblastoma: Report from the Radiation Therapy Oncology Group', *Radiation Oncology (London, England)*. BioMed Central, 3, p. 32. doi: 10.1186/1748-717X-3-32.

Choi, I. *et al.* (2014) '5-hydroxymethylcytosine represses the activity of enhancers in embryonic stem cells: a new epigenetic signature for gene regulation', *BMC Genomics*. London: BioMed Central, 15(1), p. 670. doi: 10.1186/1471-2164-15-670.

Cortázar, D. *et al.* (2011) 'Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability', *Nature*, 470. doi: 10.1038/nature09672.

D'souza, B., Meloty-kapella, L. and Weinmaster, G. (2010) 'Canonical and non-canonical Notch ligands', *Current topics in developmental*

biology, 92, pp. 73–129. doi: 10.1016/S0070-2153(10)92003-6.

Dawlaty, M. M. *et al.* (2014) 'Loss of Tet Enzymes Compromises Proper Differentiation of Embryonic Stem Cells', *Developmental Cell*. Elsevier, 29(1), pp. 102–111. doi: 10.1016/j.devcel.2014.03.003.

Dexter, J. S. (1914) 'The Analysis of a Case of Continuous Variation in Drosophila by a Study of Its Linkage Relations', *The American Naturalist*. The University of Chicago Press, 48(576), pp. 712–758. doi: 10.1086/279446.

Ehrlich, M. (2009) 'DNA hypomethylation in cancer cells', *Epigenomics*, 1(2), pp. 239–259. doi: 10.2217/epi.09.33.

Ekstrand, A. J. *et al.* (1992) 'Amplified and rearranged epidermal growth factor receptor genes in human glioblastomas reveal deletions of sequences encoding portions of the N- and/or C-terminal tails.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 89(10), pp. 4309–4313.

El-Helou, V. *et al.* (2005) 'Resident nestin+ neural-like cells and fibers are detected in normal and damaged rat myocardium.', *Hypertension (Dallas, Tex. : 1979)*. United States, 46(5), pp. 1219–1225. doi: 10.1161/01.HYP.0000187888.39665.d9.

Eleftheriou, M. *et al.* (2015) '5-Carboxylcytosine levels are elevated in human breast cancers and gliomas', *Clinical Epigenetics*, 7(1), p. 88. doi: 10.1186/s13148-015-0117-x.

Essential, I., Stem, G. and Maintenance, C. (2013) 'CD133 Is Essential for Glioblastoma Stem Cell Maintenance', pp. 857–869. doi: 10.1002/stem.1317.

Evska, I. J. O. V. Č., Evar, N. K. O. Č. and Komel, R. (2013) 'Glioma and glioblastoma - how much do we (not) know ? (Review)', pp. 935–941. doi: 10.3892/mco.2013.172.

Faria, J. *et al.* (2006) 'Interactive properties of human glioblastoma cells with brain neurons in culture and neuronal modulation of glial laminin organization', *Differentiation*, 74(9), pp. 562–572. doi: https://doi.org/10.1111/j.1432-0436.2006.00090.x.

Feng (2015) 'Overview of the roles of Sox2 in stem cell and development', *Biological Chemistry*, p. 883. doi: 10.1515/hsz-2014-0317.

Feng, S., Jacobsen, S. E. and Reik, W. (2010) 'Epigenetic reprogramming in plant and animal development', *Science (New York, N.Y.)*, 330(6004), pp. 622–627. doi: 10.1126/science.1190614.

Figueroa, M. E. *et al.* (2010) 'Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation', *Cancer Cell*, 18(6), pp. 553–567. doi: https://doi.org/10.1016/j.ccr.2010.11.015.

Foshay, K. M. and Gallicano, G. I. (2008) 'Regulation of Sox2 by STAT3 initiates commitment to the neural precursor cell fate.', *Stem cells and development*. United States, 17(2), pp. 269–278. doi: 10.1089/scd.2007.0098.

Friedmann-Morvinski, D. (2014) 'Glioblastoma heterogeneity and cancer cell plasticity.', *Critical reviews in oncogenesis*. United States, 19(5), pp. 327–336.

Fukushige, S. and Horii, A. (2013) 'DNA Methylation in Cancer : A Gene Silencing Mechanism and the Clinical Potential of Its Biomarkers', pp.

173–185. doi: 10.1620/tjem.229.173.Correspondence.

Garros-Regulez, L. *et al.* (2016) 'Targeting SOX2 as a Therapeutic Strategy in Glioblastoma', *Frontiers in Oncology*. Frontiers Media S.A., 6, p. 222. doi: 10.3389/fonc.2016.00222.

Gehring, M., Reik, W. and Henikoff, S. (2009) 'DNA demethylation by DNA repair', *Trends in Genetics*, 25(2), pp. 82–90. doi: http://dx.doi.org/10.1016/j.tig.2008.12.001.

Goodenberger, M. L. and Jenkins, R. B. (2012) 'Genetics of adult glioma', *Cancer Genetics*, 205(12), pp. 613–621. doi: http://doi.org/10.1016/j.cancergen.2012.10.009.

Gu, T.-P. *et al.* (2011) 'The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 477(7366), pp. 606–610.

Guo, L. *et al.* (2015) 'Roles of NOTCH1 as a Therapeutic Target and a Biomarker for Lung Cancer : Controversies and Perspectives', 2015.

Hackett, J. A. and Surani, M. A. (2012) 'DNA methylation dynamics during the mammalian life cycle', pp. 1–8.

Haffner, M. C. *et al.* (2011) 'Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers', *Oncotarget*, 2.

Han, H. *et al.* (2011) 'DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter', *Human Molecular Genetics*, pp. 4299–4310. doi: 10.1093/hmg/ddr356.

He, Y.-F. *et al.* (2011) 'Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA', *Science*, 333(6047), p. 1303 LP-1307.

Herceg, Z. and Ushijima, T. (2010) *Epigenetics and Cancer*. Elsevier Science (Advances in Genetics).

Holmfeldt, L. and Mullighan, C. G. (2011) 'The Role of TET2 in Hematologic Neoplasms', *Cancer Cell*, 20(1), pp. 1–2. doi: http://dx.doi.org/10.1016/j.ccr.2011.06.025.

Hughes, L. A. E. *et al.* (2013) 'The CpG Island Methylator Phenotype : What â€™ s in a Name ?', pp. 1–12. doi: 10.1158/0008-5472.CAN-12-4306.

Ignatova, T. N. *et al.* (2002) 'Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro.', *Glia*. United States, 39(3), pp. 193–206. doi: 10.1002/glia.10094.

Illingworth, R. S. and Bird, A. P. (2009) 'CpG islands – " A rough guide "', *FEBS Letters*. Federation of European Biochemical Societies, 583(11), pp. 1713–1720. doi: 10.1016/j.febslet.2009.04.012.

Iqbal, K. *et al.* (2011) 'Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 108(9), pp. 3642–3647. doi: 10.1073/pnas.1014033108.

Ishii, T. and Eto, K. (2014) 'Fetal stem cell transplantation: Past, present, and future', *World Journal of Stem Cells*. Baishideng Publishing Group Inc, 6(4), pp. 404–420. doi: 10.4252/wjsc.v6.i4.404.

Ito *et al.* (2010) 'Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification', *Nature*, 466. doi: 10.1038/nature09303.

Ito, S. *et al.* (2010) 'Role of Tet proteins in 5mC to 5hmC conversion, ES cell self-renewal, and ICM specification', *Nature*, 466(7310), pp. 1129–1133. doi: 10.1038/nature09303.

Ito, S. *et al.* (2011) 'Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine', *Science (New York, N.Y.)*, 333(6047), pp. 1300–1303. doi: 10.1126/science.1210597.

Iyer, L. M. *et al.* (2009) 'Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids', *Cell Cycle*. Taylor & Francis, 8(11), pp. 1698–1710. doi: 10.4161/cc.8.11.8580.

Jin, X. *et al.* (2013) 'Cell surface Nestin is a biomarker for glioma stem cells', *Biochemical and Biophysical Research Communications*, 433(4), pp. 496–501. doi: https://doi.org/10.1016/j.bbrc.2013.03.021.

Jones, P. A. (2012) 'Functions of DNA methylation: islands, start sites, gene bodies and beyond', *Nat Rev Genet*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 13(7), pp. 484–492.

Jones, P. A. and Liang, G. (2009) 'Rethinking how DNA methylation patterns are maintained', *Nat Rev Genet*. Nature Publishing Group, 10(11), pp. 805–811.

Kageyama, R. and Nakanishi, S. (1997) 'Helix-loop-helix factors in growth and differentiation of the vertebrate nervous system.', *Current opinion in genetics & development*. England, 7(5), pp. 659–

665.

Kalari, S. and Pfeifer, G. P. (2010) 'Identification of driver and passenger DNA methylation in cancer by epigenomic analysis', *Advances in genetics*, 70, pp. 277–308. doi: 10.1016/B978-0-12-380866-0.60010-1.

Kanamori, M. *et al.* (2007) 'Contribution of Notch signaling activation to human glioblastoma multiforme.', *Journal of neurosurgery*. United States, 106(3), pp. 417–427. doi: 10.3171/jns.2007.106.3.417.

Ko, M. *et al.* (2010) 'Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2', *Nature*, 468(7325), pp. 839–843. doi: 10.1038/nature09586.

Kofler, N. M. *et al.* (2011) 'Notch Signaling in Developmental and Tumor Angiogenesis', *Genes & Cancer*. Sage CA: Los Angeles, CA: SAGE Publications, 2(12), pp. 1106–1116. doi: 10.1177/1947601911423030.

Kondo, Y. *et al.* (2014) 'Epigenetic dysregulation in glioma', *Cancer Science*, 105(4), pp. 363–369. doi: 10.1111/cas.12379.

Kriaucionis, S. and Heintz, N. (2009) 'The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain', *Science*, 324(5929), p. 929 LP-930.

Kushwaha, G. *et al.* (2016) 'Hypomethylation coordinates antagonistically with hypermethylation in cancer development: a case study of leukemia', *Human Genomics*, 10(2), p. 18. doi: 10.1186/s40246-016-0071-5.

Ladd-acosta, C. and Fallin, M. D. (2016) 'The role of epigenetics in genetic and environmental epidemiology', 8, pp. 271–283.

Lake, R. J. *et al.* (2009) 'In vivo analysis of the Notch receptor S1 cleavage.', *PloS one*. United States, 4(8), p. e6728. doi: 10.1371/journal.pone.0006728.

Lapidot, T. *et al.* (1994) 'A cell initiating human acute myeloid leukaemia after transplantation into SCID mice', *Nature*, 367(6464), pp. 645–648.

Lasky, J. L. and Wu, H. (2005) 'Notch signaling, brain development, and human disease.', *Pediatric research*. United States, 57(5 Pt 2), p. 104R–109R. doi: 10.1203/01.PDR.0000159632.70510.3D.

Lathia, J. D. *et al.* (2015) 'Cancer stem cells in glioblastoma', *Genes & Development* , 29(12), pp. 1203–1217.

Lee, S.-K. *et al.* (2005) 'Olig2 and Ngn2 function in opposition to modulate gene expression in motor neuron progenitor cells.', *Genes & development*. United States, 19(2), pp. 282–294. doi: 10.1101/gad.1257105.

Leelatian, N. and Ihrie, R. A. (2016) 'Head of the Class: OLIG2 and Glioblastoma Phenotype', *Cancer Cell*. Elsevier, 29(5), pp. 613–615. doi: 10.1016/j.ccell.2016.04.007.

Lester, B. M., Conradt, E. and Marsit, C. (2016) 'Introduction to the Special Section on Epigenetics', *Child development*, pp. 29–37. doi: 10.1111/cdev.12489.

Ley, T. J. *et al.* (2010) 'DNMT3A Mutations in Acute Myeloid Leukemia', *New England Journal of Medicine*. Massachusetts Medical Society, 363(25), pp. 2424–2433. doi: 10.1056/NEJMoa1005143.

Li, H. *et al.* (2011) 'Phosphorylation regulates OLIG2 cofactor choice

and the motor neuron-oligodendrocyte fate switch.', *Neuron*. United States, 69(5), pp. 918–929. doi: 10.1016/j.neuron.2011.01.030.

Lian, C. G. *et al.* (2012) 'Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma', *Cell*, 150. doi: 10.1016/j.cell.2012.07.033.

Ligon, K. L. *et al.* (2007) 'Olig2-Regulated Lineage-Restricted Pathway Controls Replication Competence in Neural Stem Cells and Malignant Glioma', *Neuron*, 53(4), pp. 503–517. doi: https://doi.org/10.1016/j.neuron.2007.01.009.

Lindahl, T. and Wood, R. D. (1999) 'Quality Control by DNA Repair', *Science*, 286(5446), p. 1897 LP-1905.

Lobry, C., Oh, P. and Aifantis, I. (2011) 'Oncogenic and tumor suppressor functions of Notch in cancer: it's NOTCH what you think', *The Journal of Experimental Medicine*. The Rockefeller University Press, 208(10), pp. 1931–1935. doi: 10.1084/jem.20111855.

Lu, W. J. *et al.* (2011) 'Inducible expression of stem cell associated intermediate filament nestin reveals an important role in glioblastoma carcinogenesis', *International Journal of Cancer*. Wiley Subscription Services, Inc., A Wiley Company, 128(2), pp. 343–351. doi: 10.1002/ijc.25586.

Maiti, A. and Drohat, A. C. (2011) 'Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: POTENTIAL IMPLICATIONS FOR ACTIVE DEMETHYLATION OF CpG SITES ', *Journal of Biological Chemistry* , 286(41), pp. 35334–35338. doi: 10.1074/jbc.C111.284620.

Malatesta, P. *et al.* (2017) 'Neuronal or Glial Progeny', *Neuron*.

Elsevier, 37(5), pp. 751–764. doi: 10.1016/S0896-6273(03)00116-8.

Manoranjan, B. *et al.* (2012) 'Medulloblastoma stem cells: where development and cancer cross pathways', *Pediatr Res*. International Pediatric Research Foundation, Inc., 71(4–2), pp. 516–522. Available at: http://dx.doi.org/10.1038/pr.2011.62.

Marquardt, H. *et al.* (1983) 'Transforming growth factors produced by retrovirus-transformed rodent fibroblasts and human melanoma cells: amino acid sequence homology with epidermal growth factor.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 80(15), pp. 4684–4688.

Matsuda, Y. *et al.* (2013) 'Nestin: Neural Stem/Progenitor Cell Marker in Brain Tumors', in Lichtor, T. B. T.-E. of the M. B. of B. T. and the T. I. (ed.). Rijeka: InTech, p. Ch. 23. doi: 10.5772/52634.

Matsuda, Y., Hagio, M. and Ishiwata, T. (2013) 'Nestin: A novel angiogenesis marker and possible target for tumor angiogenesis', *World Journal of Gastroenterology : WJG*. Baishideng Publishing Group Co., Limited, 19(1), pp. 42–48. doi: 10.3748/wjg.v19.i1.42.

McNamara, M. G., Sahebjam, S. and Mason, W. P. (2013) 'Emerging Biomarkers in Glioblastoma', *Cancers*. MDPI, 5(3), pp. 1103–1119. doi: 10.3390/cancers5031103.

Meijer, D. H. *et al.* (2012) 'Separated at birth? The functional and molecular divergence of OLIG1 and OLIG2', *Nature reviews. Neuroscience*, 13(12), pp. 819–831. doi: 10.1038/nrn3386.

Mendez-Ferrer, S. *et al.* (2010) 'Mesenchymal and haematopoietic stem cells form a unique bone marrow niche', *Nature*. Nature Publishing Group, a division of Macmillan Publishers Limited. All

Rights Reserved., 466(7308), pp. 829–834.

Mendonca, A. *et al.* (2014) 'Hydroxymethylation of DNA influences nucleosomal conformation and stability in vitro', *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1839(11), pp. 1323–1329. doi: https://doi.org/10.1016/j.bbagrm.2014.09.014.

Meng, H. *et al.* (2015) 'DNA Methylation, Its Mediators and Genome Integrity', *International Journal of Biological Sciences*. Sydney: Ivyspring International Publisher, 11(5), pp. 604–617. doi: 10.7150/ijbs.11218.

Moore, L. D., Le, T. and Fan, G. (2013) 'DNA Methylation and Its Basic Function', *Neuropsychopharmacology*, pp. 23–38. doi: 10.1038/npp.2012.112.

Morgan, T. H. and Bridges, C. B. (1916) *Sex-linked inheritance in Drosophila,* . Washington: Carnegie Institution of Washington,.

Munoz, P., Iliou, M. S. and Esteller, M. (2012) 'Epigenetic alterations involved in cancer stem cell reprogramming.', *Molecular oncology*. United States, 6(6), pp. 620–636. doi: 10.1016/j.molonc.2012.10.006.

Nagane, M. *et al.* (1996) 'A common mutant epidermal growth factor receptor confers enhanced tumorigenicity on human glioblastoma cells by increasing proliferation and reducing apoptosis.', *Cancer research*. United States, 56(21), pp. 5079–5086.

Nagarajan, R. P. *et al.* (2014) 'Recurrent epimutations activate gene body promoters in primary glioblastoma', *Genome Research*. Cold Spring Harbor Laboratory Press, 24(5), pp. 761–774. doi: 10.1101/gr.164707.113.

Najbauer, J. *et al.* (2012) 'Cellular Host Responses to Gliomas', *PLoS ONE*. Edited by M. P. Bachmann. San Francisco, USA: Public Library of Science, 7(4), p. e35150. doi: 10.1371/journal.pone.0035150.

Natsume, A. *et al.* (2010) 'Epigenetic aberrations and therapeutic implications in gliomas', *Cancer Science*. Blackwell Publishing Ltd, 101(6), pp. 1331–1336. doi: 10.1111/j.1349-7006.2010.01545.x.

Niehrs, C. (2009) 'Active {DNA} demethylation and {DNA} repair', *Differentiation*, 77(1), pp. 1–11. doi: http://dx.doi.org/10.1016/j.diff.2008.09.004.

Nishikawa, R. *et al.* (1994) 'A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 91(16), pp. 7727–7731.

Normanno, N. *et al.* (2006) 'Epidermal growth factor receptor (EGFR) signaling in cancer.', *Gene*. Netherlands, 366(1), pp. 2–16. doi: 10.1016/j.gene.2005.10.018.

Ohgaki, H. and Kleihues, P. (2007) 'Genetic Pathways to Primary and Secondary Glioblastoma', *The American Journal of Pathology*. American Society for Investigative Pathology, 170(5), pp. 1445–1453. doi: 10.2353/ajpath.2007.070011.

Ohgaki, H. and Kleihues, P. (2013) 'The De fi nition of Primary and Secondary Glioblastoma', 19(4), pp. 764–773. doi: 10.1158/1078-0432.CCR-12-3002.

Okano, M. *et al.* (1999) 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development', 99, pp. 247–257.

Okano, M. *et al.* (2017) 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development', *Cell*. Elsevier, 99(3), pp. 247–257. doi: 10.1016/S0092-8674(00)81656-6.

Ono, R. *et al.* (2002) 'LCX, Leukemia-associated Protein with a CXXC Domain, Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Having t(10;11)(q22;q23)', *Cancer Research*, 62(14), p. 4075 LP-4080.

