CHARACTERIZATION OF TUMOURIGENIC SUBPOPULATIONS IN NASOPHARYNGEAL CARCINOMA ISOLATED BY SIDE POPULATION AND CANCER STEM CELL MARKER APPROACHES

SUSAN HOE LING LING

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Susan Hoe Ling Ling** (I.C No: [redacted])

Matric No: **SHC120028**

Name of Degree: **Doctor of Philosophy**

Title of Thesis ("this Work"): **Characterization of Tumourigenic Subpopulations in Nasopharyngeal Carcinoma Isolated By Side Population And Cancer Stem Cell Marker Approaches**

Field of Study: **Biohealth Science**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

CHARACTERIZATION OF TUMOURIGENIC SUBPOPULATIONS IN NASOPHARYNGEAL CARCINOMA ISOLATED BY SIDE POPULATION AND CANCER STEM CELL MARKER APPROACHES

ABSTRACT

Nasopharyngeal carcinoma (NPC) is an epithelial malignancy of the nasopharynx. Studies on NPC tumourigenic subpopulations are often performed in NPC cell lines due to a lack of sizeable NPC tissues. Commonly used methods to identify the subpopulations include sphere-forming culture, side population (SP) assay and cancer stem cell (CSC) markers. The main objective of this study is to characterize the biological properties of tumourigenic cells in NPC isolated separately using SP assay and CSC markers from established NPC cell lines and/or early-passage NPC patient-derived xenografts (PDXs). SP assay identifies stem-like cells by their ability to extrude Hoechst 33342 dye. Using this first approach, SP cells were identified in HK1, a recurrent NPC cell line and in xeno-284, an early-passage NPC PDX established from recurrent metastatic NPC. HK1 contained 5 to 10% of SP cells while xeno-284 had less than 0.5% SP cells. HK1 SP cells significantly formed more holoclones than its non-SP (NSP) cells, an *in vitro* clone morphology closely related to self-renewal. SP cells also had higher aldehyde dehydrogenase (ALDH) activity than NSP cells, showed asymmetrical cell division and contained slow-proliferating cells. *ABCG2*, *SOX2*, *TERT*, *MYC* as well as certain Hedgehog, Notch, TGF-β and Wnt signalling pathway transcripts were significantly upregulated in SP cells. Despite significant differences seen *in vitro* and in gene expression experiments, SP and NSP cells showed an overall comparable tumour formation ability and tumour-initiating cell (TIC) frequency in nude mice. CSC markers are successfully used to identify tumourigenic stem-like cells in solid tumours. In the second approach, CD24, CD44, EpCAM and a combination of EpCAM/CD44 markers were used to isolate subpopulations of cells from C666-1, a cell line established from an undifferentiated NPC biopsy and from xeno-B110, an early-

passage NPC PDX established from a NPC biopsy. CD44br and EpCAMbr cells from C666-1 and xeno-B110 enriched for faster-growing tumourigenic cells with resulting larger tumours. Marked growth differences in xeno-B110 were associated with higher percentage of S-phase cells and mitotic figures in the marker bright groups as compared to marker dim groups. EpCAM/CD44dbr marker from xeno-B110 did not enhance for faster-growing cells or higher TIC frequency than CD44br marker alone. CD24br, CD44br and EpCAMbr cells from xeno-B110 were more enriched for TICs than their respective dim phenotypes in the first passage and retained self-renewal property upon serial transplantation *in vivo* for three successive passages. At the final passage of serial transplantation, CD24br cells were 10.55 folds more enriched with TICs than CD24dim cells, CD44br cells had a 7.07-fold enrichment over CD44dim cells and there was a 4.89-fold TIC enrichment for EpCAMbr cells over EpCAMdim cells. *KLF4* and *CDKN1A* transcripts were downregulated in all bright phenotypes which also had induced larger tumour growth, indicating a tumour suppressor role for KLF4 in NPC. Together, the study has shown that SP and CSC markers could isolate NPC cells with differential biological properties *in vitro*.

Keywords: nasopharyngeal carcinoma, biological properties, tumourigenic, side population, cancer stem cell markers

PENCIRIAN SEL-SEL TUMORIGENIK KARSINOMA NASOFARIN YANG DIISOLASI DENGAN KAEDAH "SIDE POPULATION" DAN PENANDA SEL STEM KANSER

ABSTRAK

Karsinoma nasofarin (NPC) merupakan kanser sel epitelium nasofarin. Kajian ke atas sel-sel tumorigenik NPC biasanya dilakukan dengan titisan sel NPC memandangkan kekurangan tisu NPC yang bersesuaian saiz. Kaedah-kaedah yang sering digunakan untuk mengkaji sel-sel tumorigenik termasuk kultur sfera, "side population" (SP) dan penanda-penanda sel stem kanser (CSC). Objektif utama tesis ini adalah mencirikan sifat-sifat biologi sel-sel tumorigenik NPC yang diisolasi masing-masing dengan kaedah SP dan penanda-penanda CSC daripada titisan sel NPC dan/atau NPC "patient-derived xenografts" (PDX) generasi awal. Kaedah SP mengenalpasti sel-sel bercirikan sel stem berupaya mengepam pewarna Hoechst 33342. Melalui kaedah pertama, sel-sel SP dikenalpasti dalam HK1, sejenis titisan sel NPC berulang selepas rawatan dan dalam xeno-284, sejenis NPC PDX generasi awal yang dibentuk daripada NPC berulang selepas rawatan. HK1 mengandungi 5 hingga 10% sel-sel SP manakala kurang daripada 0.5% sel-sel SP dalam xeno-284. Sel-sel SP HK1 membentuk lebih banyak "holoclone" daripada sel-sel bukan SP (NSP) ($p < 0.05$), suatu morfologi klon *in vitro* yang dikaitrapat dengan sifat pembaharuan diri yang tinggi. Sel-sel SP juga mempunyai aktiviti enzim aldehid dehidrogenasa (ALDH) yang lebih tinggi daripada sel-sel NSP. Ia menjalani pembahagian sel secara asimetrik dan tumbuh perlahan. Ekspresi transkrip *ABCG2*, *SOX2*, *TERT*, *MYC* serta sesetengah transkrip yang berfungsi dalam laluan isyarat Hedgehog, Notch, TGF-β dan Wnt ditingkatkan dengan signifikan dalam sel-sel SP. Walaupun terdapat perbezaan signifikan dalam *in vitro* dan di ekspresi gen, tiada perbezaan ketara di antara sel-sel SP dan NSP dalam keupayaan membentuk kanser dan frekuensi sel-sel "tumour-initiating" (TIC) pada tikus togel. Kaedah penanda CSC berjaya mengenalpasti sel-sel tumorigenik bercirikan sel stem

dalam kumpulan kanser pepejal. Penanda-penanda CD24, CD44, EpCAM dan kombinasi EpCAM/CD44 digunakan untuk mengasingkan sel-sel di C666-1, sejenis titisan sel daripada biopsi NPC dan di xeno-B110, sejenis NPC PDX generasi awal yang juga daripada biopsi NPC. Terdapat lebih banyak sel tumorigenik tumbuh cepat dalam sel-sel CD44br dan EpCAMbr C666-1 dan xeno-B110. Kedua-dua jenis sel ini turut menghasilkan ketumbuhan yang lebih besar. Perbezaan pertumbuhan yang lebih ketara dalam xeno-B110 disebabkan oleh peratus lebih tinggi kandungan sel-sel yang berada dalam fasa-S dan mitosis di kumpulan penanda terang berbanding dengan kumpulan penanda malap. Berbanding dengan sel-sel CD44br sahaja, sel-sel EpCAM/CD44 terang daripada xeno-B110 tidak lebih memperkayakan populasi sel yang tumbuh cepat ataupun meningkatkan frekuensi TICnya. Sel-sel CD24br, CD44br dan EpCAMbr lebih tinggi frekuensi TIC daripada sel-sel masing-masing dalam kitaran pertama dan juga menunjukkan sifat pembaharuan diri di dalam eksperimen "serial transplantation *in vivo*" sebanyak tiga kitaran berturut-turut. Di dalam kitaran "serial transplantation" terakhir, sel CD24br mengandungi frekuensi TIC 10.55 kali lebih banyak daripada sel CD24dim, sel CD44br 7.07 kali lebih banyak daripada sel CD44dim dan sel EpCAMbr sebanyak 4.89 kali lebih banyak daripada sel EpCAMdim. Sel-sel penanda terang menunjukkan tahap transkrip *KLF4* dan *CDKN1A* lebih rendah berbanding sel-sel penanda malap, yang turut menghasilkan ketumbuhan lebih besar di dalam tikus. Ini mencadangkan bahawa KLF4 berfungsi sebagai penidas kanser. Kesimpulannya, kaedah SP and penanda-penanda CSC berupaya mengisolasi sel NPC yang mempunyai sifat-sifat biologi yang berlainan *in vitro*.

ACKNOWLEDGEMENTS

This academic adventure would be impossible to begin without the generous research grant given and faith in me by the Research Management Committee of Institute for Medical Research, Kuala Lumpur. My appreciation also goes to Dr. Zubaidah Zakaria for helping me to get a FACSAria SORP in the institute.

My stepping back into the academic world at a grand old age of 42 would be unimaginable without the huge load of teaching/sharing, friendship, encouragement/support and aid from my numerous past and present lab mates of Molecular Pathology Unit, IMR, KL. My deepest gratitude especially goes to Dr. Tan Lu Ping (my co-supervisor and guiding light), Dr. Alan Khoo (for bringing me on board this study and sharing your insights), the wonderful and cheerful mates of SPF, and Chu Tai Lin, Norazlin Abd. Aziz, Tan Geok Wee and Dr. Fazlyn Reeny Abdul Razak (for your camaraderie and listening ear).

I first embarked on this study with zero knowledge of flow cytometry. The steep learning curve faced for flow techniques was only made achievable by the generous and patient guidance from Melinda Leong (BD Asia Pacific), and the team of engineers and application specialists from Biomarketing Services (M) Sdn. Bhd. My utmost gratefulness goes to my past and present flow lab mates for helping me to maintain the sorter and the flow lab.

I am most privileged to be able to work with and learn from my co-authors. Thank you immensely for your contribution and sharing of knowledge.

To my academic supervisor Assoc. Prof. Dr. Ng Ching-Ching, thank you for your guidance, support and kind reminders throughout my study.

Lastly, a big thanks to my dearest Pa, Mi, Nah, Boy, Hua Chan and Elaine for your unconditional love, emotional support and constant presence behind me in good and bad times during these years of academic struggle. You are my bedrock of support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTS OF SYMBOLS AND ABBREVIATIONS

DAB	:	3,3'-diaminobenzidine
HEPES	:	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
DAPI	:	4',6-diamidino-2-phenylindole
5-FU	:	5-fluorouracil
AT	:	adenine-thymine
ATP	:	adenosine triphosphate
ASR	:	age-standardized rate
ALDH	:	aldehyde dehydrogenase
APC	:	allophycocyanin
AJCC	:	American Joint Committee on Cancer
~	:	approximately
ABC	:	ATP-binding cassette
BAAA	:	BODIPY-aminoacetaldehyde
BAA	:	BODIPY-aminoacetate
BODIPY	:	boron-dipyrromethene
CSC	:	cancer stem cell
CRI	:	cancer-related inflammation
CO_2	:	carbon dioxide
cm	:	centimetre
CRT	:	chemo-radiotherapy
CTC	:	circulating tumour cell
R^2	:	coefficient of determination
CRC	:	colorectal cancer
cDNA	:	complementary DNA
CI	:	confidence interval
CT	:	cycle threshold
℃	:	degree Celcius
DNA	:	deoxyribonucleic acid
DMSO	:	dimethyl sulfoxide
dH_2O	:	distilled water
EBER	:	EBV-encoded small RNA
U	:	enzyme unit
EGF	:	epidermal growth factor
EGFR	:	epidermal growth factor receptor
EpCAM	:	epithelial cell adhesion molecule
ESA	:	epithelial surface antigen
EMT	:	epithelial-mesenchymal transition
EBNA	:	Epstein-Barr nuclear antigen
EBV	:	Epstein-Barr virus
et al.	:	et alia
ELDA	:	Extreme Limiting Dilution Analysis
FITC	:	fluorescein isothiocyanate
FTC	:	fumitremorgin C
GEMM	:	genetically engineered mouse model
gDNA	:	genomic DNA
Gy	:	Gray
GFP	:	green fluorescent protein
HSC	:	haematopoietic stem cell
H&E	:	haematoxylin and eosin

HBSS	:	Hanks' balanced salt solution
HNSCC	:	head and neck squamous cell carcinoma
HCC	:	hepatocellular carcinoma
HPF	:	high power field
HRP	:	horseradish peroxidase
h	:	hour
HLA	:	human leucocyte antigen
HMLE	:	human mammary epithelial cell
HPV	:	human papillomavirus
HA	:	hyaluronic acid
Ig	:	immunoglobulin
IgA EA	:	immunoglobulin A early antigen
IgA VCA	:	immunoglobulin A viral capsid antigen
IHC	:	immunohistochemistry
ISH	:	*in situ* hybridization
iPSC	:	induced pluripotent stem cell
IMRT	:	intensity-modulated radiotherapy
LMP1	:	latent membrane protein 1
LMP2A	:	latent membrane protein 2A
<	:	less than
LDA	:	limiting dilution analysis
mRNA	:	messenger RNA
μL	:	microlitre
μm	:	micrometre
μM	:	micromolar
miR	:	microRNA
mL	:	millilitre
mW	:	milliwatt
min	:	minute
MAI	:	mitotic activity index
DEAB	:	N,N-diethylaminobenzaldehyde
ng	:	nanogram
nm	:	nanometre
NPC	:	nasopharyngeal carcinoma
NSG	:	NOD-scid gamma
NOD/SCID	:	nonobese diabetic/severe combined immunodeficiency
NSCLC	:	non-small cell lung cancer
NSP	:	non-SP
OSCC	:	oral squamous cell carcinoma cell
PDX	:	patient-derived xenograft
PBS	:	phosphate-buffered saline
PE	:	phycoerythrin
PCR	:	polymerase chain reaction
PI	:	propidium iodide
qPCR	:	quantitative PCR
RT	:	radiotherapy
RG	:	reference gene
®	:	registered trade-mark
RT-qPCR	:	reverse transcription-qPCR
rpm	:	revolutions per minute
RNA	:	ribonucleic acid
s	:	second

STR	:	short tandem repeat
SP	:	side population
SPF	:	specific pathogen free
SD	:	standard deviation
SEM	:	standard error of the mean
i.e.	:	that is
x g	:	times gravity
±	:	tolerance or statistical margin of error of a quantity
™	:	trade mark
TGF-b1	:	transforming growth factor beta 1
TNM	:	tumour, node, metastasis
TIC	:	tumour-initiating cell
UV	:	ultraviolet
w/v	:	weight/volume
WHO	:	World Health Organization

LIST OF APPENDICES

CHAPTER 1: GENERAL INTRODUCTION

Cancers or tumours are complex tissues made up of multiple subpopulations of cells which interact with each other and the microenvironment they reside in (Hanahan & Weinberg, 2011; Welch, 2016). Within them are cells with heterogeneous phenotypes, functions, lineages and behaviour, to name a few. For example, there are tumour cells responsible for seeding and forming a tumour ("tumour-initiating cells" or "cancer stem cells"), cells which are fast-growing to drive tumour expansion ("proliferating cells"), slower-growing cells which evade current regimens of therapy ("quiescent cells") and cells with invasive and migratory abilities to form new growth at secondary sites ("metastatic cells"). To complicate matters, cancer cells are not "static" as cellular plasticity helps to change and shift the phenotypes and functions (ElShamy & Duhé, 2013; Hanahan & Weinberg, 2011; Vicente-Dueñas et al., 2009).

Nasopharyngeal carcinoma (NPC), a type of head and neck cancer, afflicts certain ethnicities in specific localities worldwide. It is endemic amongst the Cantonese people of Southern China and high in frequency in Northern Africa, the Inuits of Greenland, the Bidayuh tribe of Sarawak, East Malaysia and the Chinese descent in Malaysia (Torre et al., 2015; Zainal Ariffin & Nor Saleha, 2011). Advances in its therapy with the use of intensity-modulated radiotherapy (IMRT) and combination of chemo-radiotherapy (CRT) have delivered much improvement to NPC outcome. However, 10% of patients still develop recurrent NPC (Zhang et al., 2013b).

