Molecules in pathogenesis

Subcellular localisation of the stem cell markers OCT4, SOX2, NANOG, KLF4 and c-MYC in cancer: a review

Bede van Schaijik,1 Paul F Davis,1 Agadha C Wickremesekera,1,2 Swee T Tan,1,3 Tinte Itinteang1

1Gillies McIndoe Research Institute, Wellington, New Zealand
2Department of Neurosurgery, Wellington Regional Hospital, Wellington, New Zealand
3Wellington Regional Plastic, Maxillofacial and Burns Unit, Hutt Hospital, Wellington, New Zealand

Correspondence to
Dr Tinte Itinteang, Gillies McIndoe Research Institute, Wellington 6242, New Zealand;
tinte@gmr.org.nz

Received 28 September 2017
Revised 5 November 2017
Accepted 6 November 2017
Published Online First 27 November 2017

ABSTRACT
The stem cell markers octamer-binding transcription factor 4, sex-determining region Y-box 2, NANOG, Kruppel-like factor 4 and c-MYC are key factors in inducing pluripotency in somatic cells, and they have been used to detect cancer stem cell subpopulations in a range of cancer types. Recent literature has described the subcellular localisation of these markers and their potential implications on cellular function. This is a relatively complex and unexplored area of research, and the extent of the effect that subcellular localisation has on cancer development and growth is largely unknown. This review analyses this area of research in the context of the biology of stem cells and cancer and explores the potential modulating effect of subcellular localisation of these proteins as supported by the literature.

INTRODUCTION
The cancer stem cell (CSC) concept of cancer proposes that not all cancer cells participate in tumour formation and that the development and progression of cancer is driven by CSCs, a small subpopulation of cells that possess the potential for self-renewal.1 Induced pluripotent stem cells (iPSC) were first created by Yamanaka using octamer binding transcription factor 4 (OCT4), sex-determining region Y-box 2 (SOX2), Kruppel-like factor 4 (KLF4) and c-MYC through retroviral transduction in mouse fibroblasts.2 Thomson further demonstrated the same net effect with the use of the NANOG gene8 in combination with the above factors to induce pluripotency.1 These markers have also been used to identify CSC subpopulations in a variety of cancers, such as glioblastoma8 and colorectal cancer (CRC),3 and they will be the focus for this review.

Traditionally, the synthesis of transcription factors begins with specific transcription of mRNA from the genomic DNA with subsequent migration of the mRNA to the ribosomes in the cytoplasm for translation (figure 1). The protein products then migrate back into the nucleus to regulate downstream transcriptional activity.3 It has been shown that the subcellular localisation of transcription factors is affected by many cellular and molecular elements, including binding partners, protein isoforms, and localisation signals, and may influence the function of stem cell-associated proteins.3 Furthermore, the localisation of stem cell markers can change as cells progress down the stem cell hierarchy. The subcellular localisation of CSC markers is a poorly understood area, and analyses are limited by the specificity of antibodies used in the assays. Therefore, research into the effects of their cytoplasmic versus nuclear localisation may provide insights into how they induce cancer development and growth.

This review summarises the current understanding of the aforementioned stem cell markers OCT4, SOX2, NANOG, KLF4 and c-MYC, with regards to their subcellular localisation in the context of CSCs in a variety of cancers.

OCTAMER-BINDING TRANSCRIPTION FACTOR 4
OCT4 is a member of the mammalian Pit-Oct-Unc domain transcription factors encoded by the Pous/f1 gene and is involved in embryogenesis, stem cell maintenance, tumour growth and metastasis.3 Following synthesis within the cytoplasm, OCT4 is imported into the nucleus by the binding of importin-α, which recognises an RKRKR motif as a nuclear localisation signal (NLS).3 In iPSCs, studies have shown that OCT4 is localised to the nucleus in undifferentiated iPSCs generated through lentivirus-mediated transduction of OCT4, SOX2, NANOG and LIN28 in foreskin cells.10 However, OCT4 mutants containing a nuclear export signal (NES), which were primarily localised to the cytoplasm, have also been shown to maintain the undifferentiated state of human embryonic stem cells (hESCs),11 suggesting that the localisation of OCT4 is insignificant as long as nuclear processing is maintained. Immunohistochemical (IHC) staining was used to identify cytoplasmic OCT4 in CSC subpopulations in two groups of oral cavity squamous cell carcinomas (SCC) (figure 2A),12 13 with a third group showing both nuclear and cytoplasmic localisation,14 consistent with findings in other tumours such as glioma15 and the epithelial cells of CRC.1 A report by Alexander et al16 which also used the same antibody as that used by Baillie et al12 and Ram et al13 also reported cytoplasmic OCT4, and attributed this to neuroendocrine differentiation; however, the latter remains to be conclusively determined. It has also been reported that the isoforms of OCT4, OCT4_A and OCT4_B exhibit different subcellular localisations: OCT4_A being localised to the nucleus, whereas OCT4_B is found in the cytoplasm.17 The role of OCT4_B has not been elucidated18; therefore, its cytoplasmic localisation suggests that it has no transcriptional function.