Parsons, D. W. *et al.* (2008) 'An Integrated Genomic Analysis of Human Glioblastoma Multiforme', *Science (New York, N.Y.)*, 321(5897), p. 1807. doi: 10.1126/science.1164382.

Penn, N. W. *et al.* (1972) 'The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid', *Biochemical Journal*, 126(4), pp. 781–790.

Pfaffeneder, T. *et al.* (2011) 'The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA', *Angewandte Chemie International Edition*. WILEY-VCH Verlag, 50(31), pp. 7008–7012. doi: 10.1002/anie.201103899.

Phillips, T. (2008) 'The role of methylation in gene expression', *Nature Education*, 1(1), p. 116.

Purow, B. W. *et al.* (2005) 'Expression of Notch-1 and its ligands, Delta-like-1 and Jagged-1, is critical for glioma cell survival and proliferation.', *Cancer research*. United States, 65(6), pp. 2353–2363. doi: 10.1158/0008-5472.CAN-04-1890.

Raiber, E.-A. *et al.* (2012) 'Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase', *Genome Biology*, 13(8), p. R69. doi: 10.1186/gb-2012-13-8-r69.

Ramsawhook, A. *et al.* (2017) 'Medulloblastoma and ependymoma cells display increased levels of 5-carboxylcytosine and elevated TET1 expression', *Clinical Epigenetics*. London: BioMed Central, 9, p. 18. doi: 10.1186/s13148-016-0306-2.

Ramsawhook, H. A. and Lewis, C. L. (2017) 'Immunostaining of DNA modifications: computational analysis of confocal images', *JoVe*.

Rasmussen, K. D. and Helin, K. (2016) 'Role of TET enzymes in DNA methylation, development, and cancer', *Genes & Development* , 30(7), pp. 733–750. doi: 10.1101/gad.276568.115.

Robertson, K. D. *et al.* (1999) 'The human DNA methyltransferases (DNMTs) 1 , 3a and 3b : coordinate mRNA expression in normal tissues and overexpression in tumors', 27(11), pp. 2291–2298.

Robertson, K. D. (2005) 'DNA methylation and human disease', *Nat Rev Genet*, 6(8), pp. 597–610.

Rodriguez-Pinilla, S. M. *et al.* (2007) 'Sox2: a possible driver of the basal-like phenotype in sporadic breast cancer.', *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*. United States, 20(4), pp. 474–481. doi: 10.1038/modpathol.3800760.

Rosen, J. M. and Jordan, C. T. (2009) 'The Increasing Complexity of the Cancer Stem Cell Paradigm', *Science (New York, N.Y.)*, 324(5935), pp. 1670–1673. doi: 10.1126/science.1171837.

Roszkowski, K. *et al.* (no date) 'Potential Role of Methylation Marker in Glioma Supporting Clinical Decisions', pp. 1–7. doi: 10.3390/ijms17111876.

Rutka, J. T. (1997) 'Medulloblastoma.', *Clinical neurosurgery*. United States, 44, pp. 571–585.

Sadakierska-Chudy, A., Kostrzewa, R. M. and Filip, M. (2015) 'A Comprehensive View of the Epigenetic Landscape Part I: DNA Methylation, Passive and Active DNA Demethylation Pathways and Histone Variants', *Neurotoxicity Research*. Boston: Springer US, 27, pp. 84–97. doi: 10.1007/s12640-014-9497-5.

Santos, F. and Dean, W. (2004) 'Epigenetic reprogramming during early development in mammals', *Reproduction* , 127(6), pp. 643–651. doi: 10.1530/rep.1.00221.

Sarkar, A. and Hochedlinger, K. (2013) 'The Sox Family of Transcription Factors: Versatile Regulators of Stem and Progenitor Cell Fate', *Cell stem cell*, 12(1), pp. 15–30. doi: 10.1016/j.stem.2012.12.007.

Schonberg, D. L. *et al.* (2014) 'Brain tumor stem cells: molecular characteristics and their impact on therapy', *Molecular aspects of medicine*, 0, pp. 82–101. doi: 10.1016/j.mam.2013.06.004.

Schüller, U. *et al.* (2008) 'Acquisition of granule neuron precursor identity is a critical determinant of progenitor cell competence to form Hedgehog-induced medulloblastoma', *Cancer cell*, 14(2), pp. 123–134. doi: 10.1016/j.ccr.2008.07.005.

Schwartzbaum, J. A. *et al.* (2006) 'Epidemiology and molecular pathology of glioma', 2(9), pp. 494–503. doi: 10.1038/ncpneuro0289.

Sedgwick, P. (2012) 'Pearson's correlation coefficient', 4483(July), pp. 1–2. doi: 10.1136/bmj.e4483.

Shen, L. *et al.* (2013) 'Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics', *Cell*, 153(3), pp. 692–706. doi: https://doi.org/10.1016/j.cell.2013.04.002.

Shinojima, N. *et al.* (2003) 'Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme.', *Cancer research*. United States, 63(20), pp. 6962–6970.

Shukla, G. *et al.* (2017) 'Therapeutic Potential, Challenges and Future Perspective of Cancer Stem Cells in Translational Oncology: A Critical Review', *Current Stem Cell Research & Therapy*, pp. 207–224. doi: http://dx.doi.org/10.2174/1574888X11666161028143224.

Sibin, M. K. *et al.* (2015) 'CD133 and BMI1 expressions and its prognostic role in primary glioblastoma', 94(4), pp. 689–696.

Simon, J. A. and Lange, C. A. (2008) 'Roles of the EZH2 histone methyltransferase in cancer epigenetics', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 647(1–2), pp. 21–29. doi: https://doi.org/10.1016/j.mrfmmm.2008.07.010.

Singh, S. K. *et al.* (2004) 'Identification of human brain tumour initiating cells', *Nature*, 432(7015), pp. 396–401.

Skalnik, D. G. (2001) 'CpG Binding Protein Is Crucial for Early Embryonic Development', 21(22), pp. 7601–7606. doi: 10.1128/MCB.21.22.7601.

Song, C.-X. and He, C. (2013) 'Potential functional roles of DNA

demethylation intermediates', *Trends in biochemical sciences*, 38(10), pp. 480–484. doi: 10.1016/j.tibs.2013.07.003.

Stockhausen, M.-T., Kristoffersen, K. and Poulsen, H. S. (2010) 'The functional role of Notch signaling in human gliomas', *Neuro-Oncology*. Oxford University Press, 12(2), pp. 199–211. doi: 10.1093/neuonc/nop022.

Stroud, H. *et al.* (2011) '5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells', *Genome Biology*. BioMed Central, 12(6), pp. R54–R54. doi: 10.1186/gb-2011-12-6-r54.

Szulwach, K. and Jin, P. (2014) *Integrating DNA methylation dynamics into a framework for understanding epigenetic codes*, *BioEssays : news and reviews in molecular, cellular and developmental biology*. doi: 10.1002/bies.201300090.

Szyf, M., Pakneshan, P. and Rabbani, S. A. (2004) 'DNA demethylation and cancer: therapeutic implications', *Cancer Letters*, 211(2), pp. 133–143. doi: http://doi.org/10.1016/j.canlet.2004.04.009.

Tabu, K. *et al.* (2008) 'Promoter hypomethylation regulates CD133 expression in human gliomas', pp. 1037–1046. doi: 10.1038/cr.2008.270.

Tahiliani, M. *et al.* (2009) 'Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1', *Science (New York, N.Y.)*, 324(5929), pp. 930–935. doi: 10.1126/science.1170116.

Takahashi, K. and Yamanaka, S. (2006) 'Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined

factors.', *Cell*. United States, 126(4), pp. 663–676. doi: 10.1016/j.cell.2006.07.024.

Takai, H. *et al.* (2014) '5-Hydroxymethylcytosine Plays a Critical Role in Glioblastomagenesis by Recruiting the CHTOP-Methylosome Complex', *Cell Reports*. Elsevier, 9(1), pp. 48–60. doi: 10.1016/j.celrep.2014.08.071.

Takebayashi, H. *et al.* (2000) 'Dynamic expression of basic helix-loop-helix Olig family members: implication of Olig2 in neuron and oligodendrocyte differentiation and identification of a new member, Olig3', *Mechanisms of Development*, 99(1), pp. 143–148. doi: http://dx.doi.org/10.1016/S0925-4773(00)00466-4.

Taylor, T. E., Furnari, F. B. and Cavenee, W. K. (2012) 'Targeting EGFR for Treatment of Glioblastoma: Molecular Basis to Overcome Resistance', *Current cancer drug targets*, 12(3), pp. 197–209.

Teodorczyk, M. and Schmidt, M. H. H. (2015) 'Notching on Cancer's Door: Notch Signaling in Brain Tumors', *Frontiers in Oncology*, 4, p. 341. doi: 10.3389/fonc.2014.00341.

Teperek-Tkacz, M. *et al.* (2011) 'Epigenetic reprogramming: is deamination key to active DNA demethylation?', *Reproduction* , 142(5), pp. 621–632. doi: 10.1530/REP-11-0148.

Thakkar, J. P. *et al.* (2014) 'Epidemiologic and Molecular Prognostic Review of Glioblastoma', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 23(10), pp. 1985–1996. doi: 10.1158/1055-9965.EPI-14-0275.

Thompson, B. J. *et al.* (2007) 'The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia.', *The Journal of experimental medicine*. United States, 204(8), pp. 1825–1835. doi: 10.1084/jem.20070872.

Thon, N., Kreth, S. and Kreth, F.-W. (2013) 'Personalized treatment strategies in glioblastoma: MGMT promoter methylation status', *OncoTargets and therapy*. Dove Medical Press, 6, pp. 1363–1372. doi: 10.2147/OTT.S50208.

Tortora, G. *et al.* (2007) 'Overcoming resistance to molecularly targeted anticancer therapies: Rational drug combinations based on EGFR and MAPK inhibition for solid tumours and haematologic malignancies.', *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*. Scotland, 10(3), pp. 81–100. doi: 10.1016/j.drup.2007.03.003.

Trépant, A.-L. *et al.* (2015) 'Identification of OLIG2 as the most specific glioblastoma stem cell marker starting from comparative analysis of data from similar DNA chip microarray platforms', *Tumor Biology*, 36(3), pp. 1943–1953. doi: 10.1007/s13277-014-2800-5.

Tronick, E. and Hunter, R. G. (2016) 'Waddington , Dynamic Systems , and Epigenetics', 10(June), pp. 1–6. doi: 10.3389/fnbeh.2016.00107.

Utt, C. A. L. N. *et al.* (2004) 'The Oligodendroglial Lineage Marker OLIG2 Is Universally Expressed in Diffuse Gliomas', 63(5), pp. 499–509.

Verhaak, R. G. W. *et al.* (2010) 'Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1.', *Cancer cell*. United States, 17(1), pp. 98–110. doi: 10.1016/j.ccr.2009.12.020.

Vinson, C. and Chatterjee, R. (2012) 'CG methylation', *Epigenomics*, 4(6), pp. 655–663. doi: 10.2217/epi.12.55.

Voldborg, B. R. *et al.* (1997) 'Epidermal growth factor receptor (EGFR) and EGFR mutations, function and possible role in clinical trials', *Annals of Oncology*, 8(12), pp. 1197–1206. doi: 10.1023/A:1008209720526.

Waddington, C. H. (1942) 'The epigenotype', *Endeavour 1*, pp. 18–20.

Wajed, S. A., Laird, P. W. and DeMeester, T. R. (2001) 'DNA Methylation: An Alternative Pathway to Cancer', *Annals of Surgery*, 234(1), pp. 10–20.

Wang, J. *et al.* (2008) 'CD133 negative glioma cells form tumors in nude rats and give rise to CD133 positive cells', *International Journal of Cancer*. Wiley Subscription Services, Inc., A Wiley Company, 122(4), pp. 761–768. doi: 10.1002/ijc.23130.

Watanabe, K. *et al.* (1996) 'Overexpression of the EGF receptor and p53 mutations are mutually exclusive in the evolution of primary and secondary glioblastomas.', *Brain pathology (Zurich, Switzerland)*. Switzerland, 6(3), pp. 214–217.

Weinhold, B. (2006) 'Epigenetics: The Science of Change', *Environmental Health Perspectives*, pp. A160-7.

Wen, L. and Tang, F. (2014) 'Genomic distribution and possible functions of DNA hydroxymethylation in the brain', *Genomics*, 104(5), pp. 341–346. doi: https://doi.org/10.1016/j.ygeno.2014.08.020.

Weng, A. P. *et al.* (2004) 'Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia.', *Science (New York, N.Y.)*. United

States, 306(5694), pp. 269–271. doi: 10.1126/science.1102160.

Wharton, K. A. *et al.* (1985) 'Nucleotide sequence from the neurogenic locus Notch implies a gene product that shares homology with proteins containing EGF-like repeats', *Cell*, 43(3, Part 2), pp. 567–581. doi: https://doi.org/10.1016/0092-8674(85)90229-6.

Wheldon, L. M. *et al.* (2014) 'Transient accumulation of 5-carboxylcytosine indicates involvement of active demethylation in lineage specification of neural stem cells', *Cell Rep*, 7. doi: 10.1016/j.celrep.2014.05.003.

Wieduwilt, M. J. and Moasser, M. M. (2008) 'The epidermal growth factor receptor family: Biology driving targeted therapeutics', *Cellular and molecular life sciences : CMLS*, 65(10), pp. 1566–1584. doi: 10.1007/s00018-008-7440-8.

Wiese, C. *et al.* (2004) 'Nestin expression--a property of multi-lineage progenitor cells?', *Cellular and molecular life sciences : CMLS*. Switzerland, 61(19–20), pp. 2510–2522. doi: 10.1007/s00018-004-4144-6.

Wiley, H. S., Shvartsman, S. Y. and Lauffenburger, D. A. (2003) 'Computational modeling of the EGF-receptor system: a paradigm for systems biology.', *Trends in cell biology*. England, 13(1), pp. 43–50.

Wong, A. J. *et al.* (1992) 'Structural alterations of the epidermal growth factor receptor gene in human gliomas', *Proceedings of the National Academy of Sciences* , 89(7), pp. 2965–2969.

Wu, H. *et al.* (2011) 'Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells', *Nature*, 473. doi: 10.1038/nature09934.

Wu, H. and Zhang, Y. (2014) 'Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions', *Cell*, 156(0), pp. 45–68. doi: 10.1016/j.cell.2013.12.019.

Xu, Y. *et al.* (2011) 'Genome-wide Regulation of 5hmC, 5mC and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells', *Molecular cell*, 42(4), pp. 451–464. doi: 10.1016/j.molcel.2011.04.005.

Xu, Y. *et al.* (2012) 'Tet3 CXXC Domain and Dioxygenase Activity Cooperatively Regulate Key Genes for Xenopus Eye and Neural Development', *Cell*, 151(6), pp. 1200–1213. doi: https://doi.org/10.1016/j.cell.2012.11.014.

Yamamoto, S., Schulze, K. L. and Bellen, H. J. (2014) 'Introduction to Notch Signaling BT - Notch Signaling: Methods and Protocols', in Bellen, H. J. and Yamamoto, S. (eds). New York, NY: Springer New York, pp. 1–14. doi: 10.1007/978-1-4939-1139-4_1.

Yu, Z. *et al.* (2012) 'Cancer stem cells', *The International Journal of Biochemistry & Cell Biology*, 44(12), pp. 2144–2151. doi: https://doi.org/10.1016/j.biocel.2012.08.022.

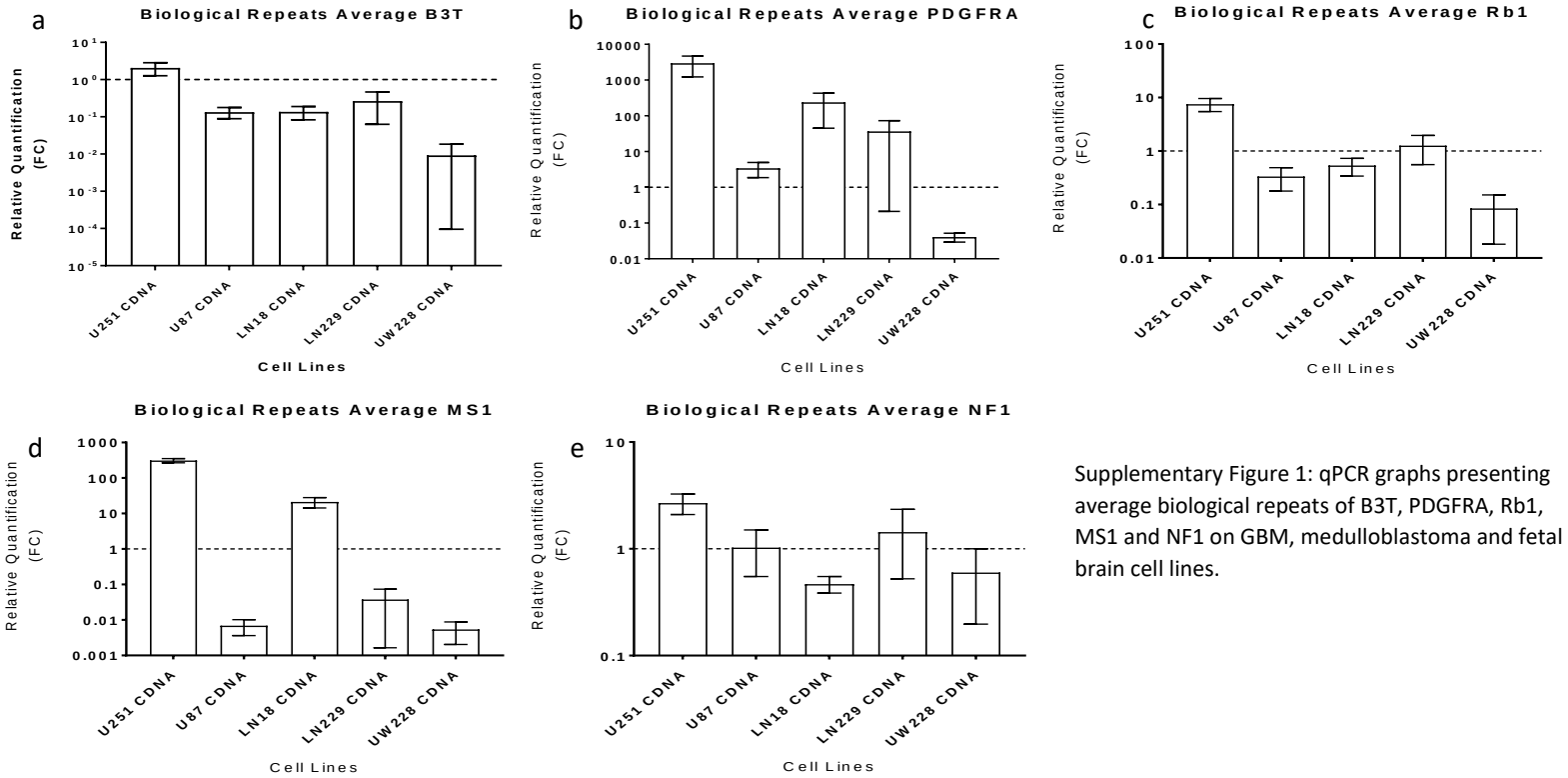
Zhang, C. *et al.* (2013) 'IDH1/2 mutations target a key hallmark of cancer by deregulating cellular metabolism in glioma', *Neuro-Oncology*. Oxford University Press, 15(9), pp. 1114–1126. doi: 10.1093/neuonc/not087.