1.1 Research question

My research question is "What are the differences within the subpopulations of tumourigenic cells in NPC?" Understandably, NPC researchers have worked on delineating the differences between subsets of tumour cells in aspects ranging from tumour initiation, growth properties such as proliferation, cell cycle, migration and self-

renewal, genetics, chemo- and radioresistance to targeted therapy (Friedrich et al., 2003; Kong et al., 2010; Lin et al., 2014a; Lun et al., 2012; Ma et al., 2013; Wang et al., 2013b; Zhuang et al., 2013). Nevertheless, majority of these researches were performed using NPC cell lines which have been in passage for many generations, such as C666-1, CNE-1 and CNE-2 (both CNE lines contained partial HeLa genome and are suspected to originate from a single individual (Chan et al., 2008), HONE-1 and SUNE-1. It is known that repeated passaging of cell lines leads to growth changes, genetic drift and/or instability and ultimately, an undesired selection of abnormal cell clones which give rise to a different presentation of the malignancy than that in real life. Hence, the findings from studies utilizing NPC cell lines may present a certain degree of deviation from the actual manifestation in NPC patients.

1.2 Study objectives

The main study objective is to characterize the biological properties of tumourigenic cells isolated using two approaches from NPC cell lines and early-passage NPC patient-derived xenografts (PDXs).

The specific objectives are:

i. to identify and isolate putative tumourigenic cells using side population (SP) and cancer stem cell (CSC) marker approaches,
ii. to characterize the tumour initiation, growth, proliferation and self-renewal properties of selected isolated cells in *in vitro* and *in vivo* models, and
iii. to determine the differential gene expressions of these cells.

1.3 Study significance

i. Patient-derived xenografts (PDXs) as a model to study NPC tumourigenesis in this work present a closer resemblance to the actual malignancy than widely-used cell lines. Data from this study will provide

more relevant information on different subpopulations of NPC, their growth properties and gene expression profiles as compared to published data from NPC cell lines.

ii. The findings on self-renewal ability of CSC marker-selected subpopulations will provide a basis for the evaluation of different NPC cell clones and their potential to sustain growth for long term as well as to induce recurrent NPC in patients. This also provides evidence for designing more effective treatment programmes to eradicate NPC cells which have longer life span.

iii. The difference in gene expression profiles of tumourigenic NPC cells identified by different CSC markers will provide a better comprehension of the association of each marker and the corresponding deregulation of genes responsible for various aspects of tumourigenesis. With this knowledge, improved diagnosis and treatment plans can be put in place in the future.

1.4 Novelty of study

This study used early-passage NPC PDXs to elucidate biological differences, if any, amongst NPC tumourigenic cells, side-by-side with authenticated NPC cell lines C666-1 and HK1. Early-passage PDXs are advocated to be used as study models for tumour heterogeneity, CSCs and therapy-related studies (Clevers, 2011; Dodbiba et al., 2015; Julien et al., 2012; Visvader & Lindeman, 2012).

To the best of my knowledge, this study is the only NPC report which systematically estimated tumour-initiating cell (TIC) frequency in a limiting dilution manner (from large cell numbers to the smallest of 10 cells) and thoroughly verified self-renewal ability for three successive passages *in vivo* using marker-selected cells from an early-

passage NPC PDX. Existing NPC publications on CSCs primarily derive their findings from the use of NPC cell lines (authenticated and unauthenticated) and utilizing *in vitro* sphere or spheroid assay to test for self-renewal ability. The sphere or spheroid assay, often used as a surrogate experiment for *in vivo* tumourigenicity, is a subject of discussion regarding its suitability to determine true self-renewal ability of a TIC (Calvet et al., 2014; Pastrana et al., 2011; Weiswald et al., 2015).

Tumourigenicity is the mainstay of this study. In order to avoid under-reporting of tumour formation, growth observation for serial transplantation experiment was performed for 150 days per serial passage. Also, NOD-scid gamma (NSG) mice (strain NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/ SzJ) were used for the serial transplantation experiment, instead of less immunodeficient mouse models such as nude and NOD/SCID mice. As shown by Quintana et al. (2008) and Zhou et al. (2014), NSG mice allowed higher and faster engraftment of human tumourigenic cells due to no functional T and B cells in addition to an absence of natural killer cells.

1.5 Correlation of chapters

Chapter 2 presents an in-depth review of literature pertaining to NPC, tumourigenicity, SP and CSC marker approaches to study non-NPC malignancies. Chapter 3 consists of literature review on past NPC publications using SP approach, the methodology of the SP approach undertaken in this study to identify tumourigenic and non-tumourigenic cells in recurrent NPC (HK1 cell line and xeno-284 PDX), results and discussion. The work has been published in Cancer Cell International (Hoe et al., 2014). My second approach to identify and characterize tumourigenic and non-tumourigenic cells using widely-used CSC markers in primary NPC (C666-1 cell line and xeno-B110 PDX) is presented in Chapter 4. Existing literature on CSC marker-related studies in NPC is also included in the chapter. The work has been published in

Scientific Reports (Hoe et al., 2017). Chapter 5 contains a general discussion on both isolation approaches, limitations encountered in this study, proposed future direction and concluding remarks.

CHAPTER 2: LITERATURE REVIEW

2.1 Nasopharyngeal carcinoma

Nasopharyngeal carcinoma (NPC) is the most common type of nasopharyngeal tumour with distinctive geographical distribution (Chan et al., 2005). According to GLOBOCAN 2012, more than 90% of new NPC cases occur in less developed countries with the highest incidence rates in Southeast Asia especially in Malaysia, Indonesia and Singapore (Figure 2.1) (Torre et al., 2015). It is also high amongst Cantonese descent in Southeast China, North Africans from Tunisia and Algeria, and amongst the Inuits of Alaska, Greenland and North Canada (Chan et al., 2005; Torre et al., 2015).

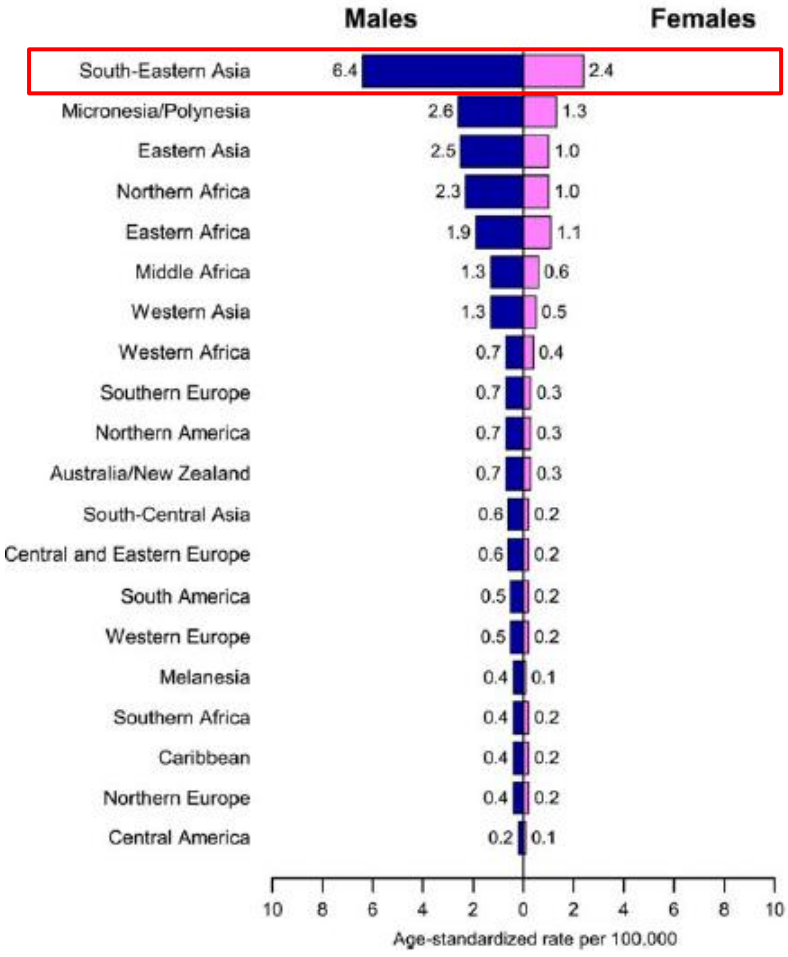


Figure 2.1: Age-standardized rate (per 100,000) of nasopharyngeal carcinoma grouped by geographical locations and sex (Torre et al., 2015).

According to the latest Malaysian cancer statistics collected from 2007 to 2011, NPC ranked the fifth most common cancer and the third highest in the overall population and amongst Malaysian males, respectively (Figure 2.2) (Azizah et al., 2016). It was also reported that NPC was the second most common cancer in Malaysian males of working age of 25 to 59 years. Further stratification revealed that local Chinese had the highest NPC incidence (age-standardized rate, ASR of 11.0 and 3.5 per 100,000 population for males and females, respectively), followed by the Malays (ASR of 3.3 and 1.3 for males and females, respectively) and the Indians (ASR of approximately 1 for each males and females). The ASR of Malaysian male and female Chinese was each third highest in the world, after China (cancer statistics from 2004 to 2007) and Singaporean Chinese (cancer statistics from 2003 to 2007) (Azizah et al., 2016; Forman et al., 2014). In the same Malaysian report, most of the local NPC patients were diagnosed at Stages III (29.3%) and IV (33.3%). An earlier study conducted in the state of Sarawak, East Malaysia with data collection from 1996 to 1998 reported that the risk of getting NPC was the highest in Bidayuh natives (2.3 and 1.9 folds higher than the average Sarawak males and females, respectively) with ASR of 31.5 (males) and 11.8 (females) (Devi et al., 2004).

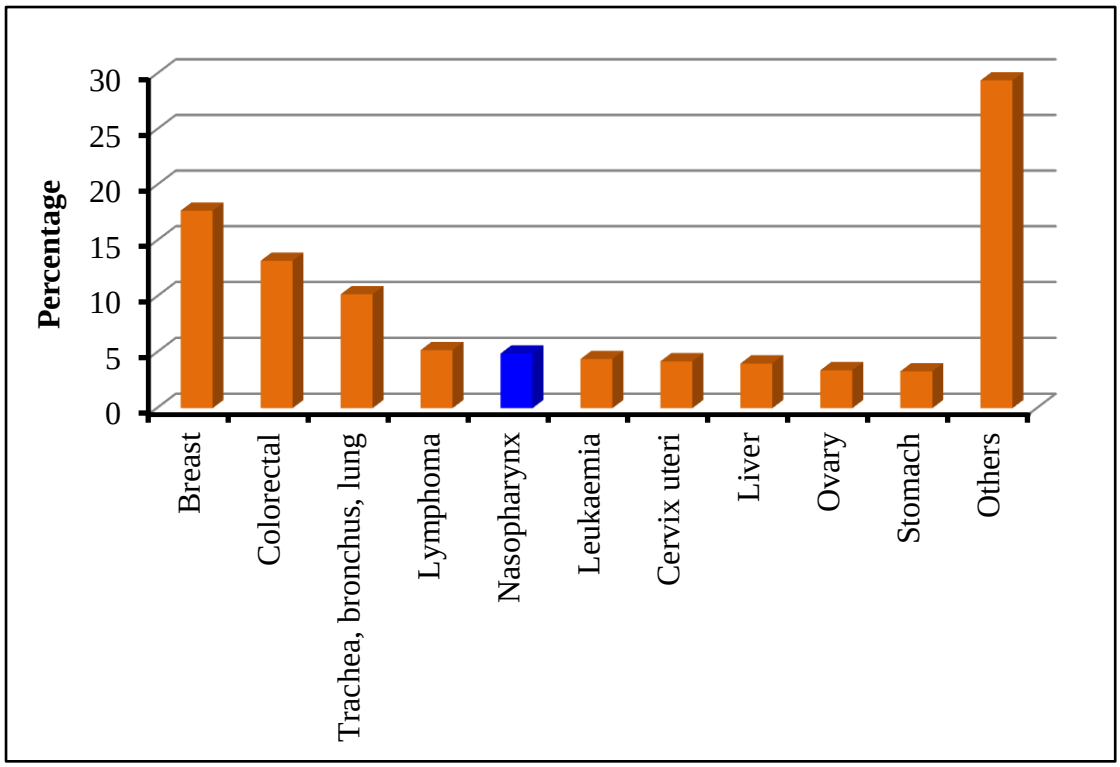


Figure 2.2: Ten most common cancers in Malaysia (2007-2011) (Manan et al., 2016).

2.1.1 Disease presentation, diagnosis and staging

Nasopharynx refers to the upper part of the throat or pharynx which is behind the post-nasal space (Figure 2.3). NPC usually presents at the fossa of Rosenműller of the nasopharynx with no definitive or specific symptoms which can range from neck lumps, blood-stained saliva, nose bleed to a blocked ear (Figure 2.4) (Khoo & Pua, 2013).

A biopsy of abnormal-looking nasopharynx is the usual first step towards a confirmed diagnosis of the malignancy (Chua et al., 2016; Khoo & Pua, 2013). This is followed by histopathological examination of the sample which categorizes it according to the 2005 WHO criteria: keratinizing squamous cell carcinoma, non-keratinizing squamous cell carcinoma (differentiated and undifferentiated), and basaloid squamous cell carcinoma (Chan et al., 2005).

Staging is then performed to assess the extent of the disease in order to determine its prognosis and choice of treatment modalities. Currently, this is done according to the American Joint Committee on Cancer (AJCC), 7th edition (Table 2.1) (Khoo & Pua, 2013).

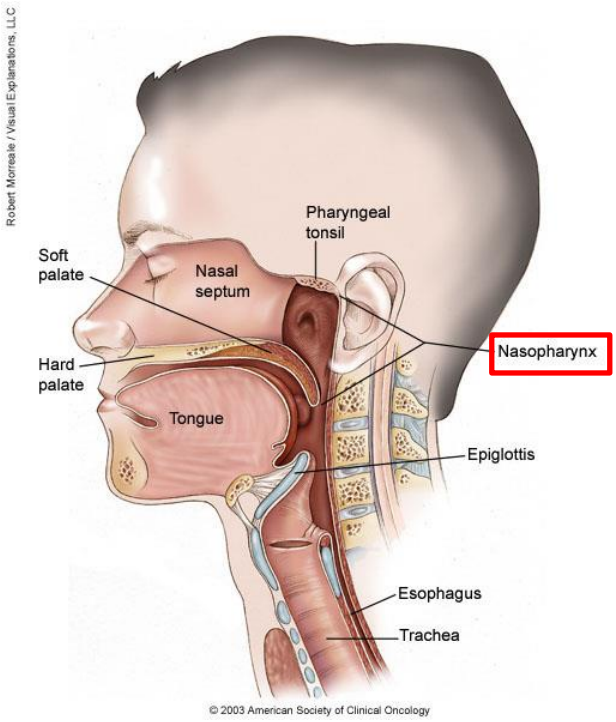


Figure 2.3: Location of nasopharynx in a drawing of head and neck region. (http://www.cancer.net/cancer-types/nasopharyngeal- cancer/medical-illustrations)

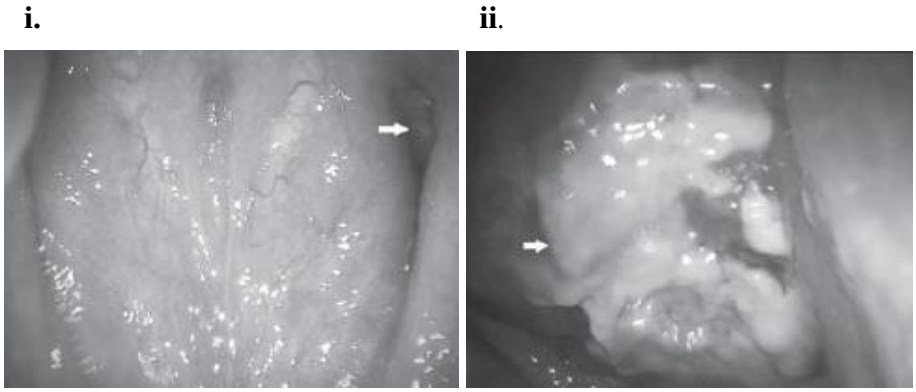


Figure 2.4: Endoscopy images of nasopharynx. **i.** Normal-looking nasopharynx with arrow pointing at fossa of Rosenműller. **ii.** Arrow points at a NPC growth. (Khoo & Pua, 2013)

Table 2.1: TNM classification for nasopharyngeal tumours according to the AJCC Staging, 7th Edition (Chan, 2010; Khoo & Pua, 2013).