SOX2 is a member of the SRY-related high mobility group (HMG) box (SOX) gene family that encodes transcription factors with a single HMG DNA-binding domain and functions to preserve developmental potential. Differential expression patterns of SOX2 have been reported in different cancer types, with oral cavity SCC (OCSCC) CSC subpopulations showing SOX2 localisation to both the nucleus and cytoplasm across different cancer types. Furthermore, it is known that hESCs express only the OCT4_A isoform, meaning that there is no evidence that OCT4_B is required to confer pluripotency. Further analysis of which OCT4 isoform is expressed coupled with the knowledge of OCT4 antibodies specificity for individual isoforms is required to fully understand the role of subcellular localisation of OCT4 in cancer development and growth.

**SEX-DETERMINING REGION Y-BOX 2**

SOX2 is a member of the SRY-related high mobility group (HMG) box (SOX) gene family that encodes transcription factors with a single HMG DNA-binding domain and functions to preserve developmental potential. Differential expression patterns of SOX2 have been reported in different cancer types, with oral cavity SCC (OCSCC) CSC subpopulations showing SOX2 localisation to both the nucleus and cytoplasm across different cancer types. Furthermore, it is known that hESCs express only the OCT4_A isoform, meaning that there is no evidence that OCT4_B is required to confer pluripotency. Further analysis of which OCT4 isoform is expressed coupled with the knowledge of OCT4 antibodies specificity for individual isoforms is required to fully understand the role of subcellular localisation of OCT4 in cancer development and growth.

**NANOG**

NANOG is a homeodomain transcription factor that directs propagation of undifferentiated hESCs and mediates induction of pluripotency. Nuclear NANOG has been confirmed in cancers including seminoma, CRC, hepatocellular carcinoma and OCSCC cell lines and tissue samples. In contrast, cytoplasmic NANOG has been observed in the invasive regions of NPC and OCSCC and in malignant cervical epithelial cells as well as mesenchymal stem cells within the stroma of cervical cancer. In glioblastoma, NANOG has been shown to be localised to both the nucleus and cytoplasm (figure 2B). These findings can be attributed to the two NLSs and one NES present in the NANOG amino acid sequence, which suggests a cellular shuttling behaviour similar to that of OCT4 and SOX2. Given that pluripotency is a dynamic state, this may explain how the levels of NANOG, SOX2 and OCT4 control development and maintenance of stem cells.

In iPSCs that express OCT4, SOX2, KLF4 and c-MYC, nuclear NANOG has been detected, suggesting that the induction of pluripotency by these factors initiates a cascade of gene activation that generates pluripotency. In contrast, conjugated fusion proteins of NANOG used in cell-free methods of inducing pluripotency have been shown to be localised to the cytoplasm and perinuclear regions of mouse embryonic fibroblasts. This suggests a range of functionalities for NANOG depending on its subcellular localisation. However, Huangfu et al demonstrate that iPSCs can be generated through ectopic expression of OCT4 and SOX2, as well as NANOG and LIN28, by means of somatic cell nuclear transfer. This infers that, while genetically dispensable as long as the other transcription factors are being expressed, NANOG still plays the nucleus, leading to the concept that a binding partner may stimulate nuclear localisation.

iPSCs show preferential nuclear localisation of SOX2, with supporting evidence demonstrating the induction of pluripotency through transferring the nuclear contents of embryonic stem cells into somatic cells. This would suggest that the reprogramming activity is confined to the nucleus rather than the cytoplasm. These results are consistent with the understanding that the mechanistic role of SOX2 in stem cell induction and formation is predominantly through its action on the cell nucleus.

![Image of protein synthesis and subsequent localisation](http://jcp.bmj.com/)

**Figure 1** Overview of protein synthesis and subsequent localisation. Following transcription in the nucleus, mRNA is transported to the cytoplasm where proteins are synthesised by the ribosomes. The resulting protein product can remain in the cytoplasm or be transported into the nucleus where it can act as a transcription factor and stimulate downstream gene expression.

**Figure 2** Immunohistochemical staining of (A) a moderately differentiated oral tongue squamous cell carcinoma demonstrating staining for OCT4 (green and white arrows) localised to the cytoplasm of the cells within the stroma and SOX2 (red) expressed throughout the cytoplasm and the nuclei (blue); (B) a grade IV glioblastoma demonstrating staining for NANOG (red) in the nuclei (blue) and glial fibrillar acidic protein (green). Nuclei are counterstained with DAPI (A, B: blue). Original magnification: 400×. OCT4, octamer-binding transcription factor 4; SOX2, sex-determining region Y-box 2.
a role in maintaining pluripotency and that its localisation in terms of its function in iPSCs remains to be conclusively determined.