Zhang, X. *et al.* (2012) 'Glioblastoma multiforme : Molecular characterization and current treatment strategy (Review)', (127), pp. 9–14. doi: 10.3892/etm.2011.367.

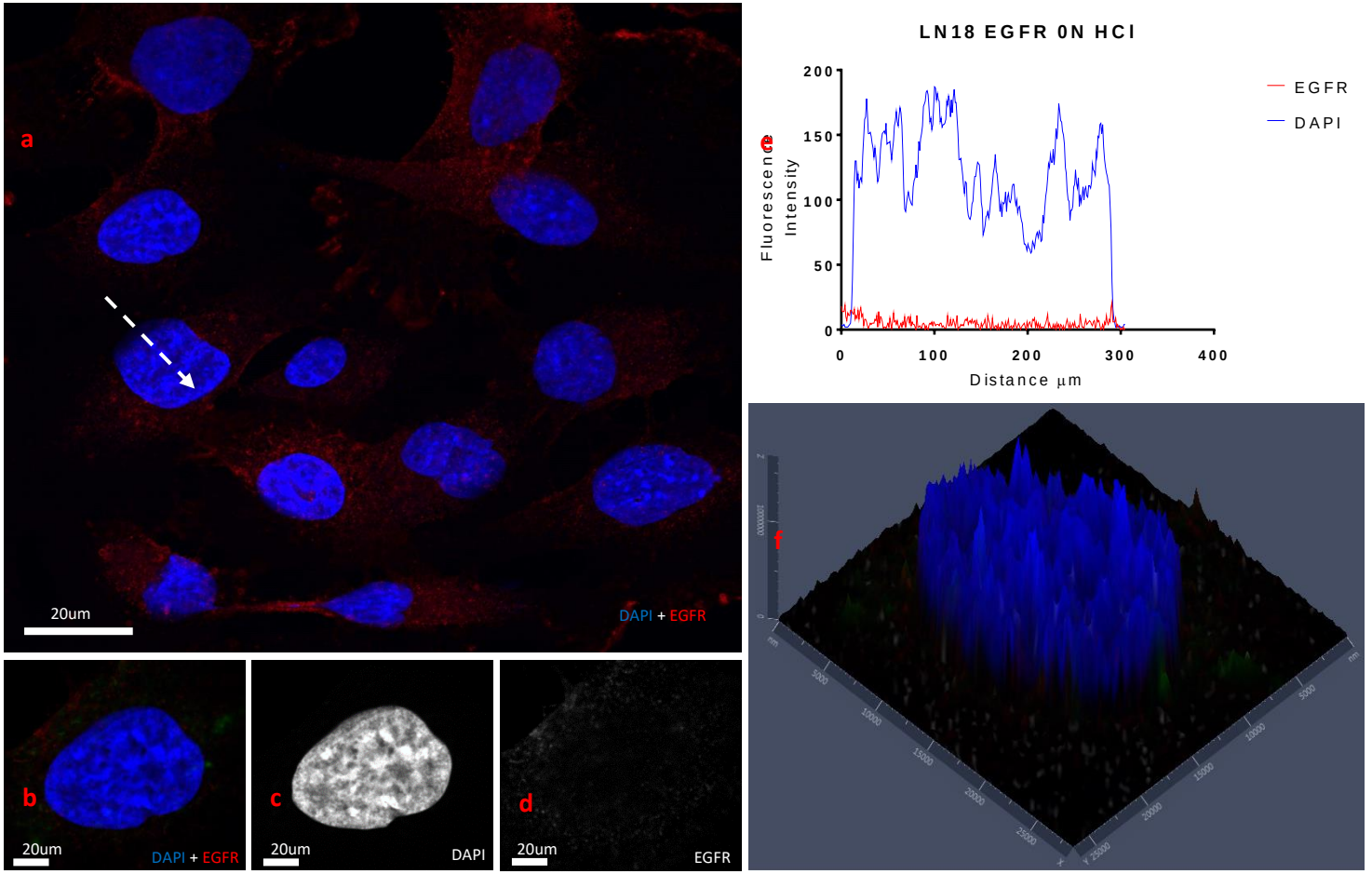
Zheng, S. *et al.* (2011) 'DNA hypermethylation profiles associated with glioma subtypes and EZH2 and IGFBP2 mRNA expression', *Neuro-*

Oncology, 13(3), pp. 280–289.

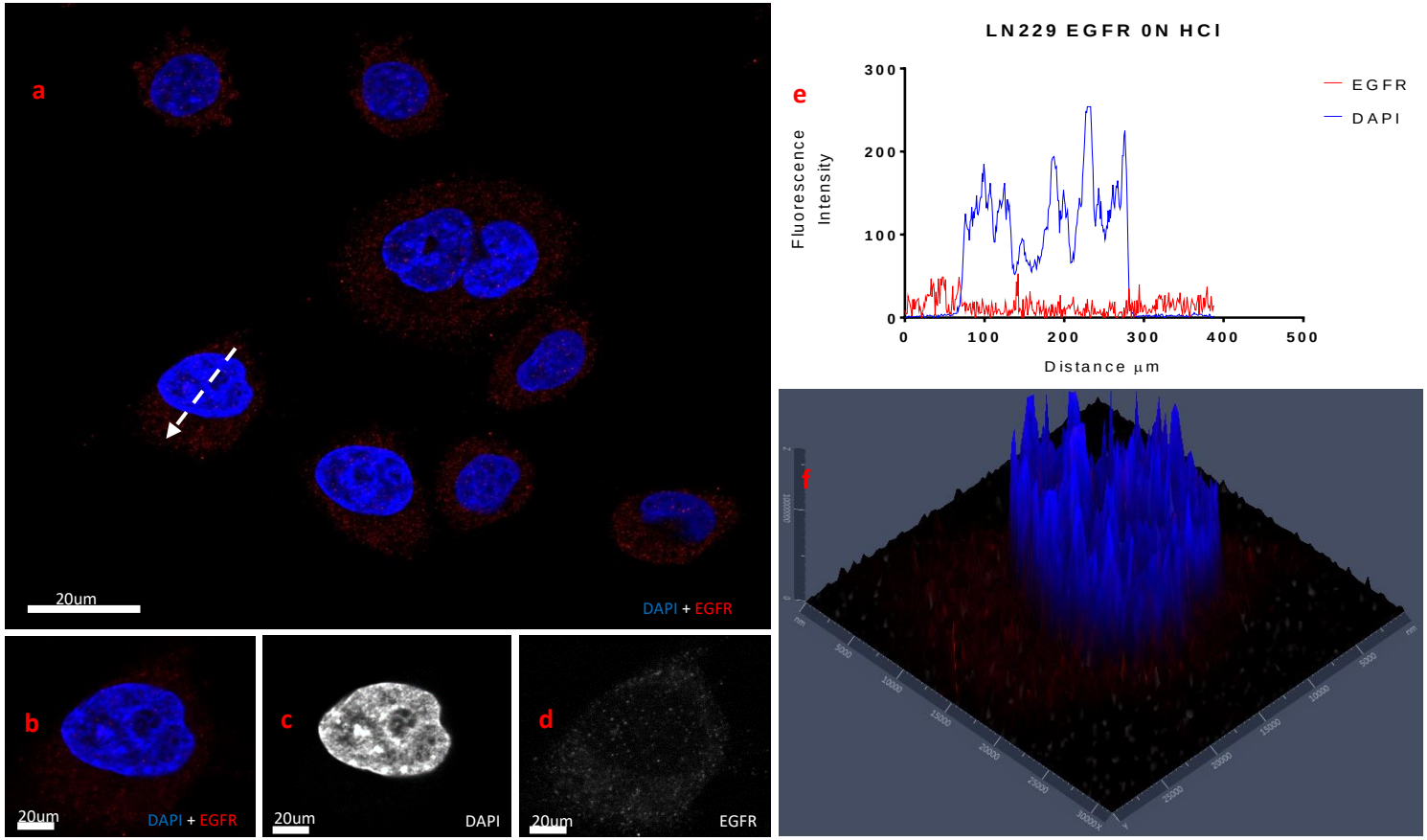
9. SUPPLEMENTARY INFORMATION



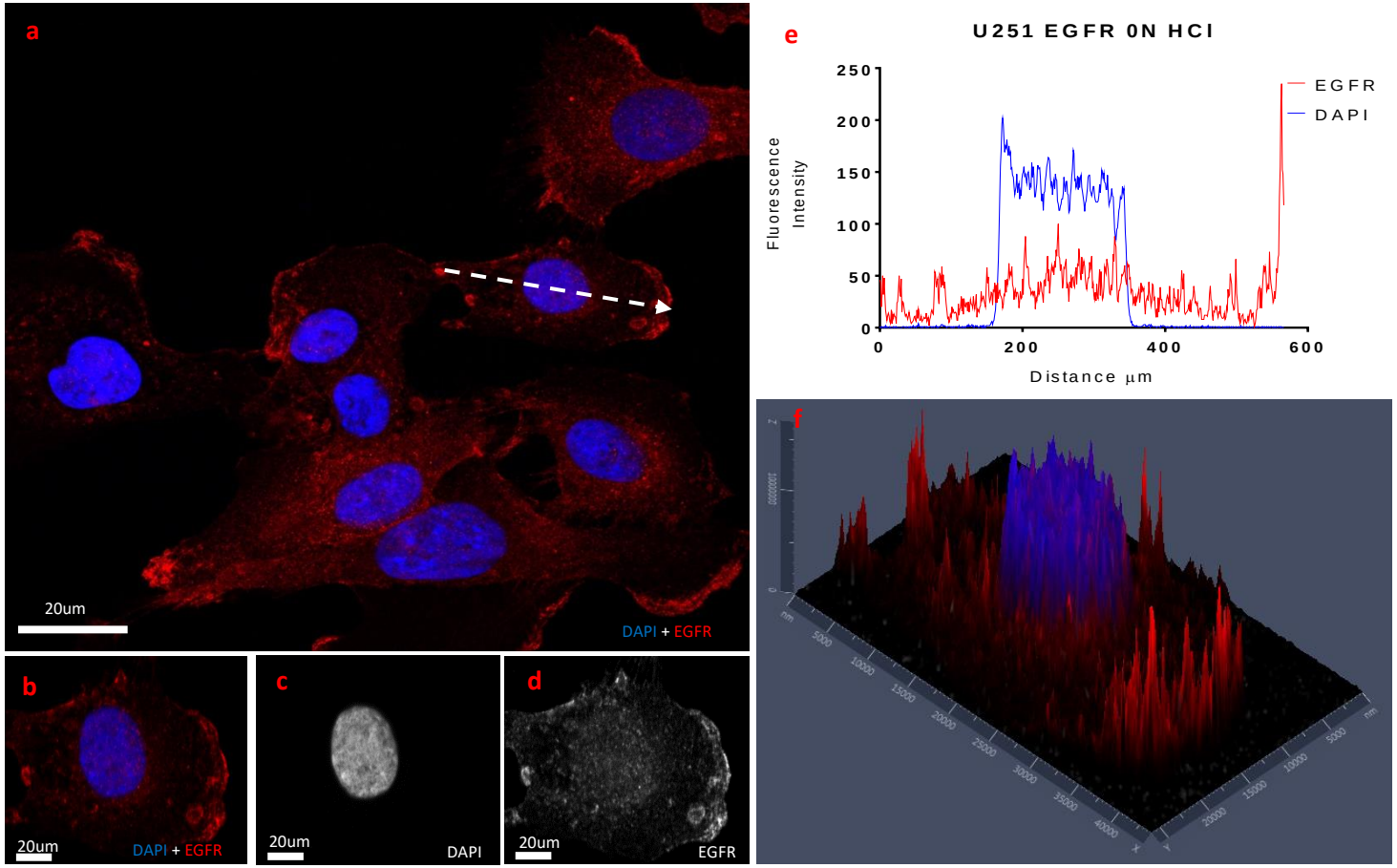
Supplementary Figure 1: qPCR graphs presenting average biological repeats of B3T, PDGFRA, Rb1, MS1 and NF1 on GBM, medulloblastoma and fetal brain cell lines.



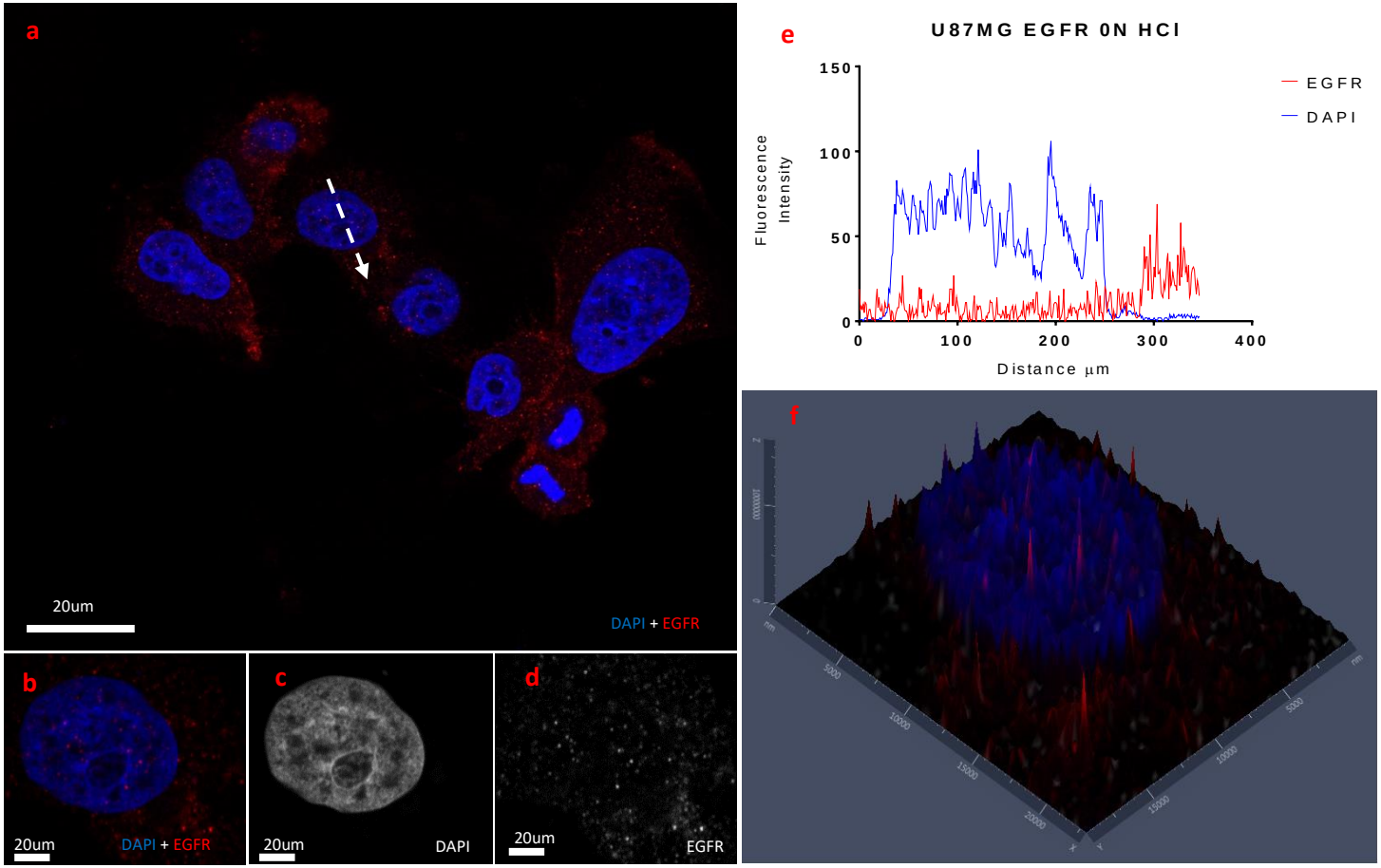
Supplementary Figure 2: **a-d** Immunostaining images from confocal microscope representing EGFR in LN18 with no treatment of HCl, includes merge of DAPI and EGFR (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and EGFR (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



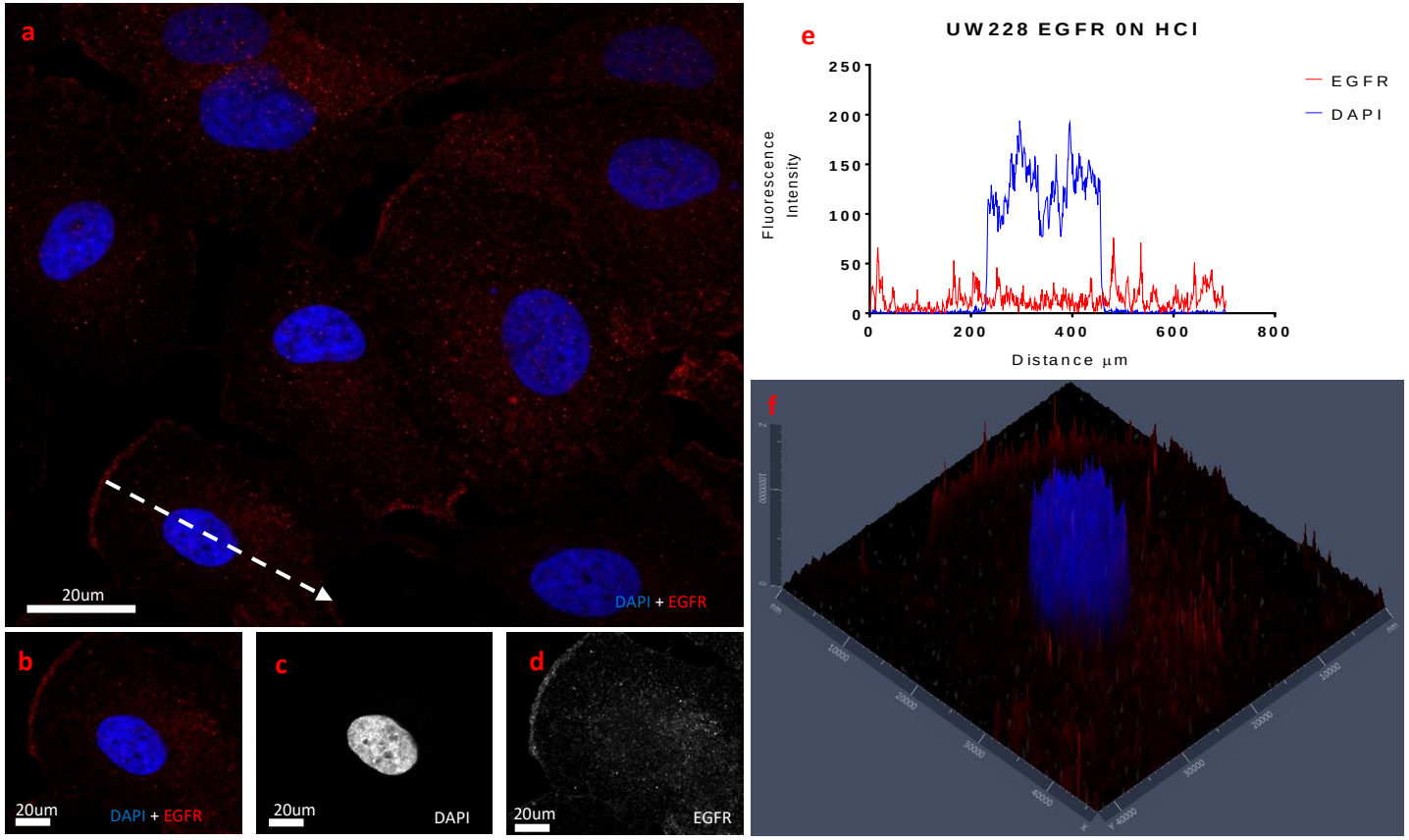
Supplementary Figure 3: **a-d** Immunostaining images from confocal microscope representing EGFR in LN229 with no treatment of HCl, includes merge of DAPI and EGFR (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and EGFR (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