Primary Tumour (T)	
T1 -	Tumour confined to nasopharynx, or extends to oropharynx and/or nasal cavity without parapharyngeal extension
T2 -	Tumour with parapharyngeal extension (posterolateral infiltration of tumour)
T3 -	Tumour involves bony structures and/or paranasal sinuses
T4 -	Tumour with intracranial extension and/or involvement of cranial nerves, hypopharynx, orbit, or with extension to the infratemporal fossa/masticator space
Regional Lymph Nodes (N)	
N0 -	No regional lymph node metastasis
N1 -	Unilateral metastasis in cervical lymph node(s), 6 cm or less in greatest dimension, above the supraclavicular fossa, and/or unilateral or bilateral, retropharyngeal lymph nodes, 6 cm or less, in greatest dimension
N2 -	Bilateral metastasis in cervical lymph node(s), 6 cm or less in greatest dimension, above the supraclavicular fossa
N3 -	Metastasis in a lymph node(s) greater than 6 cm and/or to supraclavicular fossa
N3a -	Greater than 6 cm in dimension
N3b -	Extension to the supraclavicular fossa
Distant Metastasis (M)	
M0 -	No distant metastasis
M1 -	Distant metastasis
Clinical Stage Groups (Anatomic Stage/Prognostic Groups)	
Stage I:	T1, N0, M0 (“early stage”)
Stage II:	T1, N1, M0; T2, N0, M0; T2, N1, M0 (“early stage”)
Stage III:	T1, N2, M0; T2, N2, M0; T3, N0, M0; T3, N2, M0 (“locally advanced”)
StageIVA:	T4, N0, M0; T4, N1, M0; T4, N2, M0 (“locally advanced”)
Stage IVB:	Any T, N3, M0 (“locally advanced”)
Stage IVC:	Any T, any N, M1 (“distant advanced”)

2.1.2 Aetiological factors

NPC can be attributed to three groups of causation factors.

2.1.2.1 Viral presence

As NPC is ubiquitously linked to the presence of the Epstein-Barr virus (EBV), serological testing of EBV antibody against EBV immunoglobulin A viral capsid antigen (IgA VCA) and early antigen (IgA EA) is widely used as a screening, diagnostic and/or disease monitoring tool (Chan, 2010; Gourzones et al., 2013; Lo et al., 2004). However, Chua and colleagues pointed that EBV presence may not be the sole driving force towards malignant epithelial cell transformation as the virus was undetected in biopsies from high-risk individuals for NPC (Chua et al., 2016). Another virus, the human papillomavirus (HPV), was also detected in a large study to evaluate its status in NPC from endemic South China cohorts and non-endemic US cohort (Lin et al., 2014b). Two important observations were noted in this study: there were less than 4% (3/86 cases) of EBV negative NPC with an absence of HPV in all 86 cases in the endemic cohort. Within the non-endemic cohort, approximately 91% (10/11 cases) of EBV negative NPC occurred in the whites, of which half had smoking habit and HPV-16 infection.

2.1.2.2 Dietary and lifestyle practices or occupational hazards

Besides EBV infection, a history of salted fish and other types of preserved food consumption containing N-nitrosamines (carcinogens released during food preservation) has been found to be a common risk factor in NPC patients (Feng, 2013). Contrarily, consumption of fresh fruits and/or vegetables was associated with lower risk of NPC (Jia et al., 2010; Jin et al., 2014; Wang et al., 2016a). In fact, foods with higher presence of antioxidants and phytochemicals such as dark green leafy vegetables, carrots and citrus fruits conferred stronger protective effects against NPC which may be

due to better regulation of cellular processes, reduction of oxidative DNA fragmentation and inhibition of EBV protein expression (Liu et al., 2012; Wang et al., 2016a). Other alleged factors with weak or inconsistent associations include consumption of herbal teas, slow-cooked soups and liquour, cigarette smoking, and occupational exposure to dust, chemical fumes, smoke and formaldehyde (Chan et al., 2005; Chua et al., 2016).

2.1.2.3 Genetic susceptibility

Tsao and colleagues reviewed that the incidence of NPC is 20 to 50 folds higher in endemic locations such as Southern China than in the West (Tsao et al., 2014). Additionally, there was increased risk in families with three or more relatives having NPC amongst a high-risk group of Cantonese populations in that region (Jia et al., 2004). Besides ethnicity, a susceptibility locus within the major histocompatibility complex region of chromosome 6p21, which encodes the human leucocyte antigen (HLA) class I genes, has also been named as a risk factor (Tsao et al., 2014). These genes are involved in encoding for proteins which identify and present foreign antigens, including EBV-encoded peptides, to the immune cells for elimination.

2.1.3 Prognosis

The most promising prognostic factor is the disease presenting stage (currently in use the AJCC 7th edition), whereby overall survival had dropped from 90% for Stage I to nearly 60% for Stage IVA-B in a large scale retrospective study with 2,687 consecutive NPC patients in Hong Kong (Lee et al., 2005).

Nonetheless in this genomic era, plasma cell-free EBV DNA is also being successfully utilized to screen for new NPC cases as well as in disease surveillance. For example, plasma EBV DNA concentration was found to be significantly much higher in pre-treatment recurrent NPC patients than those who were relapse-free (a median of 3,035 versus 1,202 copies/mL; $p < 0.05$) (Lin et al., 2004). In another study with 210

NPC patients (111 new patients and 99 previously reported patients) who were given induction CRT and followed for more than 6 years, a significant correlation was found between EBV DNA levels and relapse-free survival, overall survival and relapse rates in the group of new patients (Wang et al., 2013a). EBV DNA measurements at pre- and post-treatment stages were also significantly associated with poorer survival of NPC patients ($p < 0.001$) in a meta-analysis comprising of 10,732 NPC patients (Zhang et al., 2016).

2.1.4 Treatment modalities

2.1.4.1 Radiotherapy

NPC is highly radiosensitive with radiotherapy (RT) being its primary curative treatment, especially for primary NPC cases (Chan, 2010; Chua et al., 2016). IMRT is becoming a standard RT technique especially in larger well-equipped hospitals as it allows a more precise and highly controllable dose delivery of the radiation beams whilst ensuring minimal exposures to adjacent healthy tissues or organs, as compared to conventional two-dimensional RT (Lee et al., 2015; Zhang et al., 2013b). Despite studies showed more than 90% of locoregional control for T1 and T3 tumours, 5-year locoregional control rates for T4 tumours were below 80% (Lee et al., 2015; Zhang et al., 2013b).

2.1.4.2 Chemotherapy and RT

A combination of chemotherapy and RT (CRT) yielded better disease control and survival in patients with locally advanced NPC, with cisplatin (a platinum-based chemotherapy agent) being the drug of choice (Chua et al., 2016). As pointed by Lee and colleagues, NPC patients are usually diagnosed at advanced stages with more than 33% mortality within 5 years of diagnosis; hence, concurrent CRT with or without adjuvant chemotherapy has significantly improved on event-free survival versus RT

alone (5-year overall survival of 67 to 95% versus 37 to 86%, respectively), irrespective of chemotherapeutic regimens given (Lee et al., 2015).

As for recurrent and metastatic NPC cases, since NPC is also chemosensitive, first-line doublet chemotherapy (a combination of cisplatin and 5-fluorouracil (5-FU) or other agents) achieves 50 to 80% response rates, however the median survival only ranges from 12 to 20 months (Chan, 2010). Small recurrent NPC tumours (rT1 to rT2) can be also managed by surgery, brachytherapy or stereotactic radiosurgery (Chua et al., 2016).

2.1.4.3 Targeted therapies against EBV and its proteins

Despite advances achieved in the management of localized primary NPC, it is recognized that current modalities are below optimal levels for patients with advanced and recurrent NPC (as reviewed in subchapter 2.1.4.2), in addition to treatment-acquired complications and/or resistance. This leads to the search for more innovative, curative and less burdensome therapies.

A phase II clinical trial combining T-cell immunotherapy with standard chemotherapy was recently undertaken in patients with advanced EBV positive NPC with and without prior CRT (Chia et al., 2014). Gemcitabine (a third generation chemotherapy agent) and carboplatin (a second generation platinum-based agent) administration followed by up to six doses of EBV-specific T cells containing mixtures of latent membrane protein (LMP)1, LMP2 and immunodominant EBV antigens such as Epstein-Barr nuclear antigens (EBNAs) were given in four cycles ("GC-CTL" study). There were three complete and 22 partial responses, respectively, with the remaining 10 patients in stable condition amounting to an overall response rate of 71.4%. Median overall survival for the GC-CTL study was 29.9 months as compared to 17.7 and 21.4 months achieved by two independent chemotherapy trials at the same centre.

Another approach targets the LMP1 of EBV with the use of DNAzymes, i.e. synthetic single-stranded DNA catalysts which bind to their complementary sequence in target mRNAs (Lu et al., 2005). One such LMP1-binding DNAzyme is DZ1 which was evaluated in a clinical trial with 40 NPC patients receiving either DZ1 or saline, in conjuction with RT (Cao et al., 2014). At week 12, the mean tumour regression rate was significantly higher in DZ1-treated group (98%) than in saline control group (88%) and DZ1 group also had more samples with undetectable EBV DNA copy number than in the control group ($p < 0.001$). Importantly, there was no adverse effect observed due to DZ1 injections.

2.1.4.4 Targeted therapies against cancer stem cells

With the current keen interest in cancer stem cells (CSCs), there have been attempts to identify and target such cells in NPC. Briefly, CSCs are a subpopulation of cancer cells possessing self-renewal and differentiation abilities to continuously drive tumour growth. These cells will be reviewed in-depth in subchapter 2.2. A study was performed on NPC cell lines to target CSCs with nigericin, an antibiotic derived from *Streptomyces hygroscopicus* (Deng et al., 2013). Cell lines used included clones of CNE-2 with high and low metastatic abilities (S18 and S26, respectively), and HONE-1 and SUNE-1. SP cells having CSC characteristics were 2,500 folds higher in S18 cells than S26 cells. The inhibitory effect of nigericin such as toxicity, reduction of SP cells and lower tumour formation in nude mice was detected at higher level in S18 cells than S26 cells. Selective targeting of nigericin was also seen in HONE-1 and SUNE-1 cells. The mechanism of nigericin inhibition was related to the presence of polycomb group protein BMI1 (encoded by B cell-specific Moloney murine leukemia virus integration site 1 gene, *BMI1*), a marker for self-renewal, which was reduced markedly in S18 cells upon its treatment. Overexpression of BMI1 in the cells partially restored their SP content and metastatic ability.

A recent pre-clinical study in NPC used a specific inhibitor against a component of Wnt signalling pathway which is one of the most investigated pathways in cancer studies including CSCs. Chan and colleagues used ICG-001 inhibitor to block CREB-binding proteins (CBP)/b-catenin-mediated transcription of various stem cell genes and to enhance p300/b-catenin transcription which reduced CD44 and SOX2 expressions as well as sphere growth in C666-1 cell line (measurements of CD44 and SOX2 levels, and numbers of spheres were indicative of CSC presence in NPC) (Chan et al., 2015). The effects of ICG-001 alone, and combination of ICG-001 and cisplatin were also evaluated in nude mice bearing C666-1-cell induced tumours. Cisplatin alone ($p = 0.13$) and ICG-001 alone ($p = 0.16$) did not suppress tumour growth as compared to combined ICG-001 and cisplatin ($p = 0.02$). Mice treated with ICG-001 alone or combined with cisplatin had better health than mice with cisplatin only treatment.

Although there is much progress seen and excitement felt in development of targeted therapies for NPC, the severe lack of authentic NPC study models which can be used to demonstrate proof-of-principle of novel treatment approaches at the bench is a grave concern as illustrated by the above studies on CSC-targeted therapy and also reviewed by Lee and colleagues (Lee et al., 2015).

2.2 Tumourigenicity

2.2.1 Definitions

Tumourigenicity refers to the process of cells forming (benign or malignant) tumours consisting of viable and mitotically active cells in immunodeficient animal models (Bunz, 2008; Frandsen et al., 2001). It is an experimental characteristic reflecting the malignant nature of the original tumour from which the cells were derived (Bunz, 2008).

Tumour-initiating cells (TICs) are cells which are capable of initiating, inducing or forming progressively growing tumours in animals with compromised immunity such as nude, NOD/SCID and NSG mice, and need not be rare (Clarke et al., 2006; Kelly et al., 2007; Stewart et al., 2011; Valent et al., 2012).

According to the consensus reached in the American Association for Cancer Research (AACR) workshop on cancer stem cells in 2006, a *cancer stem cell* (CSC) is "a cell within a tumor that possess the capacity to self-renew and to cause the heterogeneous lineages of cancer cells that comprise the tumor." and needs to be defined experimentally by its "ability to recapitulate the generation of a continuously growing tumor". This in turn leads to putative CSCs being alternatively referred to as "TICs" and "tumourigenic cells" in literature (Clarke et al., 2006, p. 9340). The frequency of TICs or CSCs are usually measured in a xenograft transplantation experiment involving multiple doses of cell inoculation ("limiting dilution analysis", LDA) (Clevers, 2011; Valent et al., 2012).

2.2.2 Models of tumour initiation

Generally, there are two basic models which are accepted for tumour initiation, development, propagation and heterogeneity (Baccelli & Trumpp, 2012; Imrich et al., 2012; Visvader & Lindeman, 2008). The clonal or stochastic evolution model describes that most cancers originate from a single cell; genetic instability causes it to undergo uncontrolled growth resulting in the production of different clones of cells which do not organize into any hierarchy, and ultimately tumours are formed (Figure 2.5i) (Nowell, 1976; Vlashi & Pajonk, 2015). Conversely, the CSC model denotes that all cancer cells belong to a hierarchy similar to normal tissues, and only a small subset of these cells located at the apex of this organization ("CSCs") possess the ability to initiate tumours

indefinitely and to give rise to heterogeneous cell populations (Figure 2.5ii) (Shackleton et al., 2009; Valent et al., 2012; Vlashi & Pajonk, 2015).