**KRUPPEL-LIKE FACTOR 4**

KLF4 is a member of the ‘kruppel’ family of zinc-finger transcription factors involved in the regulation of the cell cycle, maintenance of pluripotency and epidermal development and is abundant in epithelial cells of the intestine, colon and thymus. It contains two independent NLSs that control nuclear import. Detection by in situ hybridisation has revealed both nuclear and cytoplasmic localisation in oesophageal SCC. In breast cancer, nuclear localisation occurs with induction of malignant transformation and in cutaneous SCCs strong nuclear staining of KLF4 is observed when compared with their adjacent non-tumour areas. In contrast, KLF4 is predominantly cytoplasmic in prostate cancer cells. A KLF4 isoform, KLF4α, lacks a NLS and is therefore located in the cytoplasm in both prostate and pancreatic cancer. Furthermore, CRM1-mediated nuclear export and subsequent interaction with PDGF-BB has demonstrated cytoplasmic KLF4 functions in cytoskeletal organisation. This shows that although the primary role of KLF4 is a nuclear transcription factor, it can also be localised to the cytoplasm to carry out a different function.

In cell-free methods of generating iPSCs, KLF4 and polyarginine fusion proteins have been shown to be localised to the nucleus with a small amount remaining in the cytoplasm. This expression pattern is also observed when KLF4-penetrating fusion proteins are expressed in mouse embryonic fibroblasts. These findings suggest a complex role for KLF4 in stem cell induction.

**C-MYC**

C-MYC is an oncprotein involved in regulating a wide range of processes including apoptosis, cell growth and division, angiogenesis and differentiation. Analyses of tissues from a range of human malignancies have shown that cytoplasmic accumulation of c-MYC is more prevalent than nuclear accumulation. In CRC, tumour progression is associated with an accumulation of cytoplasmic c-MYC, similar to that of prostate cancer cells. Ocular melanoma has been shown to express both nuclear and cytoplasmic c-MYC by IHC staining similar to that of breast cancer and testicular malignant teratomas. In pancreatic adenocarcinomas nuclear, cytoplasmic or dual localisation has been observed. These findings suggest that the localisation of c-MYC in cancer is mostly cytoplasmic, although there is potential for either cytoplasmic or nuclear c-MYC to contribute to tumour growth. C-MYC_S, an isoform formed by skipping of the N-terminal 15 amino acid residues at the N-terminus. As the two c-MYC NLSs are found close to the N-terminal end of the peptide chain, this c-MYC isoform could account for discrepancies in its subcellular localisation.

Recombinant c-MYC constructs used in cell-free reprogramming show mostly nuclear localisation with a small amount remaining in the cytoplasm, similar to that observed with other reprogramming factors, which contrasts with the primarily cytoplasmic localisation in tumour cells. Therefore, it is likely that the role of c-MYC in iPSCs is similar to its function in regular development rather than being specific to tumour growth and progression.

**CONCLUSION**

The CSC concept is supported by an increasing number of studies demonstrating the presence of subpopulations of CSCs expressing stem cell markers. However, the function of these markers in the context of their subcellular localisation is an unexplored area of research with potential for the development of novel treatment for a wide array of cancers. In this review, we outline some of the recent findings and postulate on possible mechanisms for the subcellular localisation of these markers within a variety of cancers and compare and contrast them with iPSCs. OCT4 localisation has been shown to be cytoplasmic in most tumour cells but is nuclear in iPSCs. This is supported by the localisation of SOX2 and NANOG, which are expressed in both the nucleus and cytoplasm in a range of cancers, suggesting a function dependent on their subcellular kinetics. KLF4 and c-MYC are both dispensable in the creation of iPSCs and have both shown nuclear localisation in cell-free methods of iPSC generation. However, they have also demonstrated both nuclear and cytoplasmic localisation in a range of cancers, consistent with findings of the other stem cell markers discussed in this review. These markers also act cooperatively to confer pluripotency characteristics. Therefore, one of the next steps for research into how their subcellular localisation affects cellular function could be to analyse the localisation of multiple markers simultaneously so as to determine whether the findings discussed in this review occur concurrently or independently.

This review demonstrates potential perspectives into the subcellular localisation dynamics of the stem cell markers OCT4, SOX2, NANOG, KLF4 and c-MYC in iPSCs, cancer and CSCs. However, there is clearly a lack of understanding in the literature as to the functional role of the cytoplasmic expression of these markers in relation to stem cells, and this will need to be addressed in future studies.

**Take home messages**

- Subcellular localisation of the stem cell markers octamer-binding transcription factor 4 (OCT4), sex-determining region Y-box 2 (SOX2), NANOG, Kruppel-like factor 4 (KLF4) and c-MYC in induced pluripotent stem cells and cancer stem cell may give clues to their roles in cancer development and growth.
- Many different factors contribute to the subcellular localisation of OCT4, SOX2, NANOG, KLF4 and c-MYC, including binding partners, protein isoforms and the presence or absence of nuclear localisation or nuclear export signals.
- OCT4, SOX2, NANOG, KLF4 and c-MYC exhibit both nuclear and cytoplasmic localisation across many different cell types; therefore, it may be the kinetics of protein shuttling that determine their effect on the cell.
REFERENCES