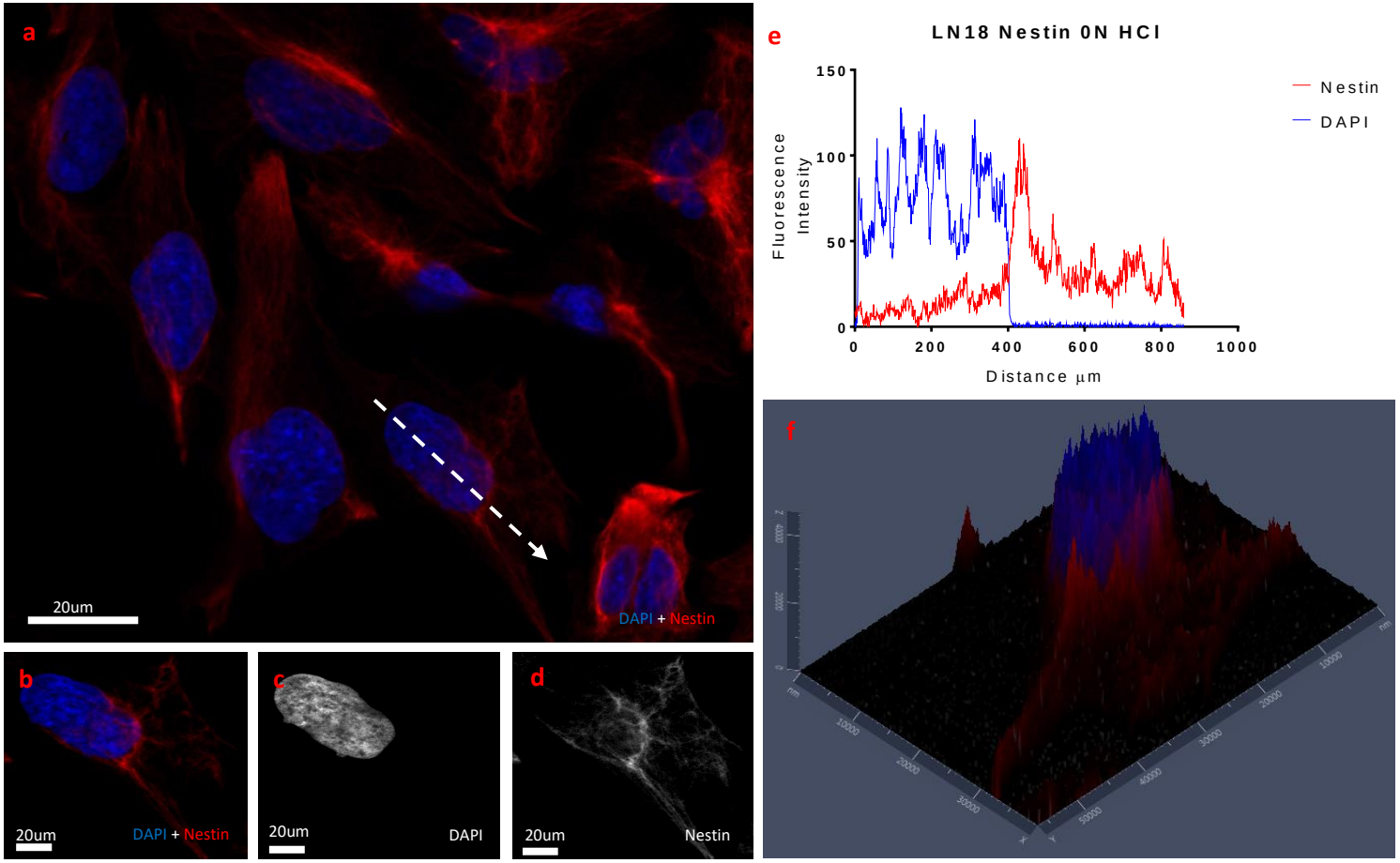
Supplementary Figure 4: **a-d** Immunostaining images from confocal microscope representing EGFR in U251 with no treatment of HCl, includes merge of DAPI and EGFR (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and EGFR (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



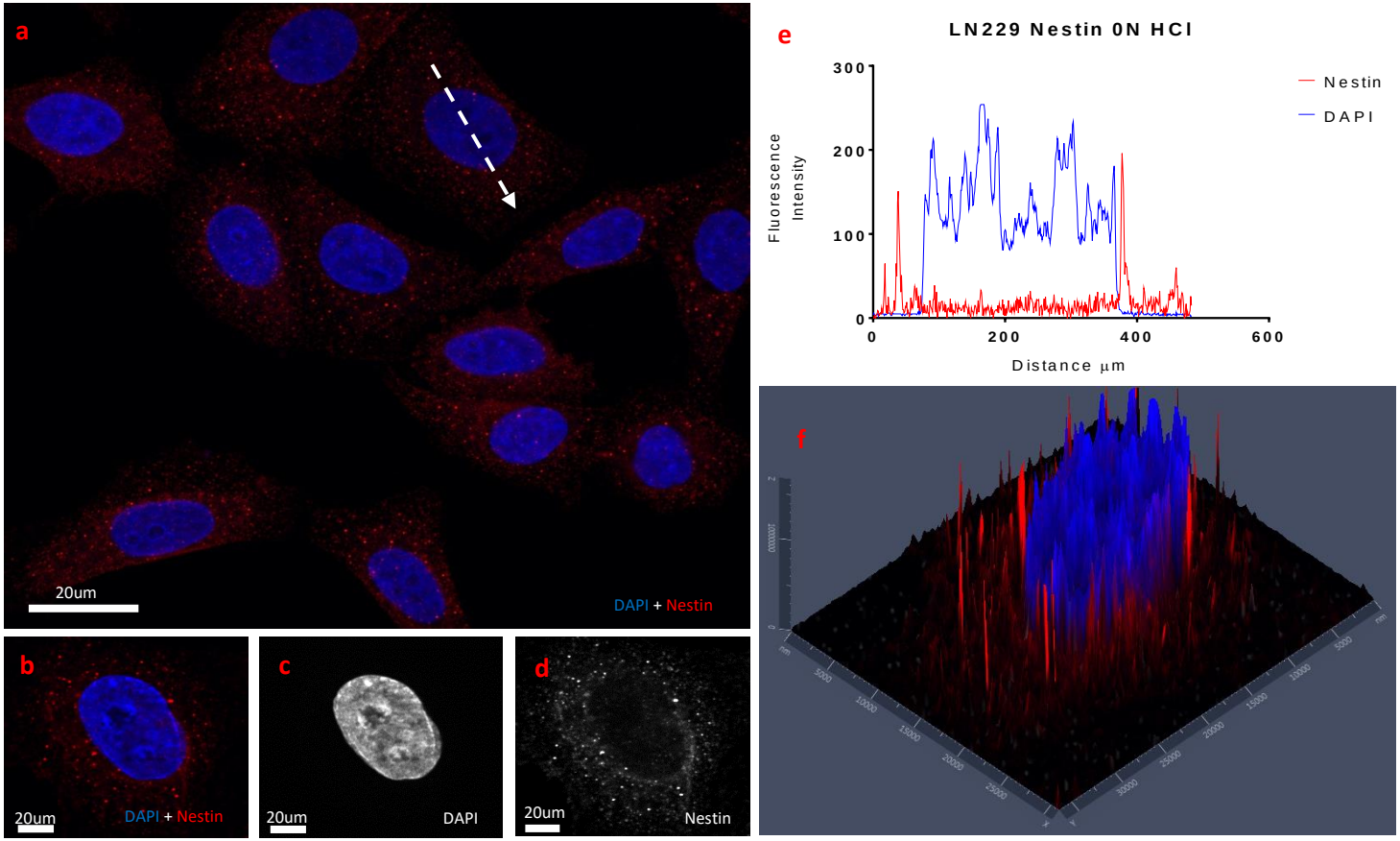
Supplementary Figure 5: **a-d** Immunostaining images from confocal microscope representing EGFR in U87MG with no treatment of HCl, includes merge of DAPI and EGFR (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and EGFR (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



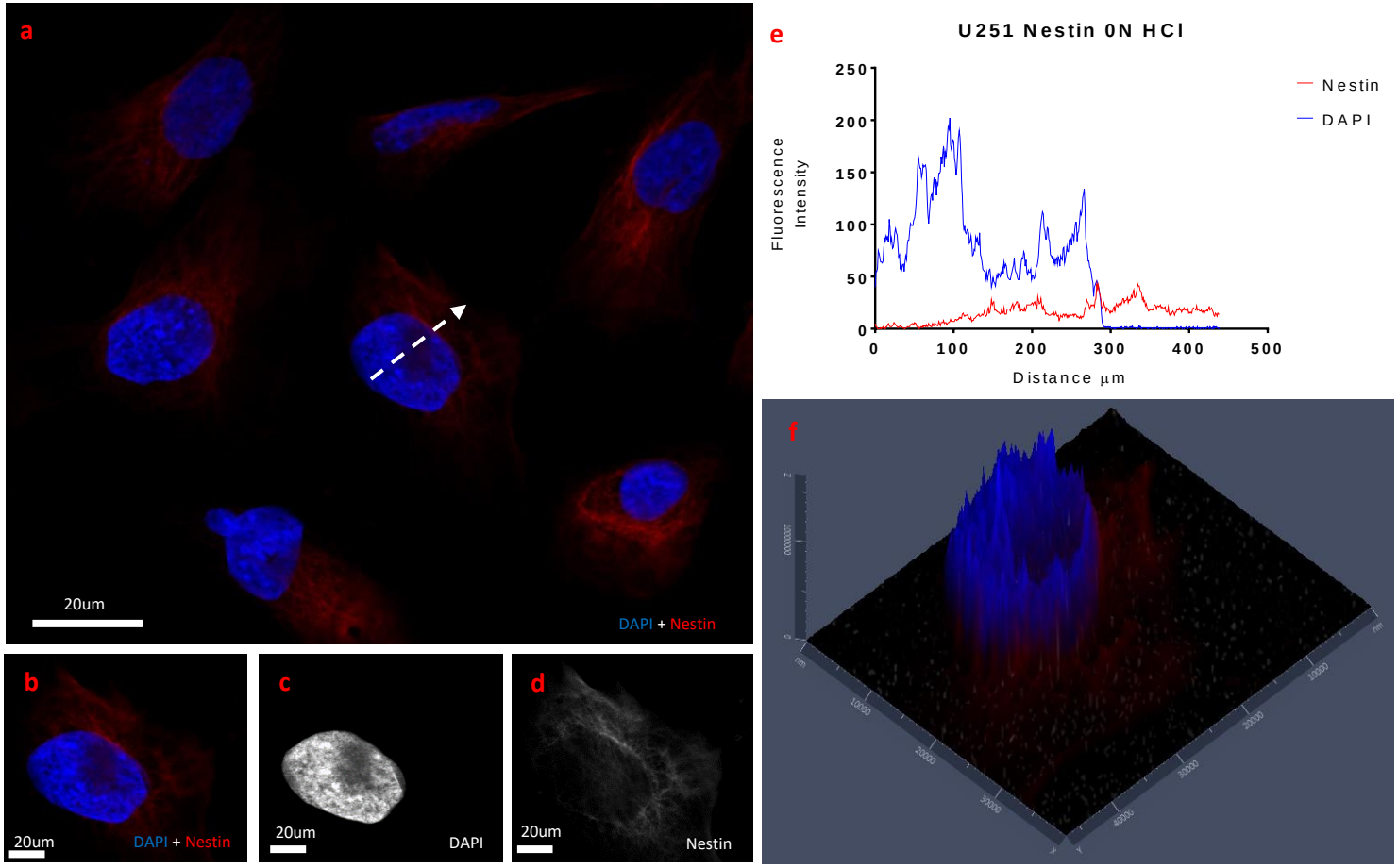
Supplementary Figure 6: **a-d** Immunostaining images from confocal microscope representing EGFR in UW228-3 with no treatment of HCl, includes merge of DAPI and EGFR (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and EGFR (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



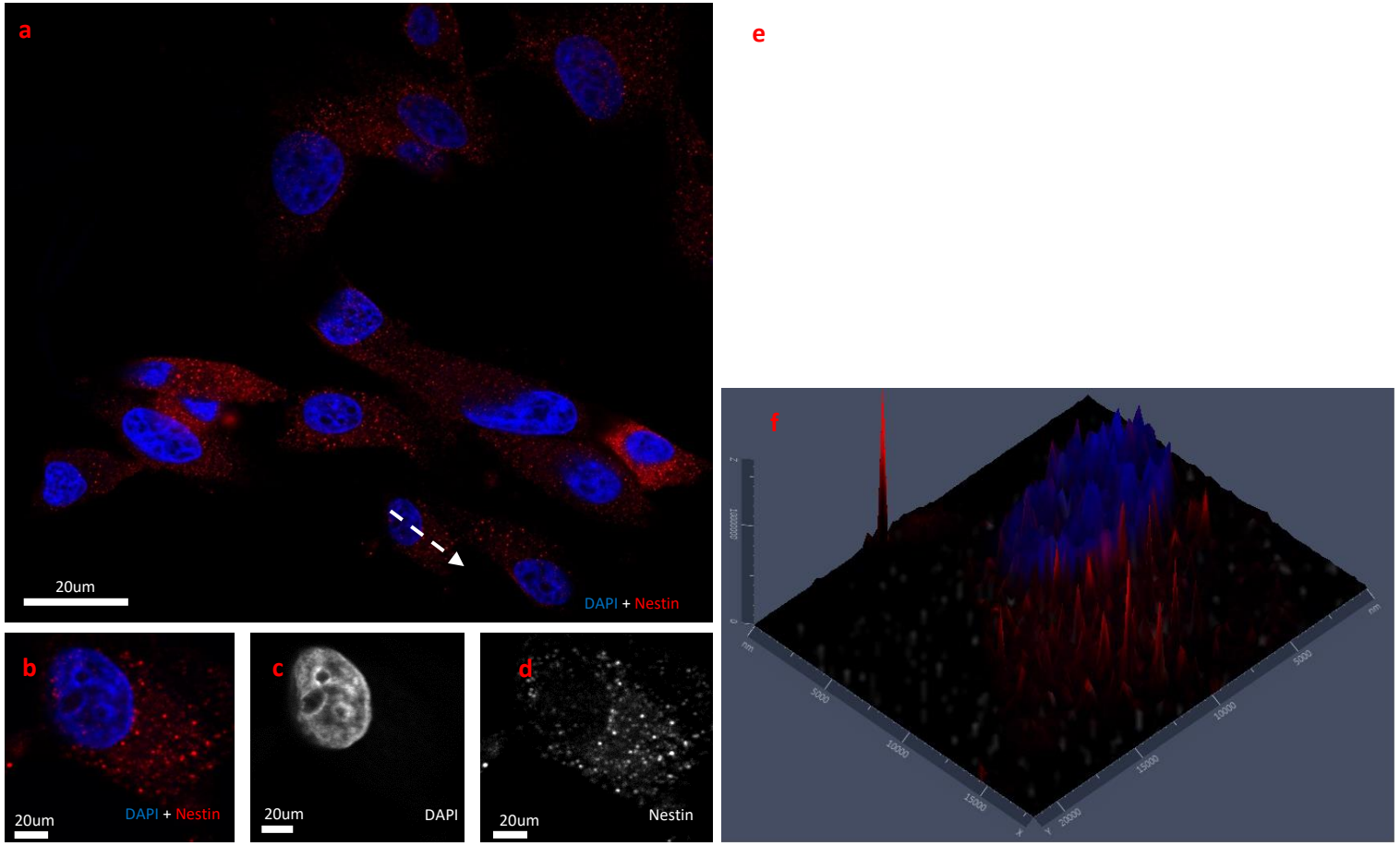
Supplementary Figure 7: **a-d** Immunostaining images from confocal microscope representing Nestin in LN18 with no treatment of HCl, includes merge of DAPI and Nestin (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Nestin (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



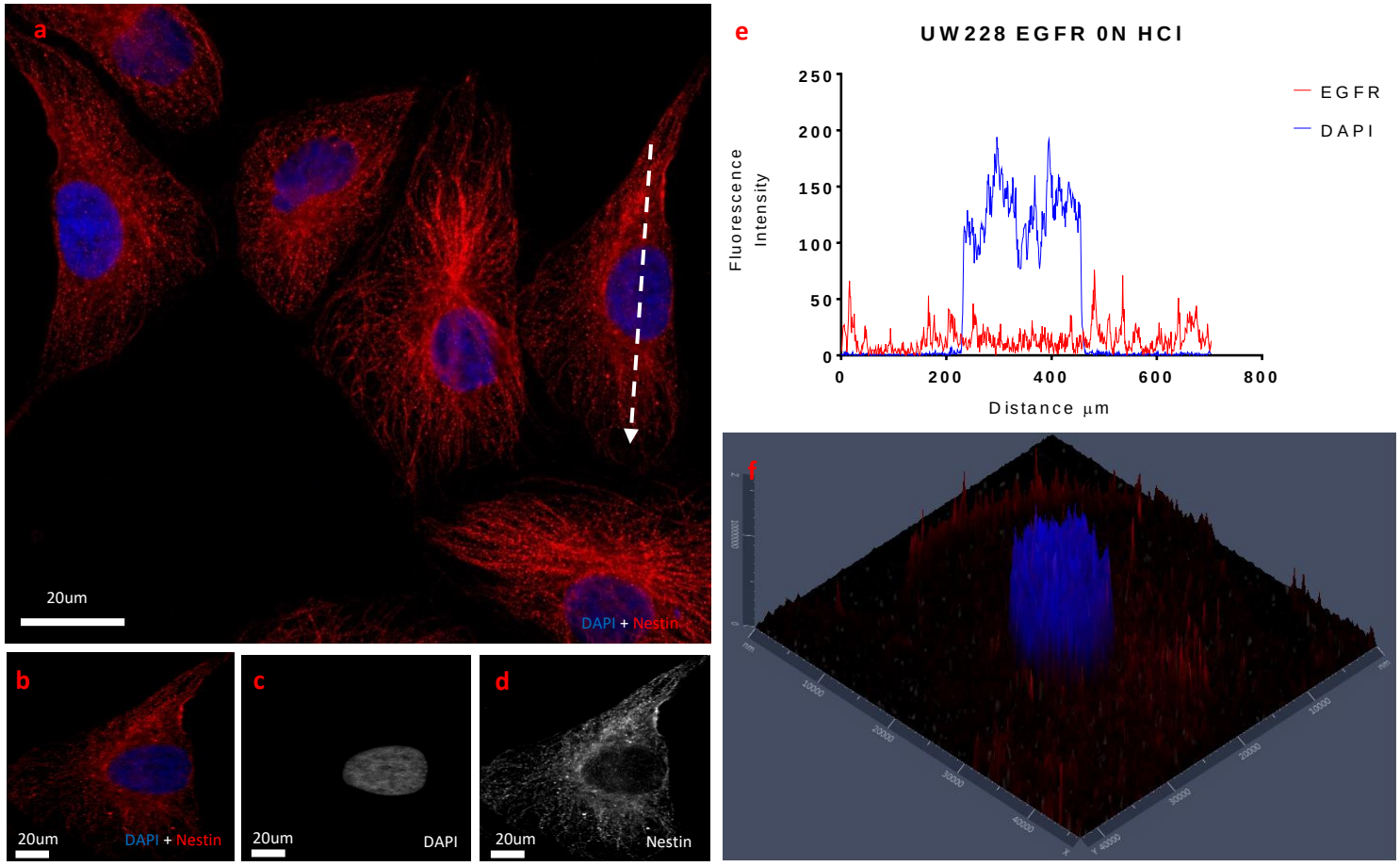
Supplementary Figure 8: **a-d** Immunostaining images from confocal microscope representing Nestin in LN229 with no treatment of HCl, includes merge of DAPI and Nestin (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Nestin (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