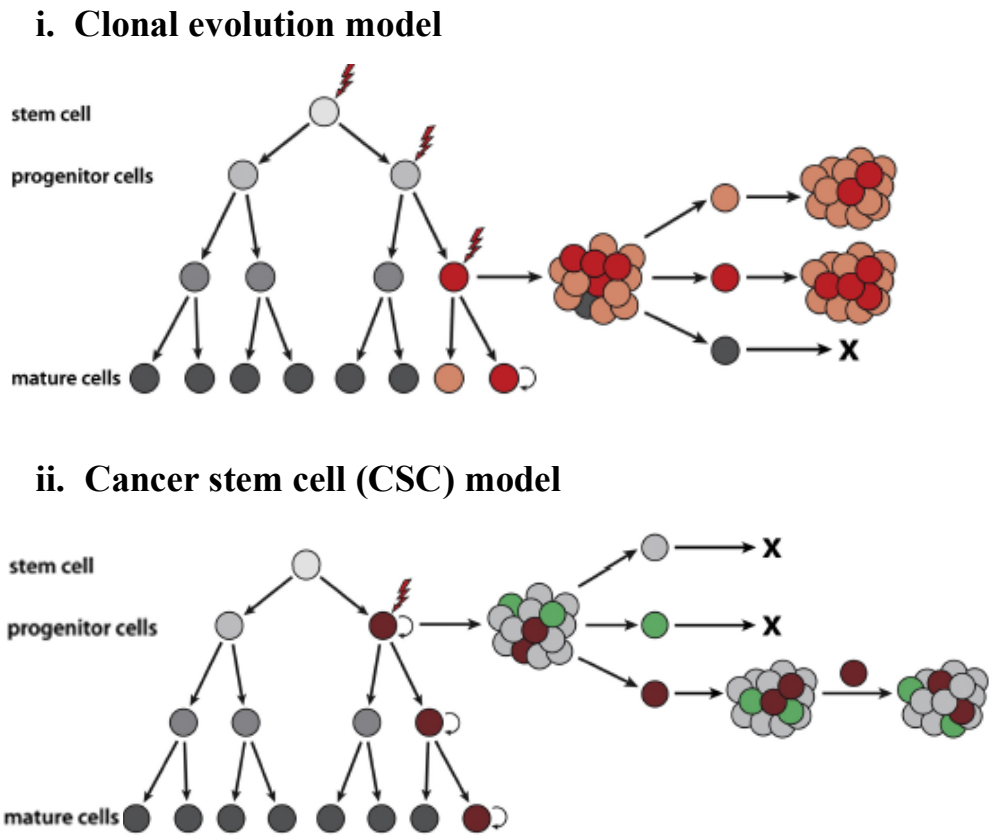


Figure 2.5: Models of tumour initiation. **i.** Clonal evolution model follows a non-hierarchical structure where growth advantage is given to any tumour cell with genetic instability. A dominant clone is produced from a cell (red) after numerous mutations. Tumour cells (red and orange) originating from this clone have similar tumourigenic potential. Other cells (grey and black) lack of ability to form tumours. **ii.** CSC model is based on a hierarchy where only a small subpopulation of cells (brown) at the apex has the ability to sustain tumour formation and generate heterogeneity through differentiation (grey and green cells). (Visvader & Lindeman, 2012)

2.2.3 Characteristics of tumourigenic cells

Table 2.2 summarizes the characteristics of tumourigenic cells in solid tumours identified by various methods.

Table 2.2: Characteristics of tumourigenic cells.

Characteristics	References
Initiate and drive tumourigenesis	Hanahan & Weinberg, 2011 Stewart et al., 2011
Abnormal cell division, growth and differentiation	Tysnes, 2010
Ability to recapitulate histology of original tumour as well as to give rise to both subpopulations of tumourigenic and non-tumourigenic cells	Tysnes, 2010
Perpetual growth ability in certain subsets of cells	Hanahan & Weinberg, 2011 Stewart et al., 2011 Xu et al., 2015
Ability to evade immune surveillance and make use of immune cells to support own growth	Krampitz et al., 2016
A certain degree of resistance to radiation and chemotherapy	Chang, 2016
Genetically altered and/or with enhanced expression of stemness-related genes and transcription factors	Chan et al., 2009 Hanahan & Weinberg, 2011
May or may not be rare depending on type of malignancy	Ishizawa et al., 2010 Kelly et al., 2007

2.2.4 Study models of tumourigenicity in NPC

As reviewed earlier, NPC is a highly radiosensitive malignancy with either RT alone or combined CRT as its main treatment modalities (Chan, 2010; Xu et al., 2016a). Hence, NPC patients are not usually subjected to surgery as a treatment modality and this leads to a near dearth of fresh NPC tumour samples for lab-based studies. If there

are, limited tissue quantity is inherent as these samples are usually from post-nasal space biopsies (Chan et al., 2008; Gullo et al., 2008). As a result, most NPC investigations into various aspects of its tumourigenesis are performed with NPC cell lines which were established years ago (Feng et al., 2013; Lun et al., 2012; Tsang et al., 2013; Xu et al., 2016b; Yang et al., 2013).

As EBV is found in almost all NPC cases except for keratinizing NPC from non-endemic areas, the most frequently used and sole authenticated EBV positive NPC cell line to date is C666-1 which was derived from an undifferentiated NPC xenograft (Chan et al., 2008; Hui et al., 1998). Other common NPC cell lines in use include HK1 (from recurrent well-differentiated NPC sample; EBV negative; authenticated), HONE-1 (poorly differentiated NPC; *in vitro* EBV loss; suspicious identity), CNE-1 (well-differentiated squamous NPC; EBV negative; suspicious identity) and CNE-2 (poorly differentiated NPC; EBV negative; suspicious identity) (Chan et al., 2008; Glaser et al., 1989; Gullo et al., 2008; Huang et al., 1980; Strong et al., 2014).

Besides sample size limitation, NPC primary cells are also known to be difficult to propagate *in vitro* for long term (Gullo et al., 2008). This led to the establishment of NPC PDXs from NPC patient samples such as C15 and C17 (Busson et al., 1988); however, they are used at a lower frequency than the above NPC cell lines at preclinical investigations mainly due to technical difficulties.

2.3 Side population assay

One of the methods to identify tumourigenic cells and putative CSCs is side population (SP) assay. It utilizes vital fluorescent dyes such as Rhodamine 123 and Hoechst 33342 to segregate cells with dye-pumping ability from those which do not. Rhodamine 123 binds to mitochondrial membrane and is also a substrate for ABCB1 drug transporter (ATP-binding cassette subfamily B1) (Bertoncello & Williams, 2004).

Hoechst 33342, on the other hand, penetrates the intact plasma membrane of live cells to bind to DNA at the AT-rich regions of its minor groove, is the preferred substrate for ABCG2 drug transporter and more commonly used for CSC studies (Wu & Alman, 2008; Golebiewska et al., 2011). Actively metabolizing cells which contain ABC transporters are able to efflux the dye, leaving these cells to be lightly-stained with Rhodamine 123 or Hoechst 33342; such cells are referred to as "SP cells".

When the Hoechst 33342-stained cells are analysed in a flow cytometer, a hummingbird profile can be seen on a Hoechst Blue versus Hoechst Red diagram where SP cells are usually positioned at the tip of "its beak" due to their low Hoechst 33342 staining intensity (Figure 2.6). The location of SP cells ("SP gate") is then confirmed with the use of ABC transporter inhibitors such as verapamil, fumitremorgin C (FTC) and reserpine (Golebiewska et al., 2011).

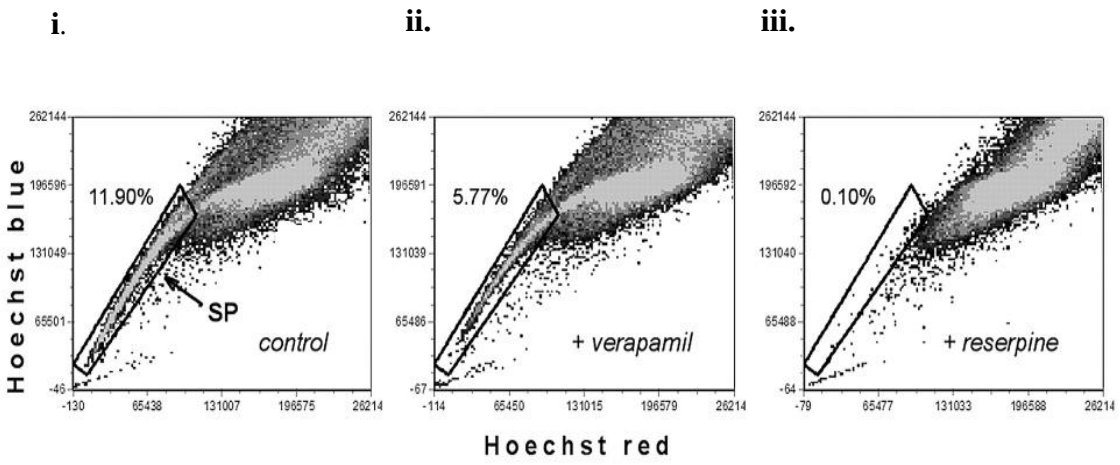


Figure 2.6: Flow cytometry dot plots showing profiles of SP cells stained with Hoechst 33342 in the presence and absence of verapamil or reserpine inhibitor. **i.** SP cells are usually positioned at the (tip of) the hummingbird's beak (without inhibitor). **ii.** Verapamil is not the optimal ABC transporter inhibitor for the cells as about half of total SP cell population was still present. **iii.** Proper setting of the SP gate is confirmed with the use of appropriate ABC transporter inhibitor such as reserpine in this experiment. (Jakubikova et al., 2011)

Historically, this assay was first used by Goodell et al. (1996) who found that SP cells isolated from murine bone marrow contained haematopoietic stem cell (HSC) marker Sca-1 positive cells, they were able to repopulate *in vivo* and were enriched in HSC activity more than 1,000 times. Post-transplantation, the SP cells conferred protection to lethally irradiated mice.

There has been a plethora of SP investigations in cancers on their biological features and functions ever since. For example, SP cells from a renal cell carcinoma cell line were able to differentiate into non-SP (NSP) cells *in vitro*, showed resistance to irradiation and 5-FU, and formed tumours in NOD/SCID mice more efficiently than NSP cells at 2,000-cell inoculation, (Huang et al., 2013). Although both phenotypes were able to form secondary xenografts, NSP cells showed growth disadvantage with smaller tumours.

Increased levels of β-catenin protein and *CCND1* gene, both members of the Wnt/β-catenin pathway, and increased expressions of stem cell markers such as *Nestin*, *OCT4*, *SOX2* and *NANOG* in osteosarcoma samples were found to contribute to self-renewal (assayed by sphere formation) and tumourigenicity potential in SP cells (Yi et al., 2015).

SP cells were both detected in melanoma patient samples and also PDXs (Luo et al., 2012). Characterization of SP and NSP cells from the xenografts found that SP cells showed resistance to paclitaxel and temozolomide (chemotherapeutic drugs for melanoma patients) which were partially attributed to the overexpression of ABC transporter *ABCB1* and *ABCB5* genes in these cells. Also, pathway analysis of their microarray data showed that NF-kB, α-β4-integrin and IL-1 pathways were significantly deregulated in SP as compared to NSP cells. All three pathways had been previously shown to be involved in apoptosis resistance after suffering DNA damage

(Arlt et al., 2002; Weaver et al., 2002; Yamamoto & Gaynor, 2001). Nonetheless, the melanoma SP study did not proceed to verify tumour-initiating and self-renewal abilities of SP and NSP cells *in vivo* and thus, the SP cells could only be deemed as "stem-like".

On the other hand, Nakayama et al. (2014) reported that SP cells from a hepatocarcinoma cell line did not conform to the CSC identity although they exhibited faster proliferation *in vitro*, formed significantly more spheroids and induced tumour formation at least 2.5 folds more than NSP cells. The conclusion was derived from the inabilities of SP cells to out-perform NSP cells in drug resistance, colony-forming and cell cycle experiments. Additionally, *in vivo* self-renewal ability, the gold standard to classify CSCs, was not tested.

Literature review of the identification of SP cells and elucidation of its roles in NPC will be covered in Chapter 3.

The ability to pump out fluorescent dyes such as Hoechst 33342 and Rhodamine-123 is associated with the presence of ABC transporter(s) expressed in the cell which uses energy released from the hydrolysis of ATP to adenosine diphosphate to transport its substrate against a concentration gradient across the membrane (Kathawala et al., 2015). To date, there are seven known subfamilies of ABC transporters (ABCA to ABCG) with at least 48 human ABC genes (Dean et al., 2001). ABCB1/MDR1/P-glycoprotein is one of the three major ABC transporters and the first one to be identified for overexpression in multidrug resistance tumour cell lines (Dean, 2009). Cells overexpressing ABCB1 protein are resistant to chemotherapy agents such as doxorubicin, vinblastine and paclitaxel (Dean et al., 2001). Likewise, overexpression of ABCC1/MRP1 transporter renders resistance to many anticancer drugs; however, it

does not confer resistance to taxanes (e.g. paclitaxel and docetaxel) (Kathawala et al., 2015).

ABCG2, the second member of G subfamily of ABC transporters and the ABC transporter of interest in this study, transports many types of substrates including chemotherapy drugs such as mitoxantrone, flavopiridol, topotecan, gefitinib, imatinib and erlotinib (Ding et al., 2010). Its cDNA was highly expressed in stem cells isolated from murine bone marrow, spleen, thymus as well as rhesus monkey bone marrow and SP phenotype was directly linked to its presence (Zhou et al., 2001). However, in a triple knockout of *Bcrp1/Mdr1a/1b* (presently known as *ABCG2/ABC1A/1B*) mouse model, there was still a presence of $0.05 \pm 0.08\%$ SP cells ($p < 0.01$) (Jonker et al., 2005). In a large panel of 150 untreated human carcinomas compromising 21 tumour types, ABCG2 protein was frequently detected in all tumour types using immunohistochemistry technique (Diestra et al., 2002). Both membraneous and cytoplasmic staining was seen in ABCG2 positive tumours; however, there was no associated disease staging or clinical outcome analysis performed in the study. A recent study in pancreatic cancer found that ABCG2 positive cells were co-expressed with gastrin, an autocrine factor linked to cell proliferation in some cancers (Wang et al., 2016b). By activating NF-κB pathway, gastrin upregulated ABCG2 expression and subsequently an increase of SP cells. Furthermore, wound healing experiments revealed that migration of pancreatic cancer cell line BxPC-3 stimulated with recombinant gastrin (rhGastrin) was more enhanced than unstimulated BxPC-3 or BxPC-3 with rhGastrin and ABCG2 knockdown.

The type of ABC transporter inhibitor to be used for a particular sample in SP assay depends on the identity of ABC transporter family member present. Golebiewska et al. (2011) reviewed that verapamil, cyclosporine A and probenecid individually inhibits

dye efflux mechanism of cells containing ABCB1 transporter, FTC specifically blocks ABCG2 transporter activity, while imatinib, an inhibitor of ABCB1 and ABCC1, may also hinder ABCG2 activity.

SP assay is not without its share of criticisms. Firstly, DNA-binding dyes such as Hoechst 33342 are known to be toxic to living cells at high concentrations and especially upon exposure to UV light (Golebiewska et al., 2011). Most of the reviewed SP studies above had used 5 or 10 μM of Hoechst 33342 which conferred low toxicity to cells at the usual incubation period of 90 min (Fried et al., 1982). Nonetheless, different cell types responded dissimilarly to Hoechst 33342 as Fried *et al.* found out in a comparative study using HeLa S-3, a human cervical cancer cell line; SK-DHLZ, a human lymphoma cell line; and normal human bone marrow cells. The concern of toxicity suffered by NSP cells due to an accumulation of Hoechst 33342 which would lead to subsequent poor viability and experimental bias in these cells was unfounded. Evidences presented by Huang et al. (2013), Nakayama et al. (2014) and Yi et al. (2015) showed that NSP cells were viable post-sorting and were able to proliferate, to form spheroids and *in vivo* tumours amongst other abilities, albeit at a lower scale than SP cells. The differences in the experimental data could be attributed to biology rather than viability.

Secondly, there is considerable variability in reporting the percentage of SP cells in publications using the same sample type or cell line. For example, breast carcinoma cell line, MCF7 stained with 5 μM of Hoechst 33342 for 90 min at 37 ºC in three independent studies gave a range of 0.39 to 7.5% of SP cells (Engelmann et al., 2008; Han & Crowe, 2009; Nakanishi et al., 2010). One of the plausible reasons leading to the data inconsistency is the lack of a clear reporting of step-by-step gating strategy to

enable cross-publication comparisons. This could be rectified by standardizing the reporting of data in SP assay as reviewed by Golebiewska et al. (2011).

2.4 CSC markers

Solid tumours are diagnosed clinically by histopathology with or without the expression of disease or cell type specific markers such as EBV-encoded small RNAs (EBERs) for NPC (Gourzones et al., 2013), and cytokeratin, estrogen receptor, progesteron receptor and HER2neu for triple-negative breast carcinoma (Dent et al., 2007). In addition to tumour subtypes, heterogeneity in cell proliferation, differentiation, tumourigenic potential and regeneration, phenotype and response to therapy is also often seen in tumour cells (Visvader & Lindeman, 2008, 2012). As CSC model is one of the models used to explain for tumourigenesis (reviewed in subchapter 2.2.2), surface markers associated with CSCs ("CSC markers") have been acknowledged to play an essential role in distinguishing tumourigenic from non-tumourigenic cells (Shackleton et al., 2009).

Table 2.3 shows a selection of seminal and recent publications which used CSC markers to isolate tumour subpopulations for studies on biological differences in cancers. The list of markers is not an exhaustive one of all published reports. Other not frequently in use markers include CD29, CD90, CD117 and CXCR4 (reviewed in Medema, 2013; Wang et al., 2015).

Following subchapters are in-depth reviews of CD24, CD44, EpCAM and a combination of dual or triple of these markers as identification markers for CSCs and/or TICs as well as their functions or roles in tumourigenesis.

Table 2.3: Isolation of tumour subpopulations from patient or cell line samples using CSC markers.

Marker	Malignancy	Combinatory marker(s)	*In vivo* serial transplantation	References
CD24	breast cancer	CD44; CD44 and ESA	yes (four cycles)	Al-Hajj et al., 2003
	pancreatic cancer	CD44 and ESA	yes (four cycles)	Li et al., 2007
	hepatocellular carcinoma (HCC)	none	yes (one cycle)	Lee et al., 2011
	NPC	none	no	Yang et al., 2014b
	oral cancer	CD44	no	Ghuwalewala et al., 2016
CD34	acute myeloid leukaemia	CD38	no	Lapidot et al., 1994
CD44	colorectal cancer (CRC)	ESA; CD166; ESA and CD166	no	Dalerba et al., 2007
	head and neck squamous cell carcinoma (HNSCC)	none	yes (two to three cycles)	Prince et al., 2007
		none	no	Oh et al., 2013
	ovarian cancer	CD117	no	Chen et al., 2013
	gastric cancer	none	no	Yu et al., 2014

Table 2.3, continued

Marker	Malignancy	Combinatory marker(s)	*In vivo* serial transplantation	References
CD44	NPC	none	no	Janisiewicz et al., 2012 Lun et al., 2012 Yang et al., 2014a
	liver cancer	none	no	Park et al., 2016
CD133	brain cancer	none	no	Singh et al., 2003
	bone cancer	none	no	Tirino et al., 2008
	NPC	none	no	Zhuang et al., 2013
CD166	non-small cell lung cancer (NSCLC)	none	yes (two cycles)	Zhang et al., 2012b
ABCG2	NPC	none	no	Zhang et al., 2012a
	liver cancer	none	no	Zhang et al., 2013a
	breast cancer	none	no	Leccia et al., 2014
ALDH (aldehyde dehydrogenase)	breast cancer	none	no	Charafe-Jauffret et al., 2009
	prostate cancer	CD44	no	Yu et al., 2011

Table 2.3, continued

Marker	Malignancy	Combinatory marker(s)	*In vivo* serial transplantation	References
EpCAM (epithelial cell adhesion molecule)	gastric cancer	CD44	yes (three cycles)	Han et al., 2011
	liver cancer	none	no	Yamashita et al., 2013
	NSCLC	CD166	no	Norashikin et al., 2015
	mouse breast cancer	none	no	Hiraga et al., 2016

2.4.1 CD24

CD24 is a surface protein made up of 32 amino acids and heavily glycosylated with a myriad of functions in T-cell proliferation and differentiation, B-cell development and autoimmune diseases (Tan et al., 2016). P-selectin, an endothelial cell adhesion molecule, is expressed under inflammatory conditions which occur during tumour metastasis (McEver et al., 1989). As CD24 acts as an alternate ligand for P-selectin (Aigner et al., 1998), CD24/P-selectin binding facilitates the passage of tumour cells in the bloodstream during metastasis, hence CD24 is also referred to as a marker for metastasis in non-small cell lung carcinoma (Lee et al., 2010). Additionally, CD24 increases proliferation and adhesion of tumour cells to fibronectin, collagen types I and IV, and laminin (Baumann et al., 2005).

The absence or low expression of CD24 is synonymous with identifying breast CSCs as was first highlighted by Al-Hajj et al. (2003). Limiting dilution experiments were performed in NOD/SCID mice with several combinations of CSC markers on breast cancer patients and PDXs. With a combination of $CD44^{+}CD24^{-/low}$ phenotype, patient

cells induced tumour formation *in vivo* with the lowest inoculation of 5,000 cells (1/1) as compared to no growth with $CD44^{+}CD24^{+}$ phenotype (0/2). TIC frequency for the $CD44^{+}CD24^{-/low}$ cells was 159 folds higher than $CD44^{+}CD24^{+}$ cells (1 in 4,985 as compared to 1 in 795,024, respectively). PDX cells selected for $CD44^{+}CD24^{-}$ phenotype had increased tumourigenicity potential (10/10 with 1,000 cells) as compared to $CD44^{+}CD24^{+}$ (0/10 with 1,000 cells) and successfully formed tumours with 10,000 cells at second passage. $CD44^{+}CD24^{-/low}$ cells were also successfully serial-passaged for four cycles in NOD/SCID mice with similar tumourigenicity. They also found that differences in growth for the cells were not due to differences in cell cycle as both tumourigenic and non-tumourigenic cells had similar cell cycle distribution. The $CD44^{+}CD24^{-/low}$ cells displayed typical CSC features as they were able to form tumours with enhanced TIC frequency, to self-renew *in vivo* and to give rise to phenotypically diverse cells.

The phenotype of $CD44^{high}CD24^{low}$ was also recently examined in OSCC (Ghuwalewala et al., 2016). In comparison to $CD44^{high}CD24^{high}$ and $CD44^{low}CD24^{high}$ cells, $CD24^{low}$ cells displayed the usual characteristics associated with CSC phenotype. Additionally, the authors found that $CD44^{high}CD24^{low}$ cells underwent epithelial-mesenchymal transition (EMT) with a lower E-cadherin (marker for epithelial cells) to vimentin (marker for mesenchymal cells) ratio seen in these cells, contained 1.5-fold more SP cells and had an enhanced ability to resist 5-FU treatment than $CD44^{low}CD24^{high}$ cells. However, the functionality of these CSCs was not evaluated in *in vivo* studies.

On the other hand, the presence of CD24 expression ($CD24^{+}$ phenotype) was more associated with tumour-initiating and stemness than $CD24^{-}$ phenotype in chemoresistant HCC PDXs (Lee et al., 2011). *CD24* mRNA in cisplatin-treated PDXs was 2.9 folds

higher than untreated PDXs. *In vivo* growth investigations revealed that $CD24^+$ cells from HCC cell lines had higher tumour incidence and TIC frequency (1 in 2,367 cells) but grew at a slower rate *in vitro*, were able to self-renew *in vivo* more efficiently (96.7% tumour formation) and to metastasize more efficiently than $CD24^-$ cells (TIC frequency of 1 in 69,669 cells). By using knockdown and overexpression experiments, it was found that these traits were regulated by *NANOG*, a self-renewal marker, upon the binding of phosphorylated STAT3 to its promoter.

Taken together, the CD24 phenotype ($CD24^+$ or $CD24^{-/low}$) responsible for tumourigenicity and/or stemness is tumour type-dependent.

Literature review of CD24 in NPC studies will be covered in Chapter 4.

2.4.2 CD44

CD44 is a surface protein with complex structure and functions. A full length *CD44* gene is made up of 20 exons and 19 introns: 10 exons are found in all isoforms ("constant" exons) with the remaining ones ("variable" exons) present in various combinations due to alternative splicing in the membrane-proximal stem region (Figure 2.7) (Yan et al., 2015). CD44s is the smallest, standard isoform or variant lacking all variable exons and is expressed on most vertebrate cells. CD44v isoforms are only expressed on some epithelial cells, during certain conditions such as embryonic development and lymphocyte activation, and in some cancers (Zöller, 2011). Most of the investigations on CD44 as a CSC or TIC marker used panCD44 antibodies which recognize both CD44s and CD44v isoforms (Zöller, 2011). Likewise, my study used a CD44 antibody (clone G44-26) which was regularly reported in flow cytometry-based CSC studies, and recognizes both CD44s and CD44v (Biddle et al., 2013). CD44 protein is a receptor for hyaluronan (hyaluronic acid, HA) which is a major component of the extracellular cell matrix, as well as a co-receptor for growth factors and

cytokines. Its binding with HA aids in cell-cell and cell-matrix interactions which further defines its many roles in cancer cell dissemination processes such as cell migration and survival signal transmission (Naor et al., 2009).

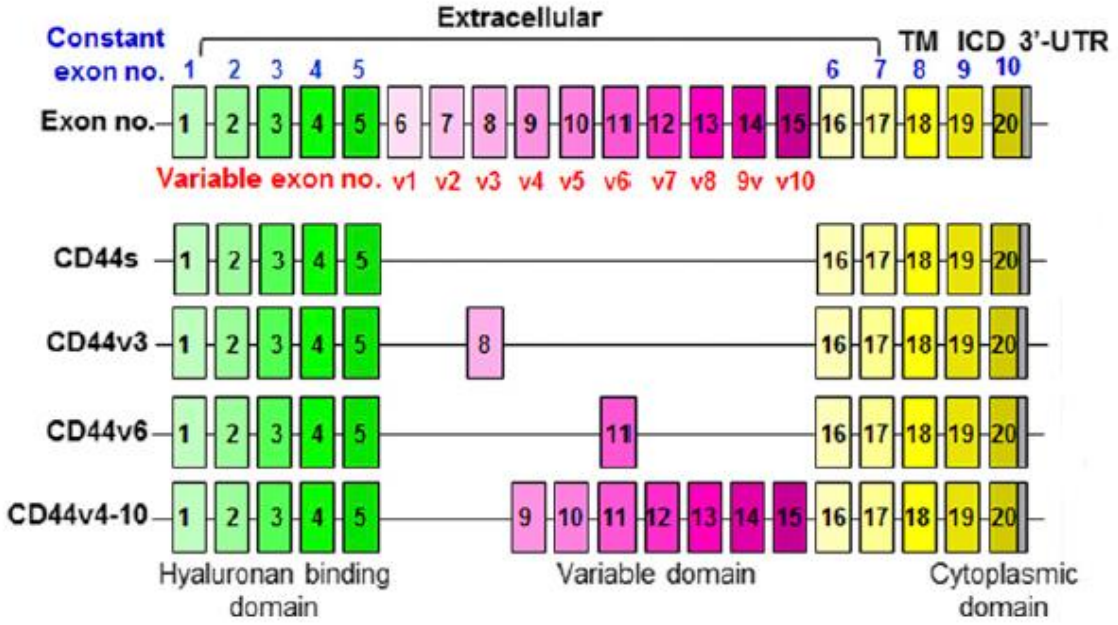


Figure 2.7: Diagrammatic structure of *CD44* gene. CD44 comprises of exons which are constant in every CD44 mRNA and protein ("constant exons", green and yellow), with some other exons which are variably found in CD44 splicing isoform (CD44v) mRNAs and proteins ("variable exons", pink). The standard CD44 (CD44s) does not contain any variable exon. ICD, intracellular cytoplasmic domain; TM, transmembrane region; UTR, untranslated region. (Yan et al., 2015)

Zöller (2011) wrote that CD44 is a major receptor for hyaluronan (HA) and the binding of HA to CD44 had important implications for TIC homing, migration and adhesion processes. CD44 is also associated with genes controlling stem cell characteristics via Wnt pathway, of which CD44 is a transcriptional target. Highly conserved Wnt pathway is involved in embryonic development and tissue homeostasis, which causes tumour formation when deregulated (Wend et al., 2010).

Prince et al. (2007) found that $CD44^+$ phenotype from patient and xenograft cells from HNSCC fulfilled the defining characteristics of CSCs, namely the abilities to self-renew and to differentiate. $CD44^+$ cells isolated from one patient induced tumour formation in mice from as low as 5,000 cells (3/3; overall TIC frequency 1 in 1) as compared to no growth with more than 40,000 $CD44^-$ cells (0/9). Moreover, $CD44^+$-induced tumours recapitulated the original tumour histology. Serial passaging was successfully performed for two to three cycles with $CD44^+$ but not $CD44^-$ cells. *BMI1*, a self-renewal gene, had a more than 4-fold upregulation in $CD44^+$ cells but was barely detected in its counterpart.

CD44 tumour initiation advantage was also shown in a combination study with EpCAM marker in CRC (Dalerba et al., 2007). Tumours formed with 5,000 and fewer $EpCAM^{high}CD44^+$ cells (TIC frequency 1 in 492 cells), whereas there was no growth from 10,000 $EpCAM^{low}CD44^-$ cells (TIC frequency 1 in 205,998 cells). The $EpCAM^{high}CD44^+$ phenotype showed asymmetric division ability by giving rise to similar proportions of $EpCAM^{high}CD44^+$ and $EpCAM^{low}CD44^-$ to those of parental growths.

Pathway-related studies showed that biological characteristics of $CD44^{+/high}$ cells may be effected by AKT, Hedgehog and TGF-β1 pathways, amongst others. $CD44^{high}$ in combination with $CD133^{high}$ phenotype selected for CRC cells which were more protective against radiation at all studied doses from 2 to 6 Gy, and AKT isoforms 1, 2 or both increased the expression of CD44 (Sahlberg et al., 2014). Both isoforms belong to the AKT pathway which is partially responsible for conferring anti-apoptosis and radiation resistance abilities. On the other hand, the Hedgehog pathway was instrumental in causing chemotherapy resistance of $CD44^+$ cells in gastric cancer (Yoon et al., 2014). CD44 showed a synergistic relationship with TGF-β1 pathway in inducing

EMT and CSC characteristics in HCC cells (Park et al., 2016). $CD44^{+}TGF\text{-}\beta1^{+}$ cells showed lower E-cadherin but higher N-cadherin expression levels in addition to increased levels of AKT, GSK-3β and β-catenin. They also formed larger spheres, an alternative measurement of CSC presence *in vitro*. Upon TGF-β1 inhibition or CD44 knockdown, HCC cells could not exhibit both characteristics. Separately, the increase of $CD44^{+}$ cells in CRC as a result of hypoxia was found to be related to Wnt/β-catenin pathway (Dong et al., 2016).

Although the utility of $CD44^{+}$ as a marker for CSCs and/or TICs as well as its interplay in major deregulated pathways for cancers has been vastly examined and verified in various solid tumours as reviewed above, Oh et al. (2013) reported that $CD44^{-}$ cells in HNSCC spheroids behaved similarly to $CD44^{+}$ cells. Both subpopulations were able to regenerate spheroids, had increased levels of OCT4, SOX2 and nestin proteins, possessed similar chemoresistance to cisplatin and percentage of SP cells, and similar tumour-forming ability in nude mice (TIC frequencies of 1 in 2,309 cells and 1 in 4,579 cells for $CD44^{+}$ and $CD44^{-}$, respectively). Subsequently, Zheng & Franzmann (2013) pointed out the possibility of phenotype conversion as one of the plausible reasons for causing the attributes of $CD44^{-}$ cells in the former study.

Literature review of CD44 in NPC studies will be covered in Chapter 4.

2.4.3 EpCAM

Since its discovery in 1979, epithelial cell adhesion molecule (EpCAM) was referred to by numerous names including epithelial surface antigen (ESA) and human epithelial antigen (HEA125). The nomenclature has since been set as EpCAM or CD326 (Trzpis et al., 2007). EpCAM is a surface marker, comprising of 314 amino acids and functions as an epithelial-specific intercellular cell adhesion molecule with additional involvement in cellular signalling, cell migration, proliferation and differentiation

(Patriarca et al., 2012). Although EpCAM is expressed by both normal and malignant epithelial tissues, its expression is low on the basolateral membrane of normal epithelial cells as opposed to high expression level on apical membrane in cancer cells (Ni et al., 2012).

Despite the establishment of EpCAM's positive role in cell proliferation and elucidation of mechanisms involved as mentioned above and proven by Münz et al. (2004), Kroepil et al. (2013) and others, EpCAM-transfected pancreatic cell lines showed lower or no difference in proliferation, and lower abilities to invade and migrate (Akita et al., 2011). In fact, there was a positive correlation between its expression and overall survival in pancreatic cancer patients (56.2% of EpCAM-high patients survived their cancer for 3 years versus only 19.2% for EpCAM-low patients, $p < 0.01$). The authors argued that the clinical significance of EpCAM depends on type of malignancy as other studies had found that it was both a good and a bad prognostic marker in cancers such as kidney, gastric, breast and gallbladder. Separately, EpCAM also had no effect on cell proliferation in a prostate carcinoma study (Massoner et al., 2014). Although EpCAM expression was incremental from low to high grade and metastatic lesions in patients, both parental and EpCAM-knockdown prostate carcinoma cell lines had comparable growth *in vitro*.