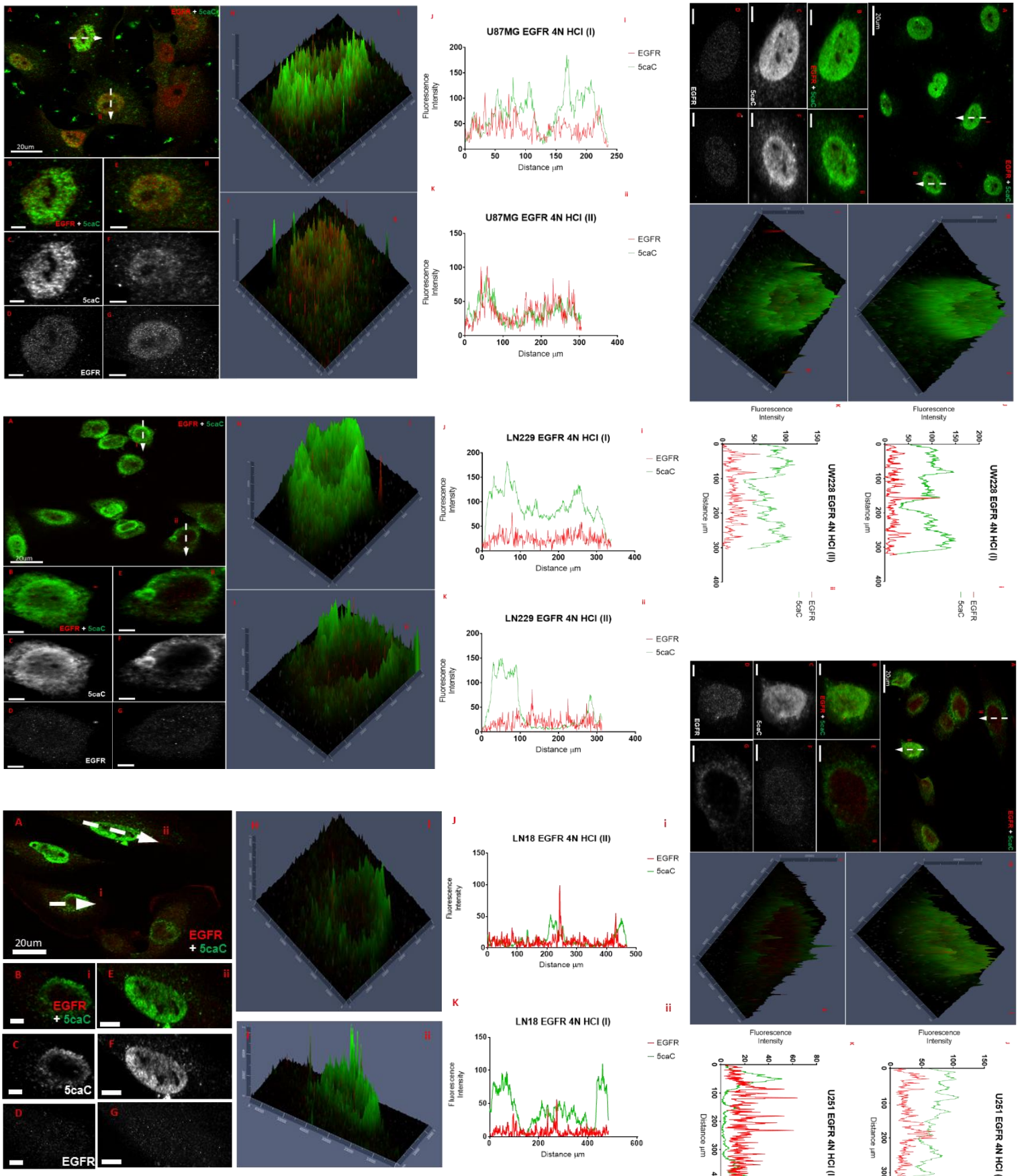
Supplementary Figure 9: **a-d** Immunostaining images from confocal microscope representing Nestin in U251 with no treatment of HCl, includes merge of DAPI and Nestin (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Nestin (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



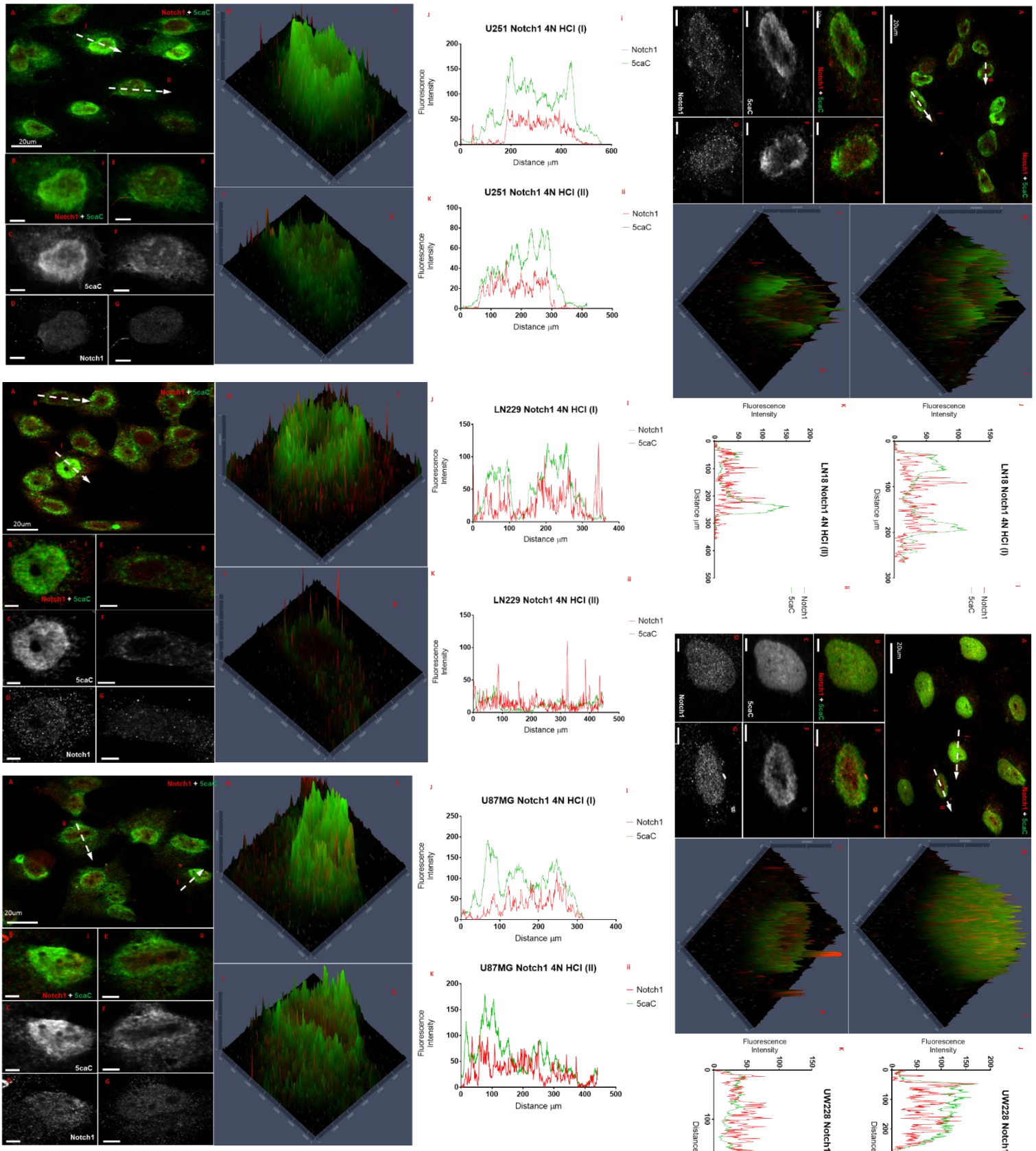
Supplementary Figure 10: **a-d** Immunostaining images from confocal microscope representing Nestin in U87MG with no treatment of HCl, includes merge of DAPI and Nestin (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Nestin (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



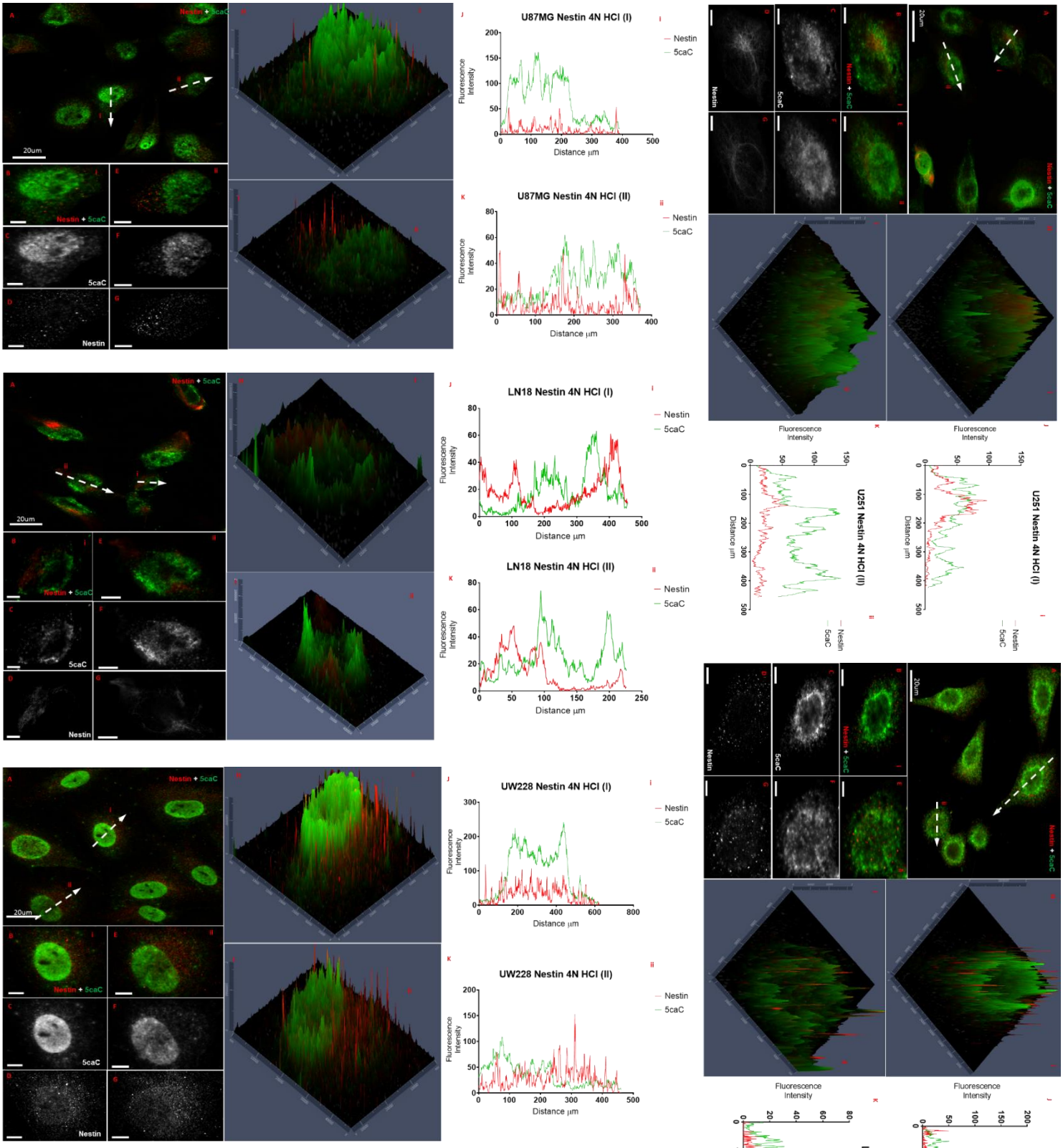
Supplementary Figure 11: **a-d** Immunostaining images from confocal microscope representing Nestin in UW228-3 with no treatment of HCl, includes merge of DAPI and Nestin (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Nestin (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



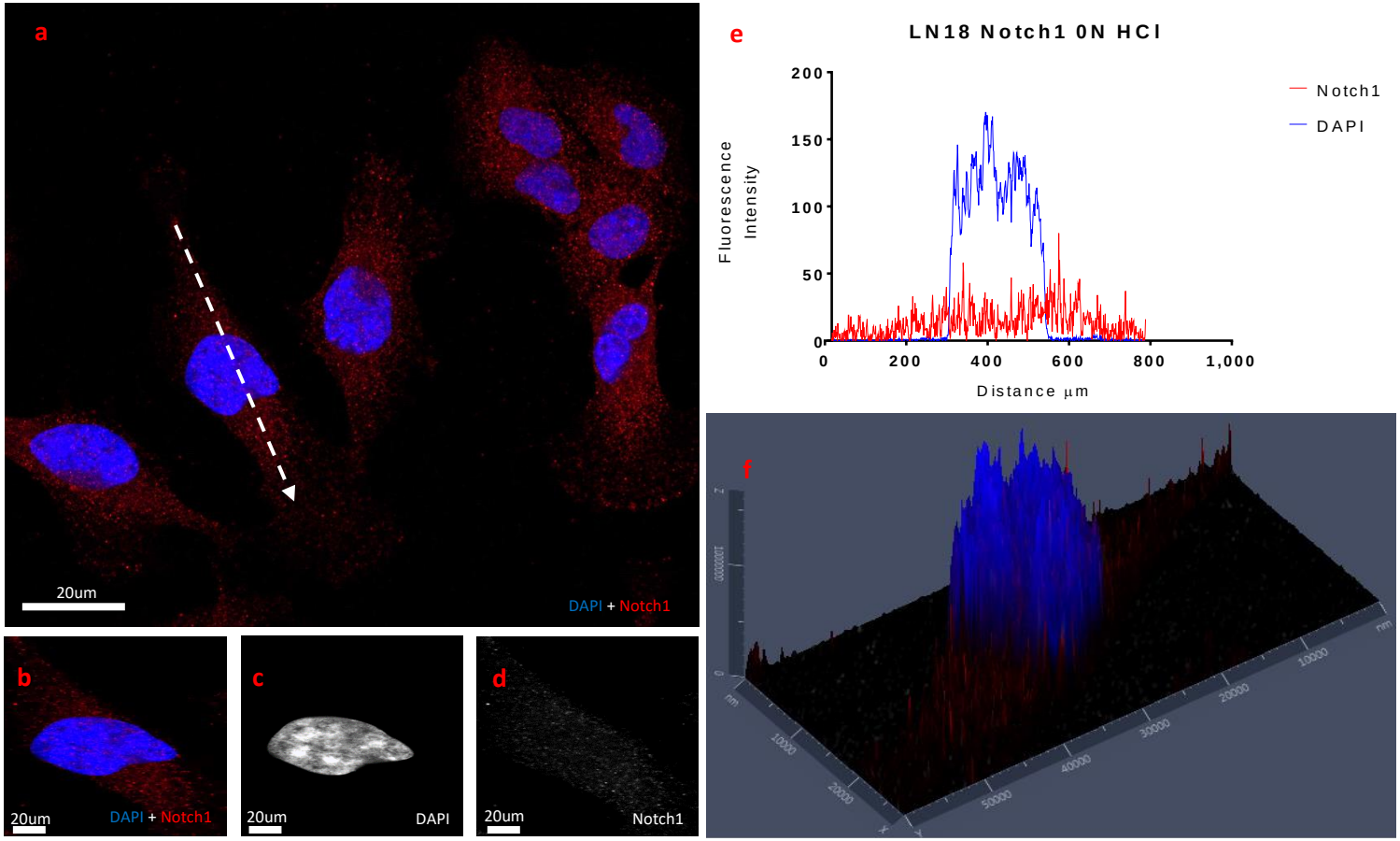
Supplementary Fig 12: Analysis of each of cell line stained for EGFR and 5caC; 2.5D Plots representing staining location for EGFR and 5caC on a single cell within each cell line; Fluorescence intensity graph representing strengths for each signal emitted by EGFR and 5caC with average values from different cells within each cell line.