$EpCAM^{pos}$ cells from mouse breast cancer cells displayed certain features of CSCs (Hiraga et al., 2016). Although there was no difference in *in vitro* proliferation, invasion and migration abilities between $EpCAM^{pos}$ and $EpCAM^{neg}$ cells, there was a significant increase of tumour-initiating potential in $EpCAM^{pos}$ cells at the lowest inoculation of 100 cells (TIC frequency of 1 in 95 cells compared to 1 in 281 cells for $EpCAM^{neg}$ cells) as well as in the development of bone metastasis by $EpCAM^{pos}$ cells. Nonetheless, lentiviral transduction of EpCAM into $EpCAM^{neg}$ cells could not establish

the features of $EpCAM^{pos}$ phenotype in these cells, which showed that EpCAM alone was most likely insufficient to induce stemness in mouse breast cancer.

EpCAM marker has also been used in combination studies with other common CSC markers to explore its functions and roles. Its role in tumour initiation cannot be overlooked in the seminal study on tumourigenic breast cancer cells (Al-Hajj et al., 2003). The presence of ESA positive cells in the combination of $CD44^{+}CD24^{-/low}$ formed 100% tumours in all inoculations from 1,000 to 100 cells (24/24) as compared to no growth in $CD44^{+}CD24^{-/low}ESA^{-}$ cells (0/24).

The combination of $EpCAM^{+}$ and $CD44^{+}$ were found to be identifying for CSCs in gastric cancer samples as these cells were consistently more tumour-initiating and were able to self-renew *in vivo* as compared to $EpCAM^{-}CD44^{-}$ and the remaining two combinations (Han et al., 2011). $EpCAM^{+}CD44^{+}$-resulting tumours were successfully passaged for three cycles; yet, $EpCAM^{-}CD44^{-}$ cells failed to induce secondary growths.

$CD166^{+}EpCAM^{+}$ marker was selective for CSC-like cells in NSCLC cell line, A549 (Norashikin et al., 2015). Although there were promising results from *in vitro* and *in vivo* growth experiments, moderate upregulation of pathways regulating drug metabolism and ABC drug transporters, and downregulation of p53, apoptosis and ECM-receptor interaction pathways, *in vivo* self-renewal ability of $CD166^{+}EpCAM^{+}$ cells was not evaluated in the study.

In the search for markers to identify circulating tumour cells (CTCs) which are involved in disseminating metastases, EpCAM by virtue of it being an epithelial cell marker, was selected as the marker for CTCs in the first and only US Food and Drug Administration-approved test kit for clinical detection of metastatic breast, colorectal and prostate cancer (Hofman et al., 2011). An increased understanding of CTCs and

EMT-MET processes later led to disputes concerning the use of EpCAM to monitor treatment efficacy and to detect disease recurrence (reviewed in Gabriel et al., 2016).

Literature review of EpCAM in NPC studies will be covered in Chapter 4.

CHAPTER 3: EVALUATION OF STEM-LIKE SIDE POPULATION CELLS IN A RECURRENT NASOPHARYNGEAL CARCINOMA CELL LINE

3.1 Introduction

Chapter 3 illustrates the use of side population (SP) assay as my first approach to identify, isolate and study biological as well as gene expression differences between tumourigenic and non-tumourigenic cells in NPC. HK1 is a cell line established from a well-differentiated recurrent NPC sample (Huang et al., 1980), while xeno-284 is a PDX line established in our laboratory from a poorly differentiated recurrent metastatic NPC sample. Both HK1 and xeno-284 are EBV negative. The presence of the SP subpopulation was first assessed in HK1 and xeno-284 cells. Due to the scarcity of SP cells in xeno-284, sorting of SP and NSP subpopulations for comparison of *in vitro* growth properties could only be performed in HK1. Aldehyde dehydrogenase (ALDH) activity which is usually high in stem cells and CSCs was also determined in the sorted HK1 cells. Gene expression studies were performed in sorted HK1 to identify stem cell related genes and pathways which may be responsible for *in vitro* observations. Finally, *in vivo* tumourigenicity experiments were performed using sorted HK1 for duration of up to seven weeks to evaluate the tumour-initiating ability of its SP and NSP cells.

3.2 Brief Literature Review

Nasopharyngeal carcinoma (NPC) is the most common malignancy arising from the nasopharynx and its causation is closely associated with the Epstein-Barr virus, environmental as well as dietary factors (Lo et al., 2004). Majority of NPC cases present in late stages which is largely due to the hidden location of the tumour in addition to either no or apparently trivial symptoms ignored by patients or even medical professionals (Khoo & Pua, 2013; Pua et al., 2008). Also, NPC disease recurrence, therapeutic resistance and metastasis remain major clinical problems (Wei & Sham, 2005).

The cancer stem cell (CSC) model is one of the models used to describe tumour initiation (Baccelli & Trumpp, 2012; Imrich et al., 2012; Visvader & Lindeman, 2008). It hypothesizes that there is a hierarchy within the tumour cell population and only a rare subset of cancer cells has the ability to self-renew and to differentiate, leading to the recapitulation of the original tumour (Clarke et al., 2006). As such, CSCs are believed to be important contributors to the process of metastasis, and reactivation of proliferating CSCs is thought to be one of the underlying causes of disease recurrence. CSCs are found to behave differently from the rest of tumour cells; amongst others, they have enriched tumour-initiating potential and have efficient drug extrusion systems to evade most chemotherapeutic drugs (Alison et al., 2012). These cells undergo asymmetric divisions to give rise to daughter cells: one will be stem-like and the other does not show stem cell characteristics. Some CSC-enriched subpopulations were slower in proliferation (reviewed in Moore & Lyle, 2011), whilst others reported an equal or rapid proliferation rate than the non-CSC subpopulations (Akunuru et al., 2011; Cao et al., 2011).

A limiting dilution analysis (LDA) in animal models entails inoculating a series of cell numbers into the hosts to quantify frequency of cells possessing a particular function which are present in a mixed and larger population (Hu & Smyth, 2009; Rosenbloom et al., 2015). As it is a dose-response experiment in which each cell number elicits a positive or negative response in the host, LDA is typically used in stem cell and CSC studies to estimate the frequency of TICs. Hu and Smyth (2009) designed a highly-cited free online software called Extreme Limiting Dilution Analysis (ELDA; http://bioinf.wehi.edu.au/software/elda/) to calculate TIC frequencies for stem cell research, amongst other analyses. TIC frequency is computed as a central estimate and confidence interval. Three key features of this software are its ability to handle extreme data situations with 0% or 100% responses, to compare frequencies across multiple

populations, and to analyse non-Poisson distribution. ELDA also accommodates well the small numbers of replicates usually seen in stem cell and CSC research by using likelihood ratio tests, instead of t-test methods which can give misleading results in small sample sets.

As CSCs are known to possess drug extrusion ability, SP assay, originally established by Goodell et al. (1996) for haematopoietic stem cells, was used for identification of putative stem cells and progenitors in solid tumours (Fukaya et al., 2010; Kato et al., 2010; Lim et al., 2011). The ability of SP cells to extrude the Hoechst 33342 dye, causing them to appear as dimly stained cells in flow cytometry dot plots, is dependent on the activity of the ATP-binding cassette (ABC) drug transporter family which includes ABCB1, ABCC1 and ABCG2 (Wu & Alman, 2008). Verapamil is a potent inhibitor for ABCB1 which also weakly inhibits ABCG2 activities, while FTC specifically inhibits ABCG2 (Duan et al., 2004; Robey et al., 2007).

In NPC research, SP assay was first used in CNE-2 a poorly-differentiated NPC cell line (Wang et al., 2007). SP cells were identified with the use of verapamil indicating the involvement of ABCB1 drug transporter in CNE-2. The presence of *ABCB1* gene, however, was not confirmed in qPCR. Instead, *ABCG2* gene was measured and found to be lowly expressed in SP cells with no expression in NSP cells. Although the growth rate of SP cells during culture was faster than NSP cells, both cells exhibited similar proliferation index and S-phase fraction in cell cycle analysis. SP cells also displayed the ability to differentiate into NSP cells by asymmetric cell division as the percentage of original SP cells decreased after nearly three weeks in culture. A 4-week *in vivo* tumourigenicity experiment revealed that SP cells formed more tumours in NOD/SCID mice than NSP cells in a limiting dilution manner from 200,000 cells to 10,000 cells. The SP cells also showed higher resistance to X-ray which was attributed to a higher

presence of Smoothened (SMO) protein in the cells. SMO is a receptor protein in the Hedgehog pathway which has been implicated in tumourigenesis, embryonic development and EMT, amongst others (Takebe et al., 2015).

A follow-up study by the group found that cancer cells with stem-like features could be induced by DNA damage (Liang et al., 2010). UV and mitomycin C treatment increased the proportion of SP cells by approximately 3 folds in CNE-2 S26 cells which were related to deregulated levels of key regulators of cell cycle and mitosis Mad2, Aurora B and Cdh1. Besides DNA damage, the presence of a latent membrane protein of the EBV, LMP2A, could also increase SP population, in part, through PI3K/Akt pathway (Kong et al., 2010).

Ma et al. (2013) reported the influence of another pathway on SP phenotype. Treatment with epidermal growth factor (EGF) increased SP cells by 2.5 and 1.6 folds in CNE-1 and CNE-2 cells, respectively and also raised the levels of phosphorylated epidermal growth factor receptor (EGFR), AKT and ERK1/2. Gefitinib, an inhibitor of the EGFR pathway, reduced SP cells to almost none in both cell lines, in addition to a substantial reduction in phosphorylated EGFR. Likewise, PI3K and ERK inhibitors only abolished the phosphorylation of AKT and ERK1/2, respectively. Knockdown of *EGFR* expression led to reduction of spheroid formation in parental CNE-2 cells. β-catenin was also influential in maintaining the SP phenotype as sh-β-catenin-transduced cells showed a reduction of SP percentages with and without addition of EGF. It was thus concluded that SP phenotype and stemness properties in NPC were regulated by EGFR/PI3K/AKT pathway and mediated by β-catenin.

Many important findings on SP and CSC phenotypes, stemness-related functions and cellular properties in NPC were ascertained from investigations performed on CNE-1

and CNE-2 cell lines as reviewed above. Chan et al. (2008) performed an extensive authentication of CNE-1 and CNE-2 in addition to other commonly used NPC cell lines and xenografts. Due to the highly similar short tandem repeat (STR) profiles of these two cell lines, it was implied that they were genetically similar and derived from one single cell line despite both lines being established from two different NPC tumours. Repeated profiling of both lines sourced from other laboratories in Hong Kong and Singapore also gave similar results. It was further reported that HPV-18 genome was present in CNE-1 and CNE-2 and their STR profiles were similar to the STR profile of HeLa, postulating that CNE-1 and CNE-2 could be developed "from the fusion of HeLa with an unknown cell line by somatic cell hybridization" (Chan et al., 2008, p. 2170).

NPC cell lines are not currently available in cell line repositories such as American Type Culture Collection (ATCC) with high biological standards in the acquisition, authentication, production and preservation of reference cell lines. In my study, HK1 cells had been previously obtained from the University of Hong Kong and STR profiling had been performed in the Department of Chemistry, Malaysia. HK1 cells used in my study shared a 100% match of STR profile with the one published by Chan et al. (2008) (Appendix C).

3.3 Materials

3.3.1 Cell line

HK1 cells were obtained as a gift from Professor George SW Tsao (University of Hong Kong). The cell line was established from a recurrent NPC patient having well-differentiated squamous carcinoma with prior radiation therapy (Huang et al., 1980).

3.3.2 NPC patient-derived xenograft (PDX)

Xeno-284 was established from a recurrent NPC patient with poorly differentiating squamous cell carcinoma with prior concurrent chemo-radiotherapy (manuscript in preparation).

3.3.3 Cell lines and culture reagents

3.3.3.1 Complete medium

Total volume of 100 mL

RPMI-1640 medium (with phenol red)	89.5 mL
Fetal bovine serum	10 mL
10,000 U/mL Penicillin-Streptomycin	500 μL

All items were obtained from Thermo Fisher Scientific Inc., MA, USA.

3.3.3.2 Freezing medium

Total volume of 10 mL

Complete medium	5 mL
Dimethyl sulfoxide Hybri-Max (DMSO)	1 mL
Fetal bovine serum	4 mL

DMSO was purchased from Sigma-Aldrich, MO, USA.

3.3.3.3 1X phosphate-buffered saline (PBS)

Total volume of 100 mL

Deionized water (dH_2O)	100 mL

Phosphate-buffered salt	1 tablet

Phosphate-buffered salt was obtained from Takara Biotechnology, Shiga, Japan.

3.3.3.4 1X PBS supplemented with 1X antibiotic-antimycotic

Total volume of 100 mL

1X PBS	99 mL
100X antibiotic-antimycotic	1 mL

Antibiotic-antimycotic was purchased from Thermo Fisher Scientific Inc., MA, USA.

3.3.4 HBSS+ buffer

Total volume of 100 mL

Hanks' balanced salt solution (HBSS)	98 mL
Fetal bovine serum	2 mL
HEPES	0.2383 g

HBSS and HEPES were obtained from Thermo Fisher Scientific Inc., MA, USA and Sigma-Aldrich, MO, USA, respectively.

3.3.5 Reagents for RT-qPCR

3.3.5.1 75% ethanol

Total volume of 10 mL

Absolute ethanol	7.5 mL
dH_2O	2.5 mL

Absolute ethanol (molecular biology grade) was purchased from Merck Millipore, MA, USA.

3.3.5.2 Reverse transcription (RT) mix

Total volume of 10 μL (equivalent to 1 reaction of RT)

5X buffer BC3	4 μL
control P2	1 μL
RE3 reverse transcriptase mix	2 μL
RNase-free water	3 μL

All items were components from RT^2 Profiler PCR Array kits (SABiosciences, MD, USA).

3.3.5.3 quantitative PCR (qPCR) mix

Total volume of 2.7 mL (equivalent to a 96-well PCR Array)

2X RT^2 SYBR Green mastermix	1.35 mL
cDNA	102 μL
RNase-free water	1.248 mL

All items were components from RT^2 Profiler PCR Array kits (SABiosciences, MD, USA).

3.4 Methods

This study was performed according to the protocols approved by Research Management Committee of Institute for Medical Research (project code: JPP-IMR 09-

037) and Animal Care and Use Committee of the Ministry of Health, Malaysia (approval code: ACUC/KKM/02(1/2011)).

3.4.1 Cell culture

HK1 adherent cells were cultured in complete medium. The cells were maintained in a 5% CO_2-humidified incubator (model NU-4750E, Nuaire, MN, USA) at 37 °C. Cell culture experiments were carried out in a biosafety cabinet class II and culture media and reagents were warmed to 37 °C in a water bath prior to use.

3.4.1.1 Thawing of frozen culture and culturing of revived cells

A cryovial containing cryopreserved cells was removed from the liquid nitrogen tank and immediately thawed in a water bath at 37 °C with gentle swirling. Thawed content was pipetted dropwise into a 15-mL centrifuge tube containing 5 mL of cold complete medium. The cell suspension was centrifuged at 1,000 rpm for 5 min and the supernatant was subsequently discarded. The cell pellet was resuspended with 5 mL of complete medium and transferred to a 6-cm culture dish. The dish was maintained in the incubator. Medium was changed every 2 - 3 days.

3.4.1.2 Passaging of cells

HK1 cells were passaged when they were 80 - 90% confluent (i.e. logarithmic growth phase). The spent medium was discarded and the cells were gently rinsed with 3 mL of 1X PBS. The cells were trypsinized with 2 mL of TrypLE™ Express Enzyme (1X) (Thermo Fisher Scientific Inc., MA, USA) at 37 °C for 5 - 8 min. Cells were viewed under a phase contrast inverted microscope for detachment. Trypsinization process was stopped with the addition of 2 - 3 mL of complete medium. The cell suspension was transferred to a 15-mL centrifuge tube and centrifuged at 1,000 rpm for 5 min. The supernatant was discarded and the cell pellet was gently resuspended with 2

mL of complete medium prior to transfer to the required numbers of new 10-cm culture dishes. Six mL of complete medium were then added with gentle swirling to ensure an even distribution of cells in each dish.

3.4.1.3 Cryopreservation

Freezing medium was freshly prepared and chilled for each cryopreservation process. After the cells had been detached, centrifuged and the supernatant was discarded, the cell pellet was gently resuspended with chilled freezing medium. Cryovials were prelabelled and chilled prior to filling with 1 - 2 X 10^6 cells per vial. The cryovials were then stored at -20 °C for 2 - 3 h, followed by -80 °C storage for overnight prior to long term storage in liquid nitrogen.

3.4.1.4 Cell counting by automated cell counter

Following cell detachment and resuspension in complete medium, the cells were counted prior to use. Ten μL of cell suspension were combined with 10 μL of trypan blue solution for cell viability and counting checks. Ten μL of the mixture were then loaded into a Countess® cell counting chamber slide. The loaded slide was then inserted into the Countess® automated cell counter (Thermo Fisher Scientific Inc., MA, USA). Image quality was first adjusted to ensure a good contrast between live and dead cells before cells were counted automatically.

3.4.1.5 Mycoplasma detection

Mycoplasma contamination in cultured cells was periodically checked with Venor GeM Mycoplasma Detection Kit for Conventional PCR (Minerva Biolabs, Berlin, Germany) following the manufacturer's protocols.

3.4.2 NPC patient-derived xenograft

Xeno-284 xenograft was maintained and passaged *in vivo* in NOD-scid gamma (NSG) mice (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/ SzJ; The Jackson Laboratory, ME, USA) as a subcutaneous growth in the SPF Animal Facility (Institute for Medical Research, Malaysia). Only xeno-284 samples of passage 12 and below were used in this study.

3.4.2.1 Harvesting of xenograft

Mouse bearing a tumour was checked for its correct identity via its ear tag. The mouse was humanely euthanized using CO_2 and/or cervical dislocation method. The exterior of the subcutaneous xenograft and its surrounding area was first disinfected with 10% w/v Povidone iodine solution (Polylab Biotech Sdn. Bhd., Malaysia). The xenograft was gently excised out and transferred to a sterile glass petri dish. Visible blood clots, blood capillaries and/or fat were removed. It was then cut length-wise to inspect for the presence of necrotic tissue which would be removed. A cross section of the xenograft tissue was preserved in 10% neutral buffered formalin (Leica Biosystems, IL, USA).

3.4.2.2 Digestion of xenograft

Xeno-284 was freshly harvested as above and rinsed with cold 1X PBS supplemented with 1X antibiotic/antimycotic. The xenograft tissue was minced finely in the presence of 5 mL of 1X collagenase/dispase solution (Roche, Mannheim, Germany) in a sterile glass 6-cm petri dish. The mixture was incubated for 60 min in a 5% CO_2-humidified incubator at 37 ºC with constant mixing. The cell suspension and undigested xenograft pieces were separated by sieving through 40-µm cap strainer. The clear flow-through was centrifuged at 900 rpm for 6 min and the supernatant was discarded. The cell pellet was resuspended in 2 mL of RBC lysis buffer (Qiagen, Hilden, Germany) followed by another centrifugation step as described earlier. The cell

pellet was resuspended with 2 - 3 mL of 1X PBS and centrifuged as described earlier. Supernatant was discarded and the pellet was resuspended with HBSS+ buffer prior to cell count and viability check.

3.4.3 SP assay with Hoechst 33342

For this assay, HK1 cells were grown in 10-cm culture dishes and were consistently used at logarithmic growth phase for every flow cytometry analysis or cell sorting. Spent medium was discarded and the cells were rinsed with 1X PBS. Cell detachment for SP assay was performed with Accutase (Millipore, MA, USA) instead of TrypLE™ Express Enzyme (1X) under the same conditions. The chemical components in Accutase are proprietary with "proteolytic and collagenolytic" activity. It was used for Hoechst 33342 experiments as it contained reportedly less proteolytic activity than trypsin-containing dissociation enzymes such as TrypLE™ Express, thus ensuring a better preservation of the biological state of HK1 during and after cell detachment (Biddle et al., 2013). Cell suspension was centrifuged as described above, supernatant was discarded and cell pellet was resuspended with 1X PBS for a wash prior to Hoechst 33342-staining.

A modification of Goodell's method was used to stain both HK1 and xeno-284 cells (Goodell et al., 1996). The cells were resuspended at a concentration of 1 X 10^6 cells per mL in HBSS+ buffer. Hoechst 33342 dye (Molecular Probes, OR, USA) was added into the cell suspension at a final concentration of 5 μM in the presence or absence of FTC (ABCG2 inhibitor) at a final concentration of 1 μM (Sigma, MO, USA) or verapamil (ABCB1 inhibitor) at a final concentration of 50 and 100 μM (Sigma, MO, USA). The cells were incubated for 90 min in a 37 ºC water bath with intermittent mixing. The centrifuge tubes containing stained cells were immediately immersed into an ice bucket and all subsequent steps were performed on ice, with cold buffer or at 4

°C. Excessive Hoechst 33342 dye was removed from the cell suspension by washing with cold HBSS+ followed by centrifugation at 1,000 rpm (HK1 cells) or 900 rpm (xeno-284 cells) for 6 min at 4 °C. The resulting pellet was resuspended to a final concentration of 1 X 10^6 cells per mL with cold HBSS+. In order to delineate host mouse cells, xeno-284 cells were further stained with H2Kd-phycoerythrin (PE) antibody (1:10 dilution, BD Pharmingen, MA, USA) for 30 min at 4 ºC. Hoechst 33342-stained HK1 and xeno-284 cells were also stained with propidium iodide (PI, 2 μM, BD Pharmingen, MA, USA) and 7-aminoactinomycin D (7-AAD, 1:400 dilution, BD Pharmingen, MA, USA), respectively, for determination of cell viability. 7-AAD (emission wavelength ~670 nm) was used as a viability dye for xeno-284 cells due to a similar emission wavelength of approximately 585 nm for PI and PE (conjugated to H2Kd antibody), leading to a technical challenge in discriminating the fluorescence signals accurately. The stained cells were analysed and sorted in a BD FACSAria II SORP cytometer (BD Biosciences, MA, USA) equipped with a 355-nm UV laser power of 50 mW. The Hoechst 33342 fluorescence was detected via a 405/20 band-pass filter (Hoechst Blue) and a 670 long pass filter (Hoechst Red).

Flow analysis was performed with BD FACSDiva software (version 6.1.3; BD Biosciences, CA, USA) in a hierarchical manner. Single cells were identified and gated using the in-built doublet discrimination gating strategy. The single cells were then analysed for viable cells, after which SP cells were identified with FTC or verapamil inhibitor. Presence of cells was reported as percentage of single, viable SP or NSP cells.

3.4.4 Clone morphology experiment

Sorted cells were plated at a low cell density of 50 cells per well in a 96-well culture plate containing complete medium. Upon seeding, microscopic inspection of each well

was performed to ensure there was no cell clumping. The cells were incubated in 5% CO_2-humidified incubator at 37 °C. Once the cells were adherent, the culture medium was changed every two days. The identification and description of different clone morphologies were performed using previously reported method (Felthaus et al., 2011; Harper et al., 2007; Roudi et al., 2016; Shigeishi et al., 2013). The numbers of holoclones, meroclones and paraclones were counted with a phase contrast inverted microscope on day 8 post-sorting. Data were recorded from three independent experiments with 12 replicate wells per SP or NSP cells in each experiment.

3.4.5 Aldehyde dehydrogenase (ALDH) experiment

Sorted cells were left for overnight recovery in the culture conditions as mentioned above. The cells were detached using Accutase as described above and ALDH staining was performed with the ALDEFLUOR staining kit (Stem Cell Technologies, Vancouver, Canada). After cell count, 0.25 X 10^6 cells from each population (SP or NSP) were resuspended in ALDEFLUOR assay buffer containing 1.5 µM ALDH activated substrate (BODIPY-aminoacetaldehyde , BAAA) ("test" sample). Half of each stained population was then mixed with 15 µM ALDH inhibitor (N,N-diethylaminobenzaldehyde, DEAB) in another tube ("control" sample). All stained samples were then incubated for 45 min at 37 ºC with constant mixing. This was followed with a centrifugation step at 1,000 rpm for 5 min at 4 °C. The stained cells were washed, resuspended in cold ALDEFLUOR assay buffer containing 2 µM PI and analysed in BD FACSCalibur (BD Biosciences, MA, USA). SKBR3 cell line was used as a staining positive control for the experiment. Analysis gate for ALDH positive cells was drawn on stained viable parental HK1 cells with DEAB inhibitor. Presence of cells was reported as percentage of viable ALDH positive cells. Data were obtained from five independent experiments.

3.4.6 Asymmetric division experiment

Ten thousand sorted cells were re-cultured for three weeks in complete medium as described above. Hoechst 33342 staining was performed and the cells were analysed in a BD FACSAria II SORP cytometer as previously described. Data were recorded from three independent experiments.

3.4.7 Proliferation experiment

Sorted cells were left to recover from the sorting process in a 5% CO2-humidified incubator at 37 ºC for approximately two hours. The "recovered" cells were seeded at 3,000 cells per 200 µL of complete medium into each well of the E-plate 16 (Roche, Mannheim, Germany). Cell index values were recorded over a period of 125 h with an interval of 1 h for the first day, followed by every 6 h for the remaining experiment by the xCELLigence System's Real time Cell Analyser (RTCA) DP instrument (Roche, Mannheim, Germany). Cell index values represent measurements of electrical impedance of monitored cells which reflect cell growth (number and viability), morphology and adhesion ability. The cell index values of SP and NSP cells at each time point were then normalized to the control cells (parental HK1 cells which were only stained with PI and sorted from PI negative gate). Data were obtained from three independent experiments.

3.4.8 Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) experiment

Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) was performed using RT^2 Profiler Human Stem Cell and Stem Cell Signaling PCR Arrays and kits (SABiosciences, MD, USA). The 96-well Human Stem Cell PCR Array detects expression of 84 genes related to the identification, growth and differentiation of stem cells (see Appendix A for list of genes). The 96-well Human Stem Cell Signaling

PCR Array profiles the expression of 84 key genes involved in signal transduction pathways important for embryonic stem cell (ESC) and induced pluripotent stem cell (iPSC) maintenance and differentiation (Appendix B). Each array contains five housekeeping genes, genomic DNA (gDNA) control, RT control and positive PCR control.

3.4.8.1 Total RNA extraction

Total RNA was isolated using TRIzol (Invitrogen, CA, USA). Sorted cells were centrifuged at 1,000 rpm for 5 min and the supernatant was discarded. The pellet was resuspended with 750 μL of TRIZOL Reagent per 5 - 10 X 10^6 cells and the cells were lysed by repetitive pipetting. The homogenized cells were incubated for 5 min at room temperature to allow for complete dissociation of nucleoprotein complexes, followed by the addition of 200 μL of chloroform per 1 mL of TRIZOL Reagent. The tube was capped securely, vortexed vigorously for 15 s and incubated at room temperature for 2 - 3 min. The cell mixture was centrifuged at 12,000 x g for 15 min at 4 °C. Following centrifugation, the mixture was separated into three layers: lower red phenol-chloroform phase, an opaque interphase, and a colourless upper aqueous phase. The upper phase which contained RNA was carefully removed into a new tube without disturbing the interphase. RNA was then precipitated with the addition of 500 μL of 100% isopropanol per 1 mL of TRIZOL Reagent used. The mixture was incubated at room temperature for 10 min and centrifuged at 12,000 x g for 10 min at 4 °C. The supernatant was carefully aspirated and RNA could be seen as a gel-like pellet at the side or bottom of the tube. The pellet was washed with 1 mL of 75% ethanol per 1 mL of TRIZOL Reagent used. Resuspension of the pellet in 75% ethanol was properly done by vortexing. The mixture was then centrifuged at 7,500 x g for 5 min at 4 °C and

the ethanol was carefully removed. The clean RNA pellet was air-dried for 5-10 min at room temperature. RNA was dissolved in 10-20 μL of nuclease-free water.

3.4.8.2 Quantification of total RNA

The concentration and quality of the extracted RNA was determined with NanoDrop 8000 spectrophotometer (Thermo Scientific, DE, USA). Nuclease-free water was used as a "blank". Only RNA samples with a 260/280 ratio of ~ 1.8 to 2.0 were used for reverse transcription (RT).

3.4.8.3 Reverse transcription (RT)

Prior to RT, gDNA elimination was performed on the total RNA. For a 10-μL gDNA elimination reaction mixture, 100 ng of total RNA was added with 2 μL of Buffer GE and the remaining volume was made up to 10 μL with nuclease-free water. The mixture was incubated at 42 °C for 5 min and immediately placed on ice for 5 min.

Ten μL of RT mix was then added into the tube containing 10 μL of gDNA elimination mixture. The content was mixed gently by pipetting prior to incubating at 42 °C for exactly 15 min. RT reaction was stopped by incubating the mixture at 95 °C for 5 min. Subsequently, 91 μL of nuclease-free water was added into the reaction and mixing was performed by pipetting several times. The mixture was then placed on ice before proceeding with qPCR.

3.4.8.4 Quantitative PCR (qPCR) and data analysis

For a 2.7-mL mixture of qPCR components sufficient for a 96-well PCR Array, 102 μL of cDNA was mixed with 2.598 mL of qPCR master mix. The components were evenly mixed and dispensed into a nuclease-free loading reservoir. Using an 8-channel pipettor, 25 μL of qPCR mixture was added into each well of the PCR Array. The array

was tightly sealed with the provided optical adhesive film and centrifuged at 1,000 x g for 1 min at room temperature to remove bubbles. The array was then loaded into the ABI7500 Fast Real-Time thermal cycler (Applied Biosystems, CA, USA) for qPCR: activation of HotStart DNA *Taq* polymerase at 95 °C for 10 min, followed by 40 cycles of denaturing at 95 °C for 15 s and annealing/extending at 60 °C for 1 min.

Relative expression was determined using comparative cycle threshold (C_T) method. Genes with C_T values of more than 35 were removed from analysis. C_T value of each gene was normalized against the geometric mean C_T value of five housekeeping genes (C_T geoHK), with the formulae 2^{-dCT} (dC_T = C_T gene of interest – C_T geoHK). The normalized values were used to calculate fold change ratios between SP and NSP cells ($=2^{-ddCT}$). Data were obtained from three independent experiments.

3.4.9 *In vivo* tumourigenicity experiment

Inoculations of SP and NSP cells into nude mice were performed in a limiting dilution manner with 1,000, 100 and 10 cells. Sorted cells were mixed with BD Matrigel™ basement membrane matrix (BD Biosciences, MA, USA) and inoculated subcutaneously into 5 – 6-week-old nude mice (Balb/c nu/nu; Animal House of IMR, Malaysia). Tumour latency data, defined as the period between inoculation day and the first day of tumour detection, were recorded. Once a palpable growth was detected, tumour volume was recorded every two days. Tumour volume (mm^3) was calculated from 0.5 X ($width^2$ X length). All tumours were harvested once the length or width reached 10 mm (animal ethics' consideration) or on day 50 post-inoculation (to maintain an equal tumour burden amongst the mice). Limiting dilution analysis was performed using the Extreme Limiting Dilution Analysis (ELDA) software (Hu & Smyth, 2009). Data for each group of mice were from three independent experiments.

3.4.10 Haematoxylin and eosin (H&E) staining

SP and NSP tumours were formalin-fixed for 24 h and paraffin-embedded in an automated tissue processor Leica ASP300 S (Leica Biosystems, Victoria, Australia). Four-micrometre formalin-fixed, paraffin-embedded tissue sections from SP and NSP tumours were mounted on plain glass slides and stained with H&E in an autostainer Leica XL (Leica Biosystems, Victoria, Australia). Histopathological observations were made under a light microscope by a pathologist and all tumours were confirmed to be NPC.