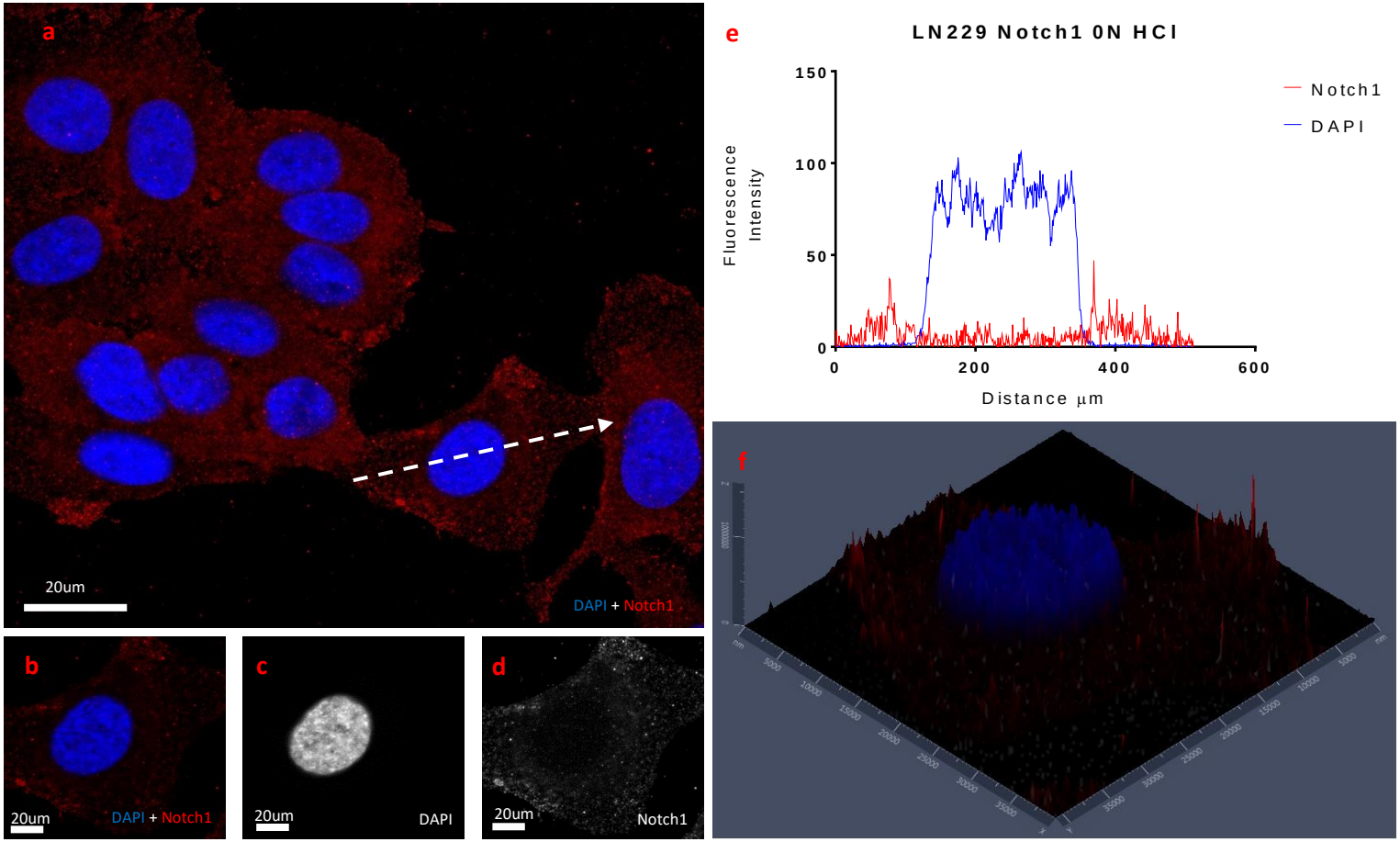
Supplementary Fig 13: Analysis of each of cell line stained for Notch1 and 5caC; 2.5D Plots representing staining location for Notch1 and 5caC on a single cell within each cell line; Fluorescence intensity graph representing strengths for each signal emitted by Notch1 and 5caC with average values from different cells within each cell line.



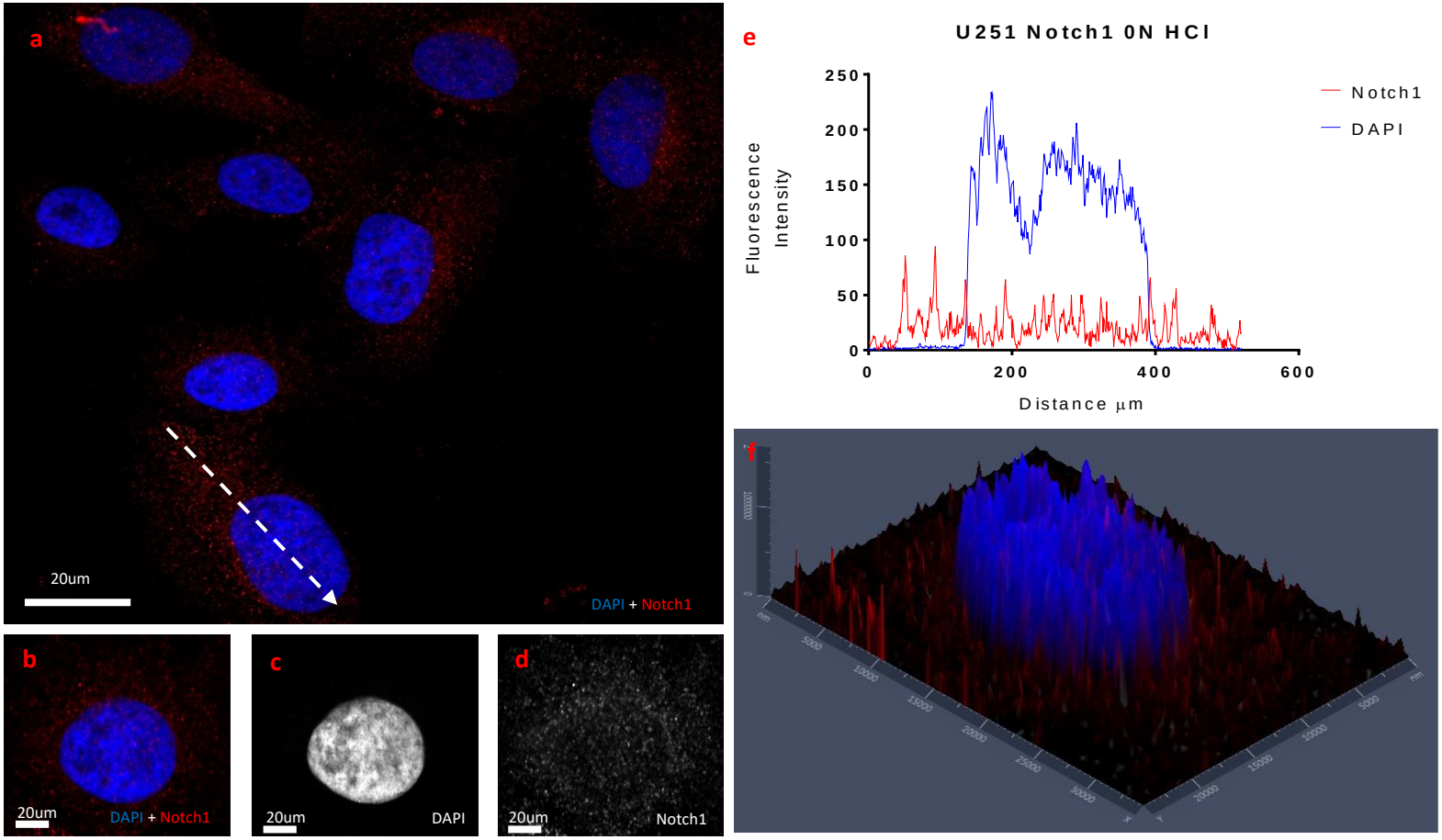
Supplementary Fig 14: Analysis of each of cell line stained for Nestin and 5caC; 2.5D Plots representing staining location for Nestin and 5caC on a single cell within each cell line; Fluorescence intensity graph representing strengths for each signal emitted by Nestin and 5caC with average values from different cells within each cell line.



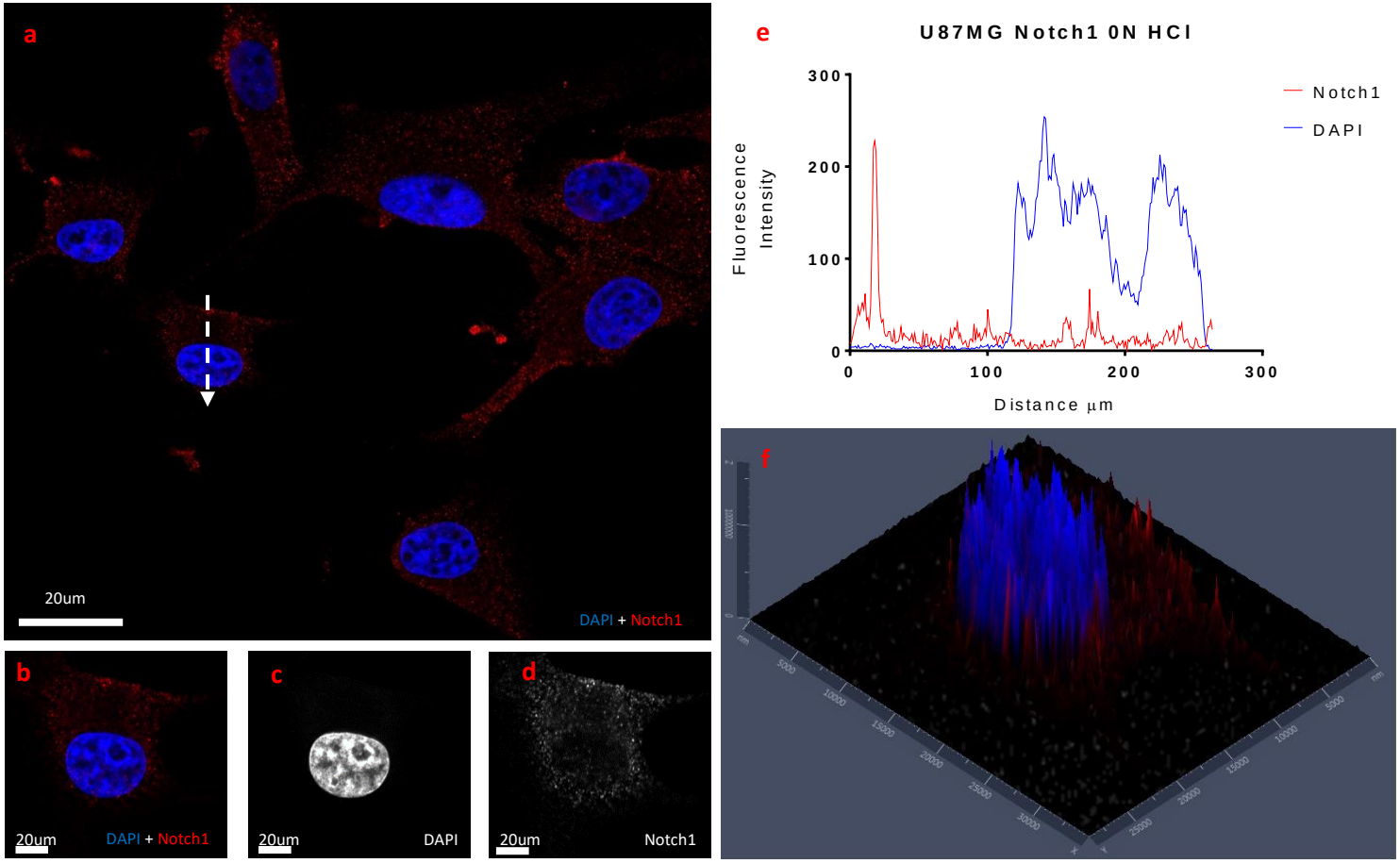
Supplementary Figure 15: **a-d** Immunostaining images from confocal microscope representing Notch1 in LN18 with no treatment of HCl, includes merge of DAPI and Notch1 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Notch1 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



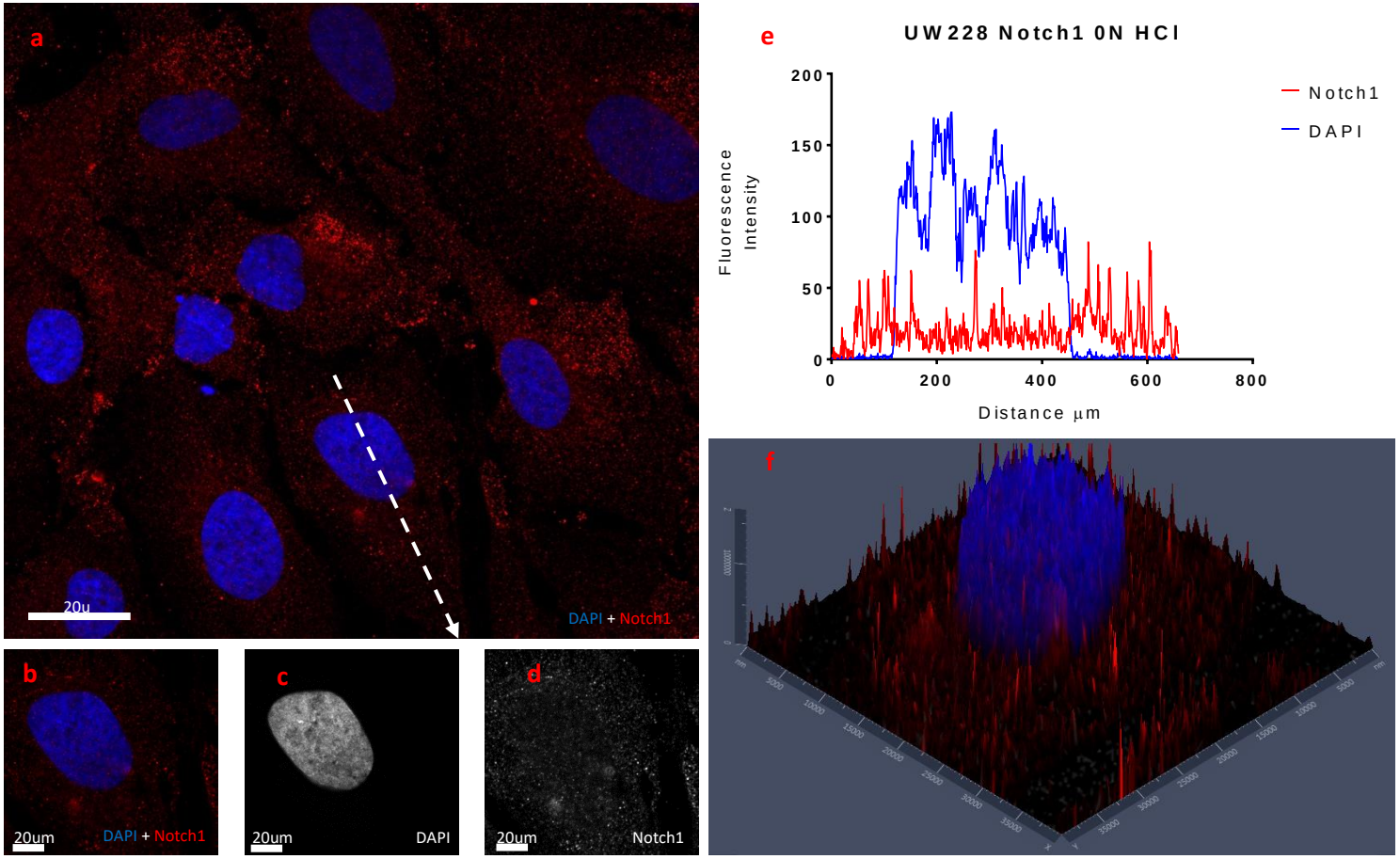
Supplementary Figure 16: **a-d** Immunostaining images from confocal microscope representing Notch1 in LN229 with no treatment of HCl, includes merge of DAPI and Notch1 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Notch1 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



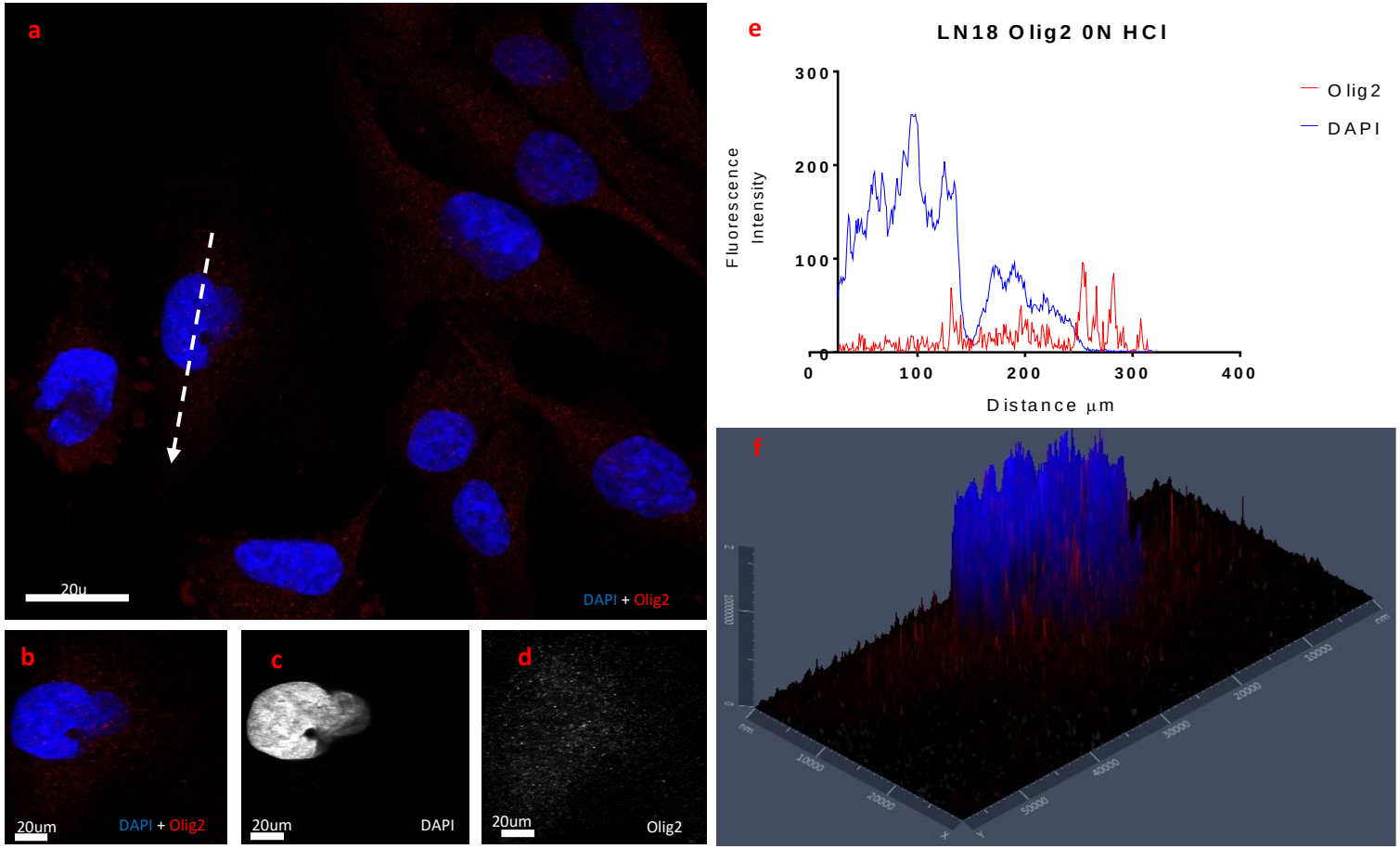
Supplementary Figure 17: **a-d** Immunostaining images from confocal microscope representing Notch1 in U251 with no treatment of HCl, includes merge of DAPI and Notch1 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Notch1 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



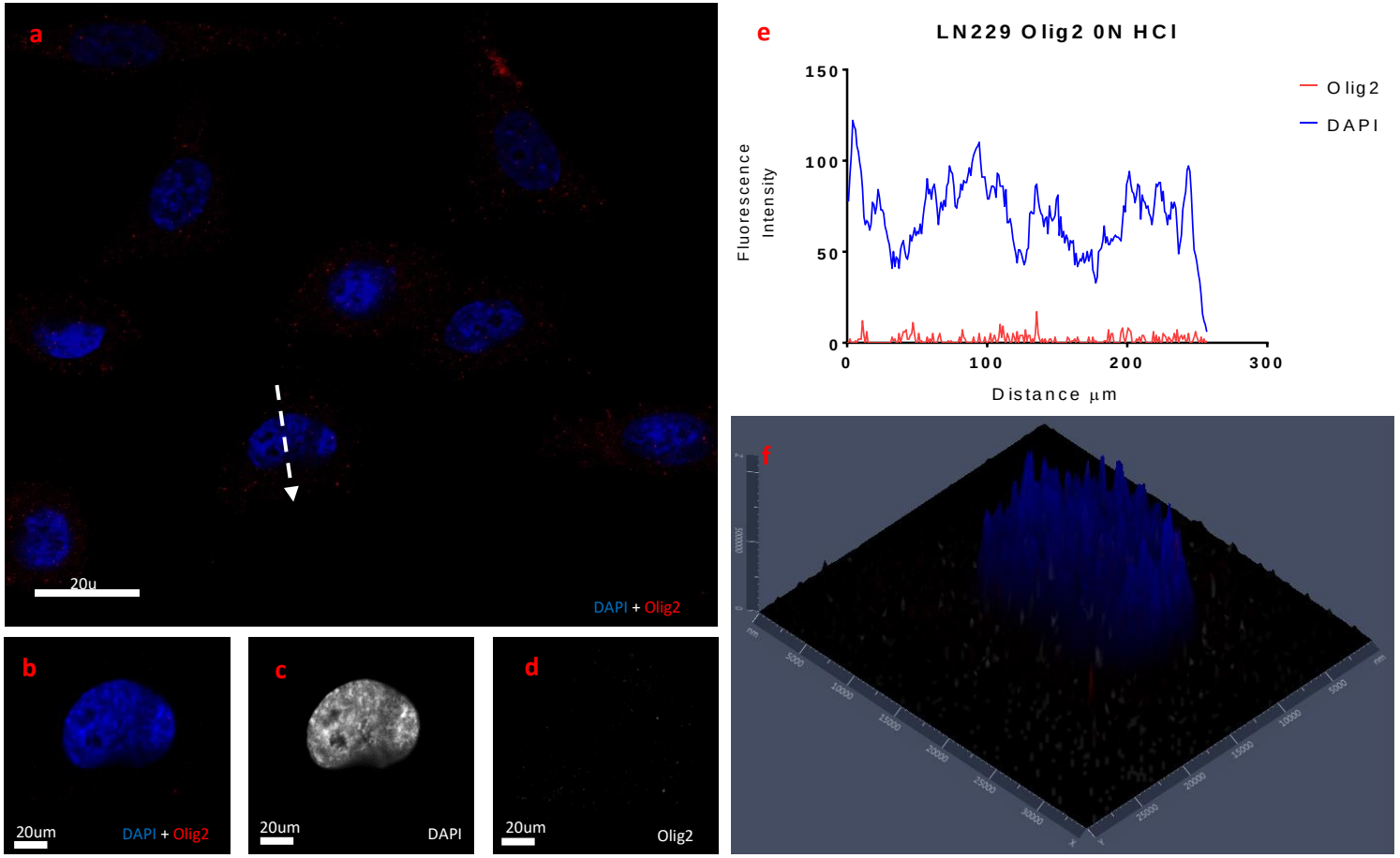
Supplementary Figure 18: **a-d** Immunostaining images from confocal microscope representing Notch1 in U87MG with no treatment of HCl, includes merge of DAPI and Notch1 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Notch1 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



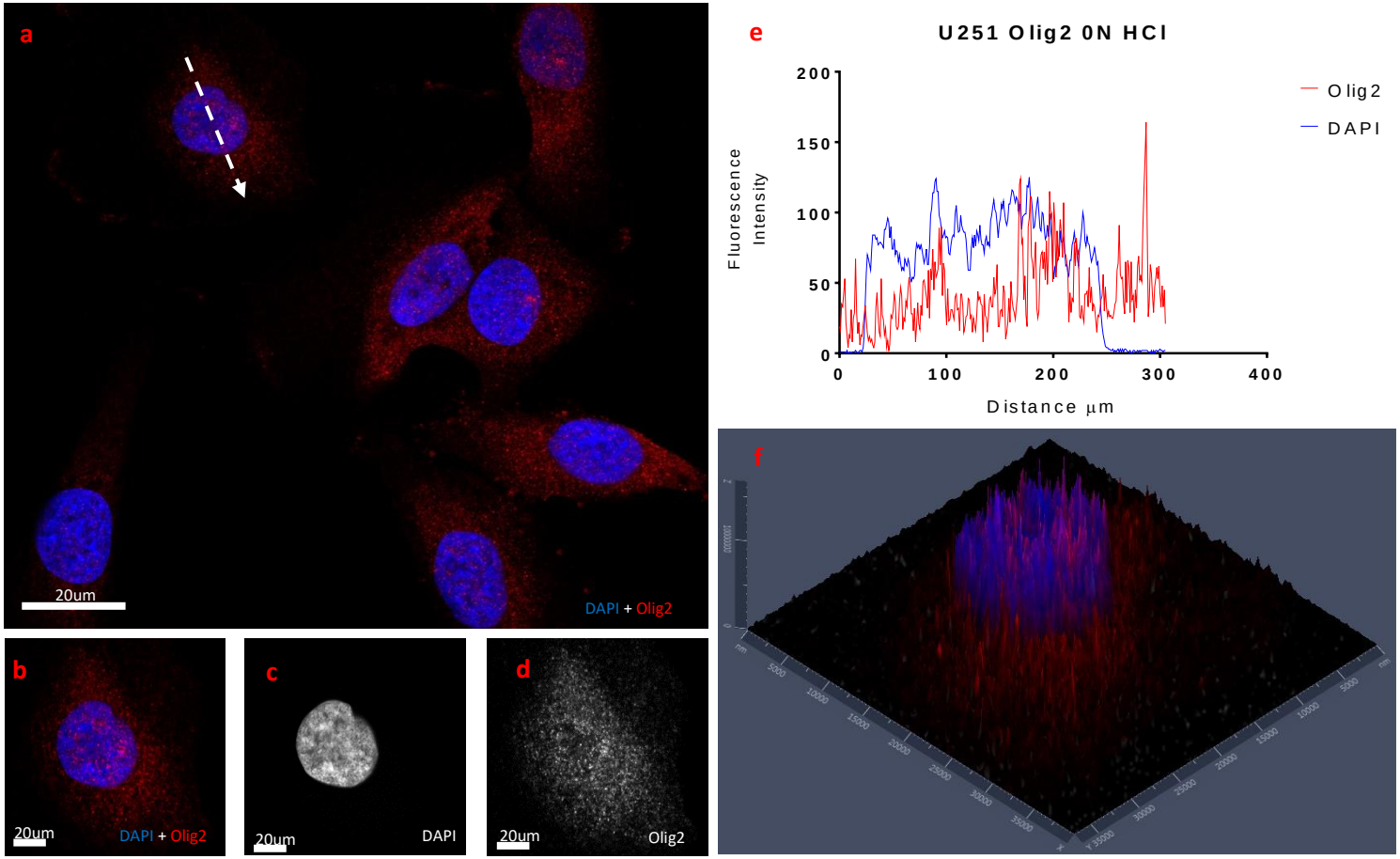
Supplementary Figure 19: **a-d** Immunostaining images from confocal microscope representing Notch1 in UW228-3 with no treatment of HCl, includes merge of DAPI and Notch1 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Notch1 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



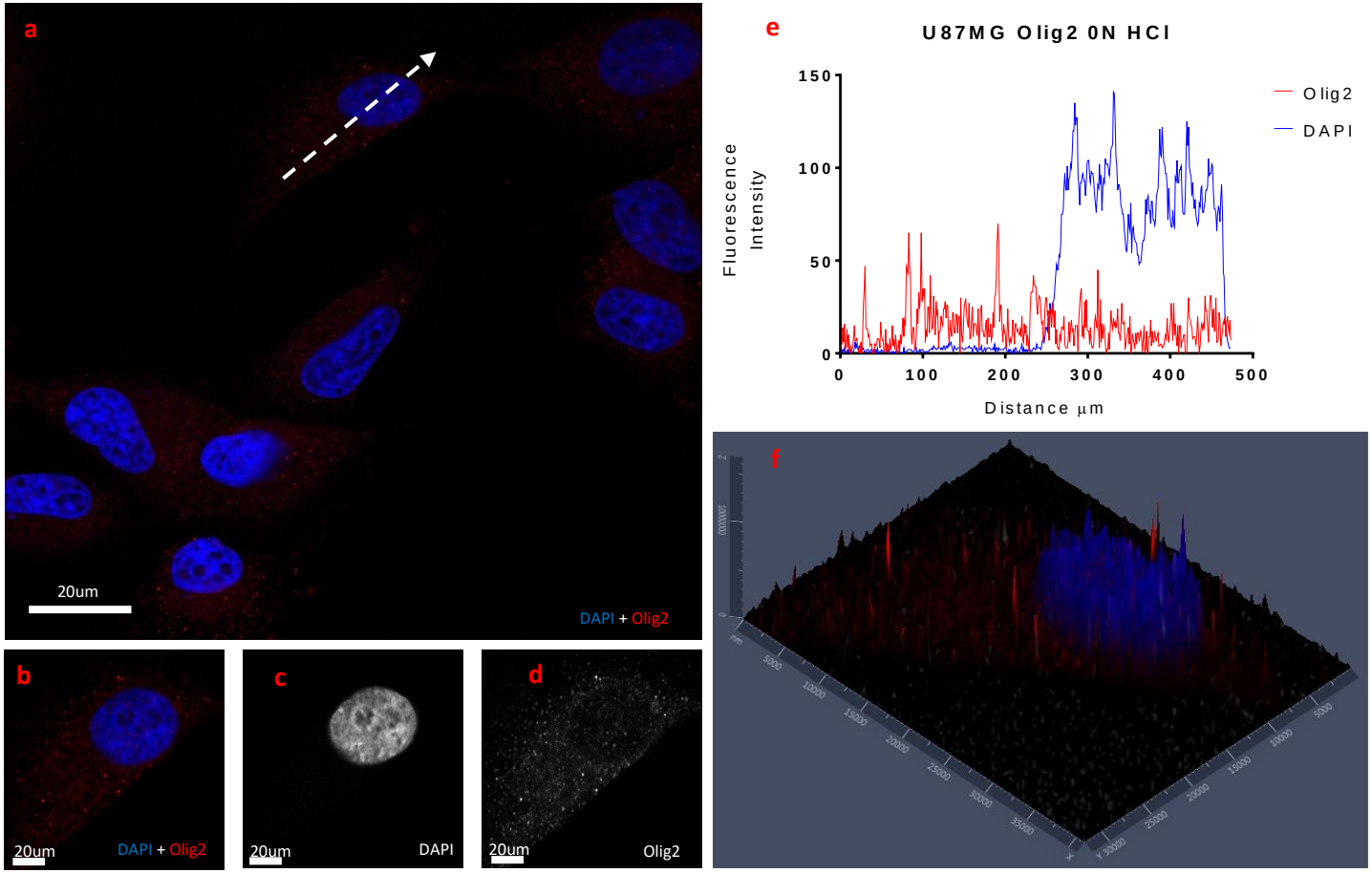
Supplementary Figure 20: **a-d** Immunostaining images from confocal microscope representing Olig2 in LN18 with no treatment of HCl, includes merge of DAPI and Olig2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Olig2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



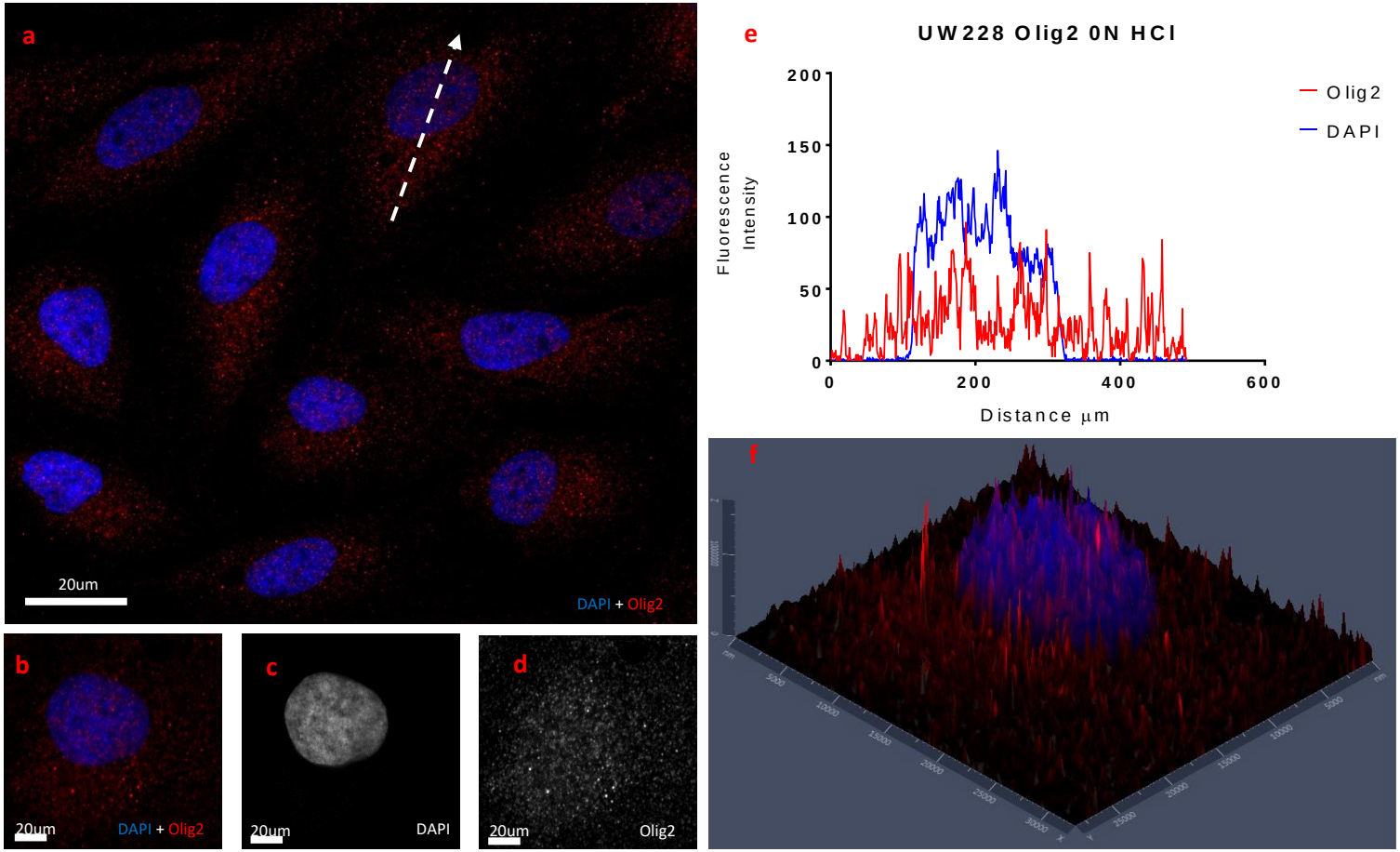
Supplementary Figure 21: **a-d** Immunostaining images from confocal microscope representing Olig2 in LN229 with no treatment of HCl, includes merge of DAPI and Olig2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Olig2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



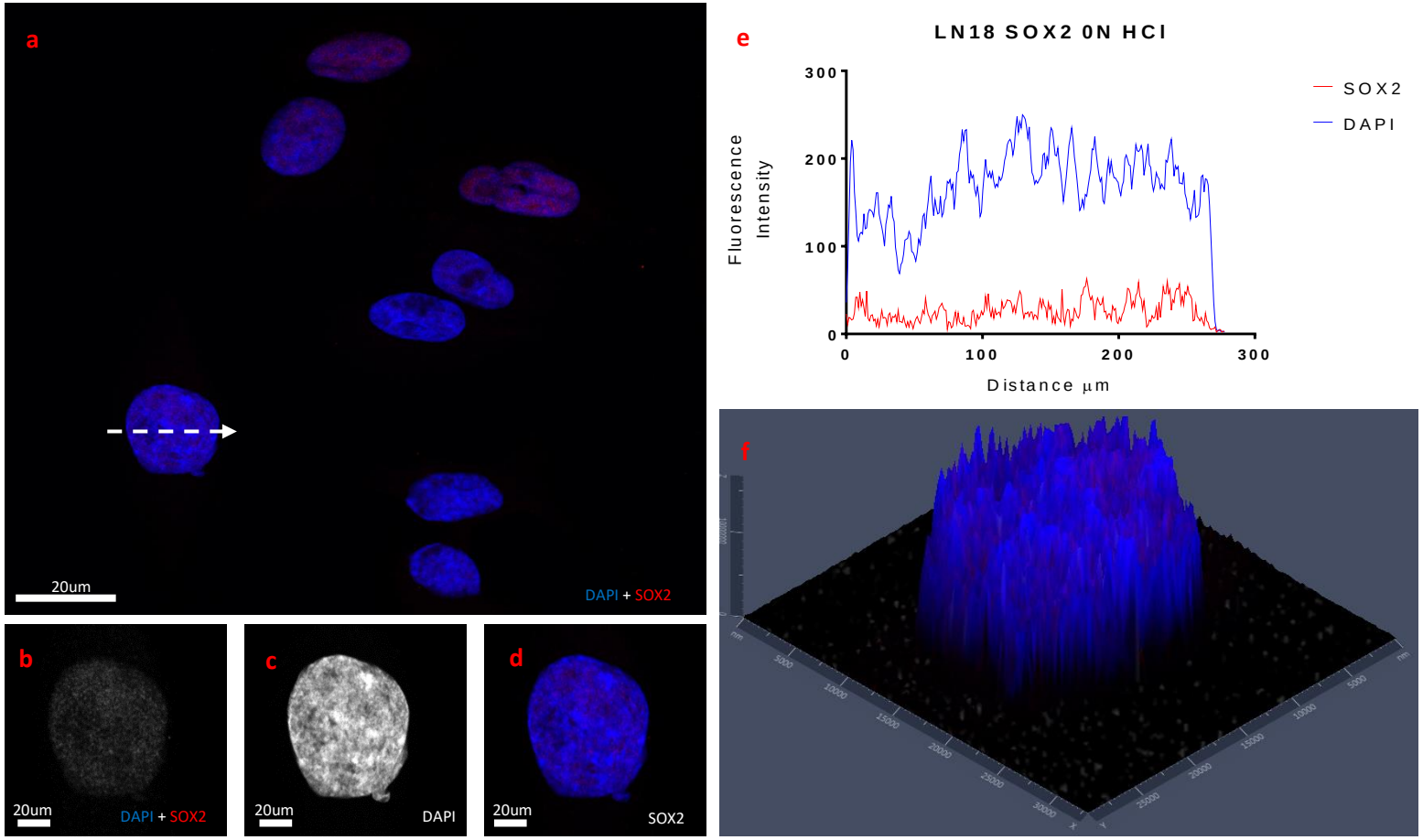
Supplementary Figure 19: **a-d** Immunostaining images from confocal microscope representing Olig2 in U251 with no treatment of HCl, includes merge of DAPI and Olig2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Olig2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



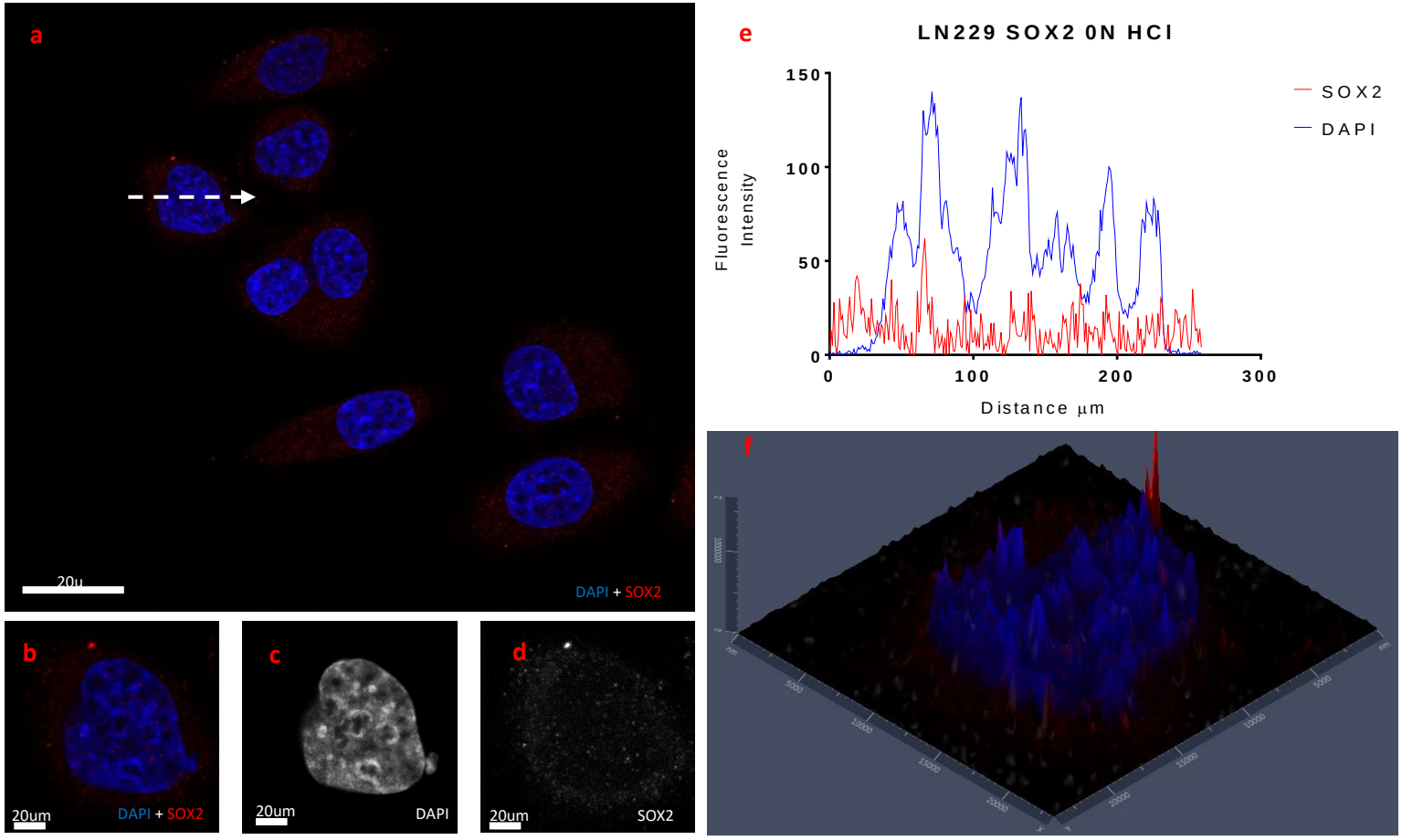
Supplementary Figure 22: **a-d** Immunostaining images from confocal microscope representing Olig2 in U87MG with no treatment of HCl, includes merge of DAPI and Olig2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Olig2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



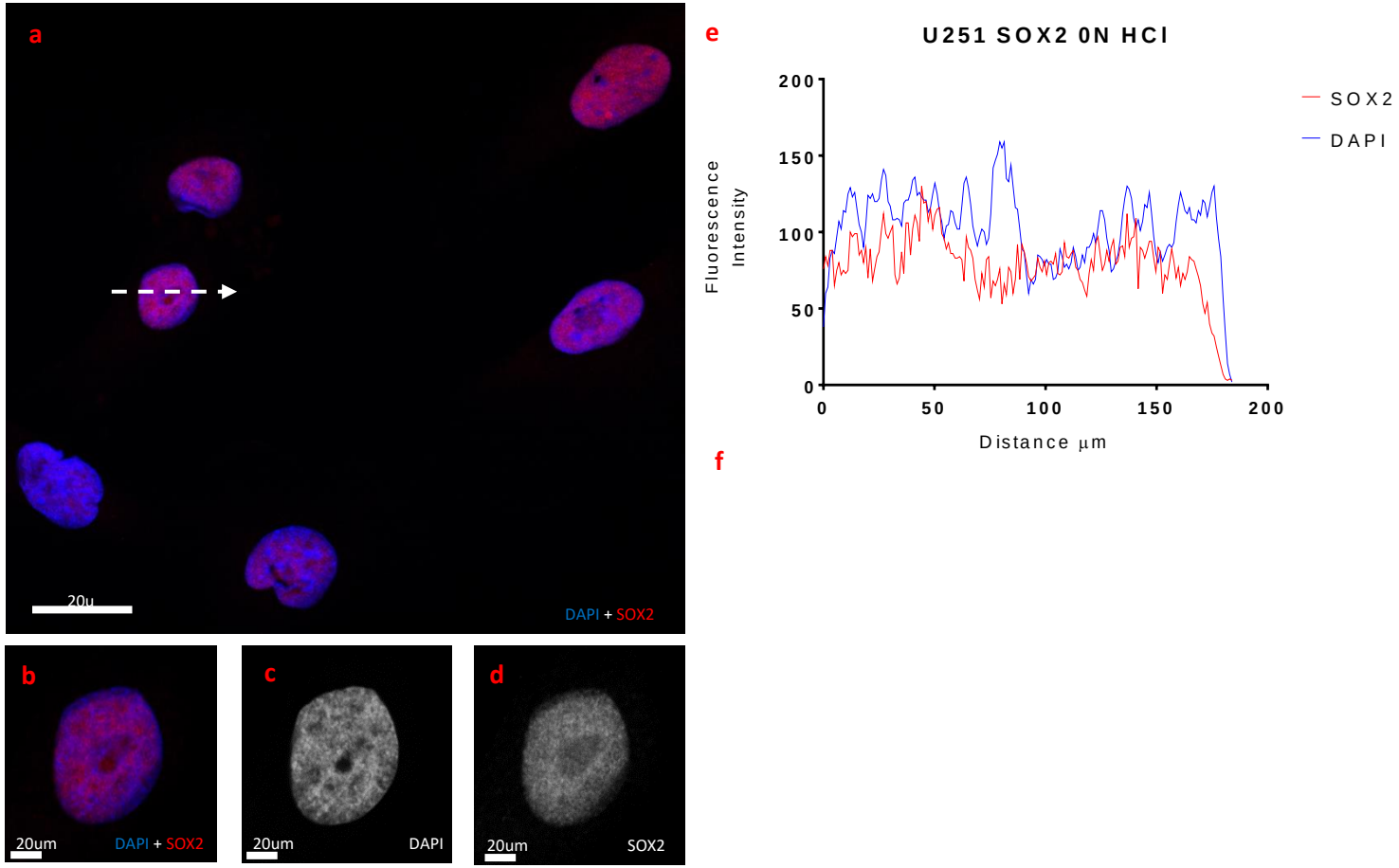
Supplementary Figure 23: **a-d** Immunostaining images from confocal microscope representing Olig2 in UW228-3 with no treatment of HCl, includes merge of DAPI and Olig2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and Olig2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



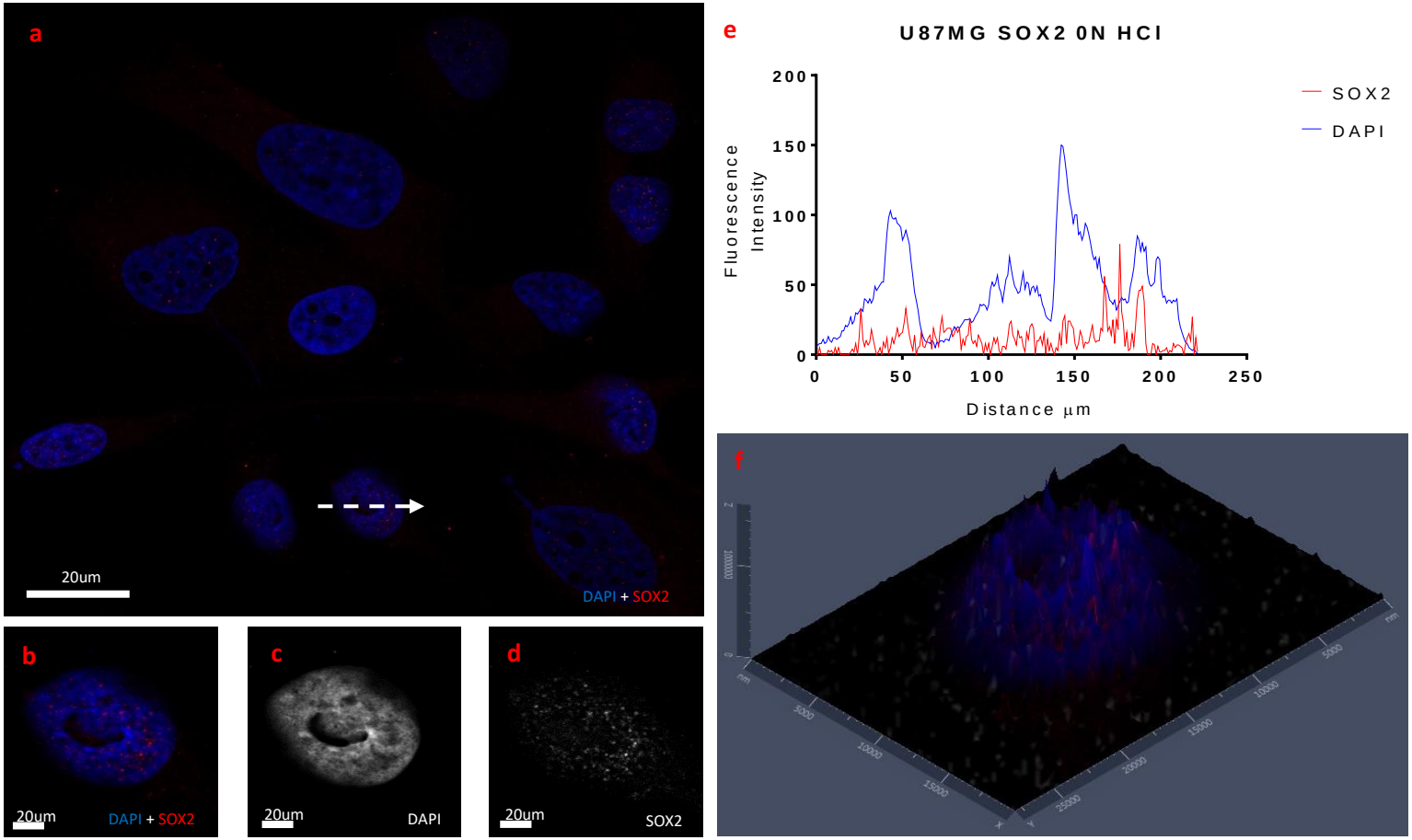
Supplementary Figure 24: **a-d** Immunostaining images from confocal microscope representing SOX2 in LN18 with no treatment of HCl, includes merge of DAPI and SOX2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and SOX2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



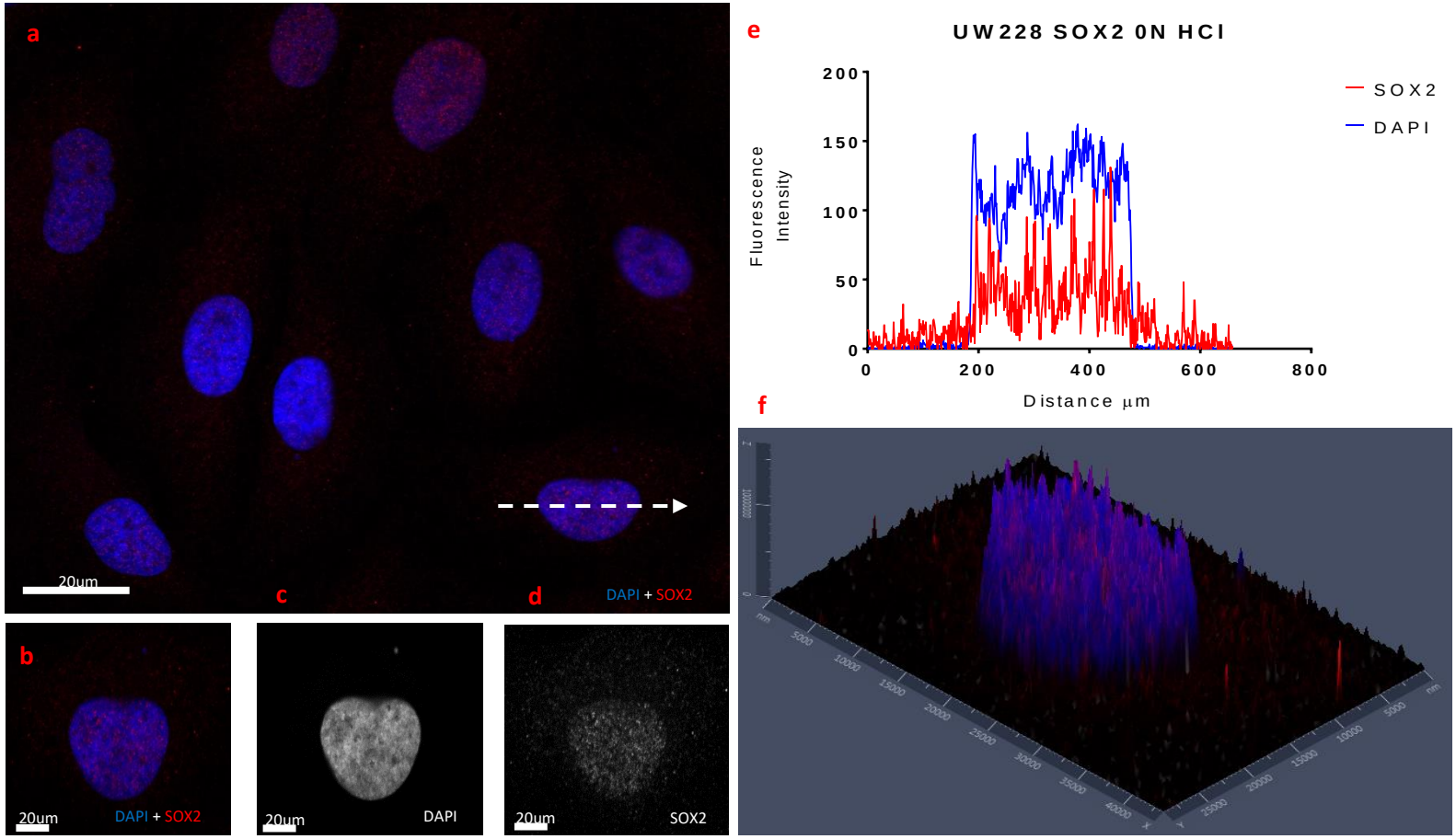
Supplementary Figure 25: **a-d** Immunostaining images from confocal microscope representing SOX2 in LN229 with no treatment of HCl, includes merge of DAPI and SOX2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and SOX2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



Supplementary Figure 26: **a-d** Immunostaining images from confocal microscope representing SOX2 in U251 with no treatment of HCl, includes merge of DAPI and SOX2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and SOX2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



Supplementary Figure 27: **a-d** Immunostaining images from confocal microscope representing SOX2 in U87MG with no treatment of HCl, includes merge of DAPI and SOX2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and SOX2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.



Supplementary Figure 28: **a-d** Immunostaining images from confocal microscope representing SOX2 in UW228-3 with no treatment of HCl, includes merge of DAPI and SOX2 (blue & red channels) and grayscale single cell (arrowed) images of split channels of DAPI (blue channel) and SOX2 (red channel). **e** Fluorescence intensity graph representing signal produced by the single cell (arrowed) for red and blue channels. **f** 2.5D plot of single cell (arrowed) showing location and intensity of staining.

10. ABBREVIATIONS

2OG	2-oxoglutarate
2OGFeDO	2-oxogluterate (2OG)-Fe (II)-dependent dioxygenase
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AML	Acute myeloid leukaemia
ANOVA	Analysis of variance
B3T	Beta tubulin III
BER	Base excision repair
bHLH	Basic helix-loop-helix proteins
BTIC	Brain tumour initiating cells

CDKN2A	Cyclin-dependent kinase inhibitor 2A
cDNA	Complementary DNA
CFU-Fs	Colony-forming units-fibroblastic
CGI	CpG islands
CIMP	CpG island methylator phenotype
CNS	Central nervous system
CpG	C-phosphate-G
CSC	Cancer stem cells
CT	Cycle threshold
DAPK	Death-associated protein kinase
DIP	DNA immunoprecipitation

DMEM	Dulbecco's Modified Eagle's medium
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DSBH	Double-stranded β helix
DNMT	DNA methyltransferase
DTT	Dithiothreitol
EGF-	Epidermal growth factor-
EGFR	Epidermal growth factor receptor
ESCs	Embryonic stem cells
Fe(II)	Iron (II) Oxide
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase

GBM	Glioblastoma Multiforme
G-CIMP	Glioma-CpG island methylator phenotype
GSC	Glioma/Glioblastoma stem cells
HCl	Hydrochloric acid
IDH1/2	Isocitrate dehydrogenase ½
IGF-2	Insulin-like growth factor-2
iPSCs	Induced pluripotent stem cells
KD1/2	Knockdown1/2
MAGE-	Melanoma-associated antigen-
mESCs	Mouse embryonic stem cells
MET	Mesenchymal-epidermal transition
MGMT	O^6-methylguanine-DNA methyltransferase

MLL	Mixed lineage leukaemia
MS1	Musashi-1
NAPDH	Nicotinamide adenine dinucleotide phosphate-oxidase
NCSI	National Centre for Science Information
NECD	Notch extracellular domain
NF1	Neurofibromin 1
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NICD	Notch intracellular domain
NPC	Neural progenitor cells
NSCs	Neural stem cells
OCT4	Octamer-binding transcription factor 4

Olig2	Oligodendrocyte transcription factor-2
PBS	Phosphate-buffered saline
PBT	PBS with Tween
PCR	Polymerase chain reaction
PDGFRA	Platelet derived growth factor receptor Alpha
PI3K	Phosphoinositide 3-kinase
PRC2	Polycomb repressive complex 2
PTEN	Phosphatase and tensin homolog
qPCR	Quantitative/real-time PCR
Rb1	Retinoblastoma protein
RNA	Ribonucleic acid
RT	Reverse transcriptase

Sat2	Satellite-2 DNA sequence
SOX2	Sex determining region Y-Box 2
SSEA-1	Stage-specific embryonic antigen-1
STAT3	Signal transducer and activator of transcription 3
T-ALL	T-cell Acute lymphoblastic leukaemia
TCGA	The Cancer Genome Atlas
TDG	Thymine DNA glycosylase
TET	Ten-eleven translocation
TNF	Tumour necrosis factor
TP53	Tumour protein p53
TSA	Tyramide signal amplification

UCSC	University of California, Santa Cruz
WHO	World Health Organisation
α-KG	Alpha-ketoglutaric acid
ΔEGFR	Active mutant of EGFR