3.4.11 Statistical analysis

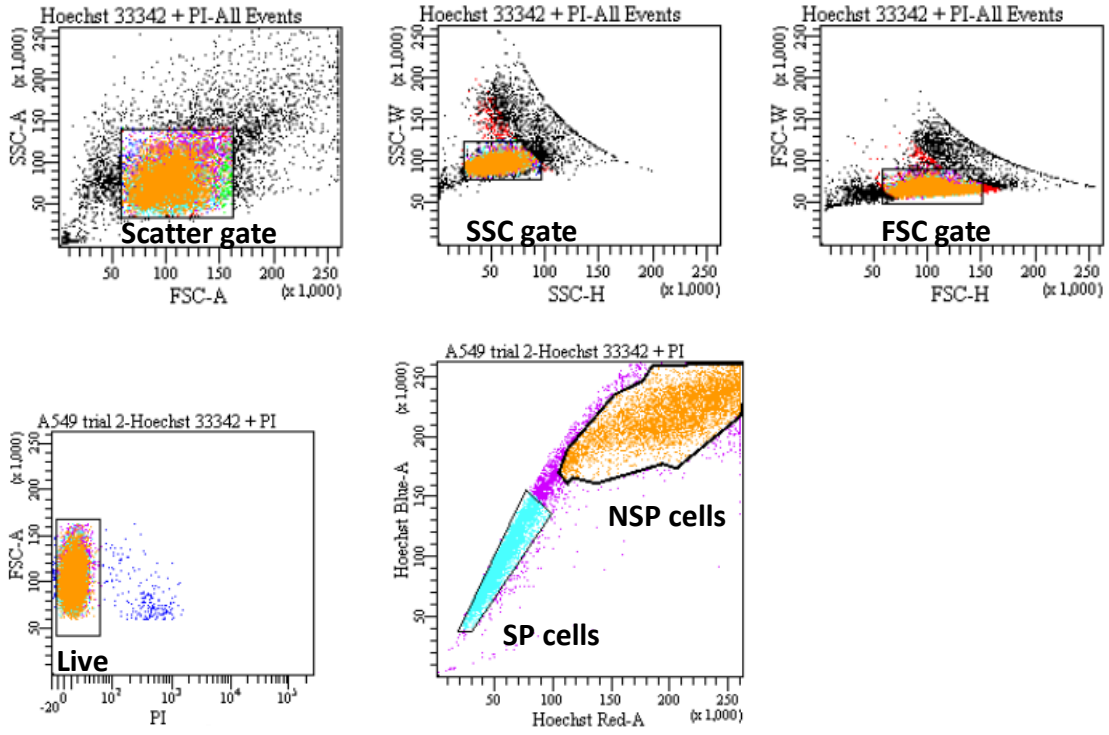
Data are reported as mean ± standard deviation (SD), mean ± standard error of the mean (SEM) or boxplot with whiskers showing the 5 - 95 percentile as indicated in the figure description. All statistical analyses were performed using paired Student's t test from the GraphPad Prism 5 software (GraphPad Software Inc., CA, USA), except for clone morphology which used unpaired Student's t test. A p-value of < 0.05 was deemed to be statistically significant.

3.5 Results

3.5.1 Optimization of Hoechst 33342 staining

Hoechst 33342 staining assay was first optimized using human lung carcinoma cell line A549 which had been reported to contain a high percentage of SP cells (Sung et al., 2008). Two main parameters were optimized: the concentration of Hoechst 33342 dye and the type of ABC transporter inhibitor. There were 31.7% of single, viable SP cells which was equivalent to 25.5% of total cells and similar to Sung et al. with the commonly used concentration of 5 μM of Hoechst 33342 (Figure 3.1).

i.



ii.

Tube: Hoechst 33342 + PI

Population	#Events	%Parent	%Total
All Events	16,368	#####	100.0
Scatter gate	13,827	84.5	84.5
SSC gate	13,677	98.9	83.6
FSC gate	13,433	98.2	82.1
Live	13,190	98.2	80.6
SP cells	4,181	31.7	25.5
NSP cells	5,843	44.3	35.7

Figure 3.1: Staining of A549 cells with 5 μM of Hoechst 33342 dye and viability dye PI. **i.** Using doublet discrimination gating strategy, stained A549 cells were identified and gated for single cells ("FSC gate"). The single cells were then gated for a subset of live or viable cells ("Live"). Finally, SP cells were identified and gated using an appropriate inhibitor ("SP cells"). **ii.** Statistics box shows the percentage of cells analysed in a hierarchical manner. Percentage of cells are calculated from the population of cells they are gated from ("%Parent") or from the total cell population ("%Total"). FSC: forward scatter, SSC: side scatter, A: area, H: height, W: width, PI: propidium iodide

Optimization of the appropriate ABC transporter inhibitor to be used for A549 was also performed (Figure 3.2). There was a total inhibition of dye efflux ability of A549 cells with FTC at a final concentration of 1 μM, unlike verapamil at a final concentration of 50 μM.

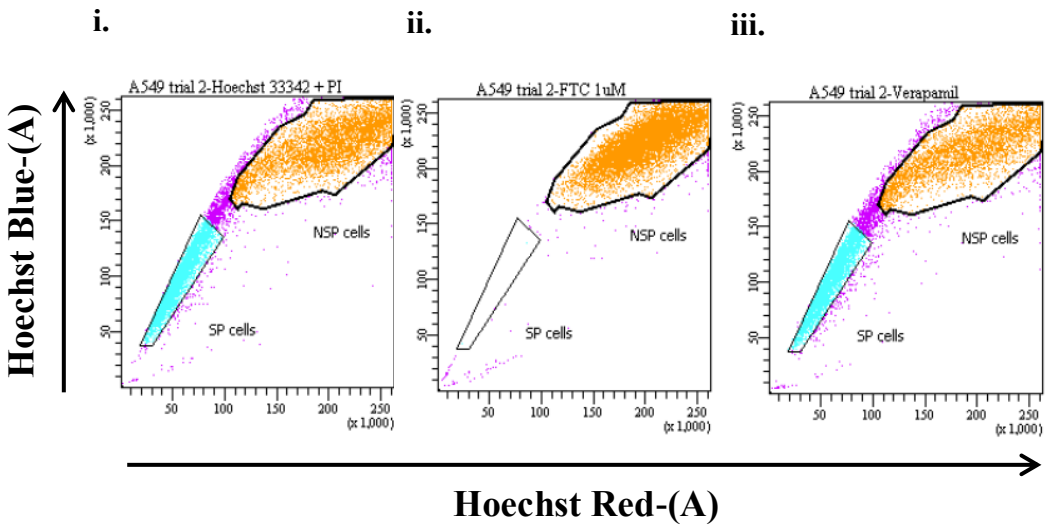


Figure 3.2: Optimization of ABC transporter inhibitor in A549. **i.** A549 cells stained with Hoechst 33342. **ii.** FTC at 1 μM was able to block the dye efflux ability of stained A549 cells which led to an absence of cells with low Hoechst staining ("SP cells"). **iii.** Verapamil at 50 μM was unable to block the dye efflux ability of stained A549 cells, leading to many cells with varied intensity levels of Hoechst staining. *Bona fide* SP gate could not be identified with verapamil.

3.5.2 Detection of SP cells in HK1 and xeno-284

HK1 cell line used in this study had been previously validated by STR profiling and found to be identical to the HK1 cell line used by others (Chan et al., 2008) (Appendix C). SP phenotype, as identified by low Hoechst 33342 blue/red fluorescence intensity, was detected in 5 - 10% of single, viable HK1 cells (representative dot plot as shown in Figure 3.3).

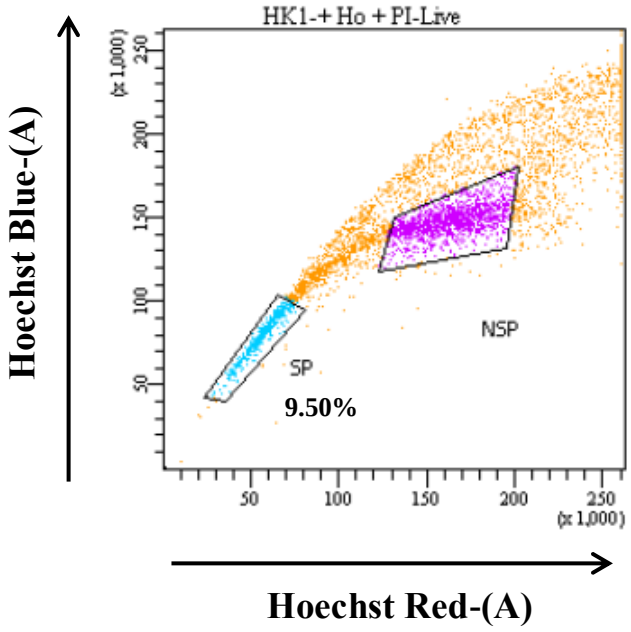


Figure 3.3: Detection of SP and NSP cells with Hoechst 33342 dye in HK1 cells. There were 9.5% of single, viable SP cells in HK1. Non-SP ("NSP") cells were also gated as control cells for downstream experiments.

The ability of these SP cells to efflux Hoechst 33342 dye was inhibited by FTC and not verapamil (representative dot plots as shown in Figure 3.4) which suggests that ABCG2 was the functional ABC transporter in these cells.

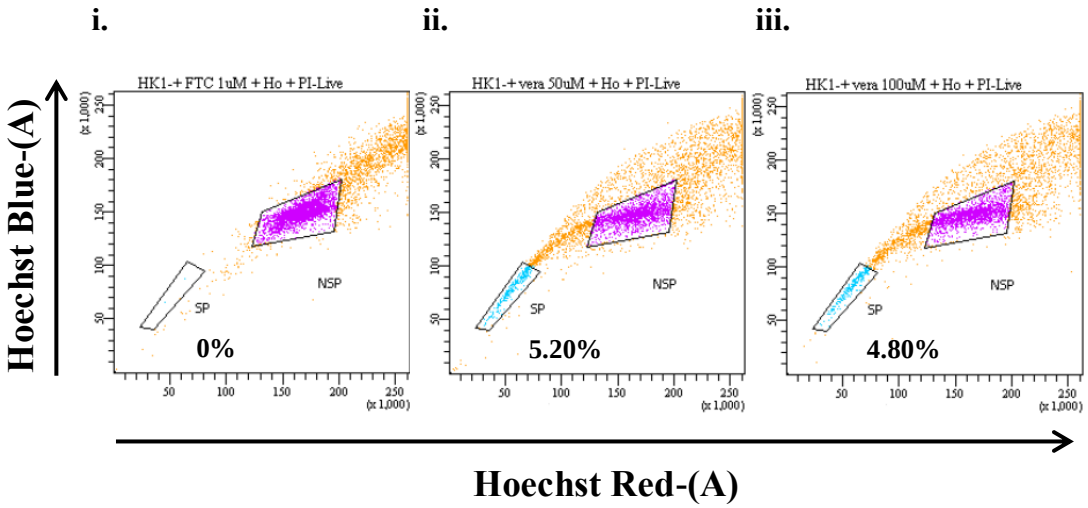


Figure 3.4: Comparison of ABC transporter inhibitors in HK1. **i.** FTC at 1 μM was the suitable ABC transporter inhibitor for HK1, instead of verapamil (50 μM and 100 μM) (**ii** and **iii**).

Compared to HK1, xeno-284 cells had very few (less than 0.5% on average) or no SP cells during repeated runs with different batch of xenograft cells (representative dot plots from two xeno-284 xenografts as shown in Figure 3.5). As such, only HK1 SP and NSP cells were used for subsequent downstream experiments.

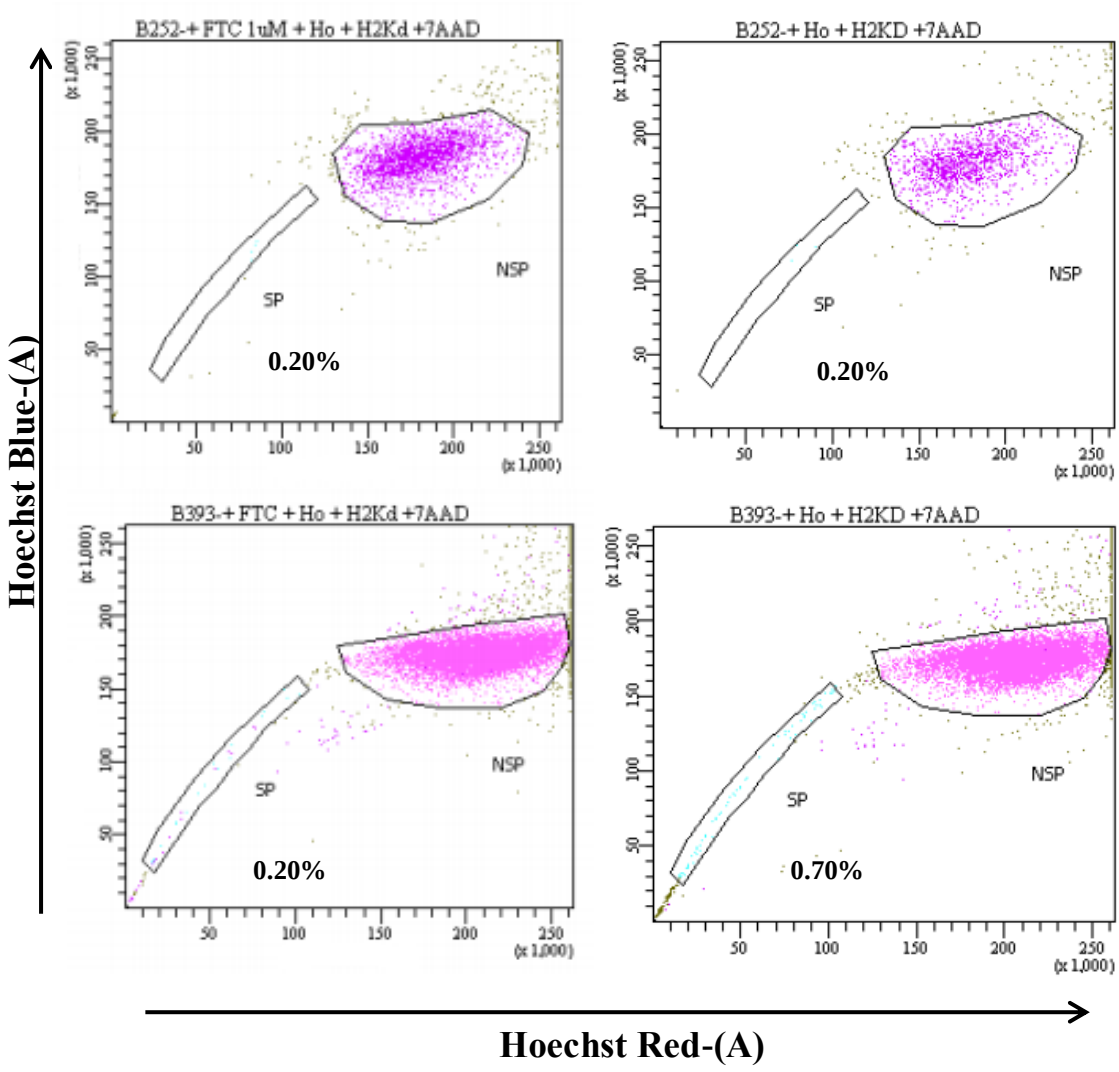
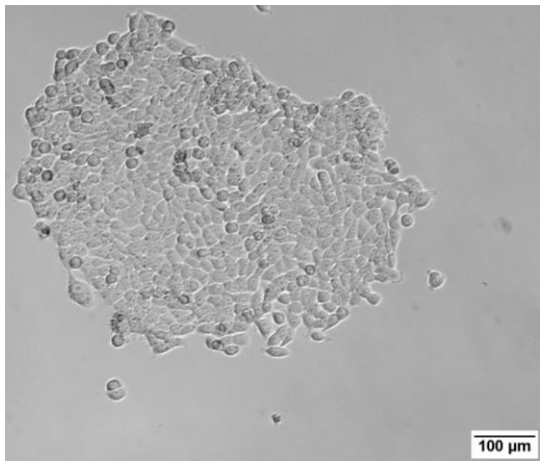


Figure 3.5: Detection of SP and NSP cells in xeno-284 xenografts.

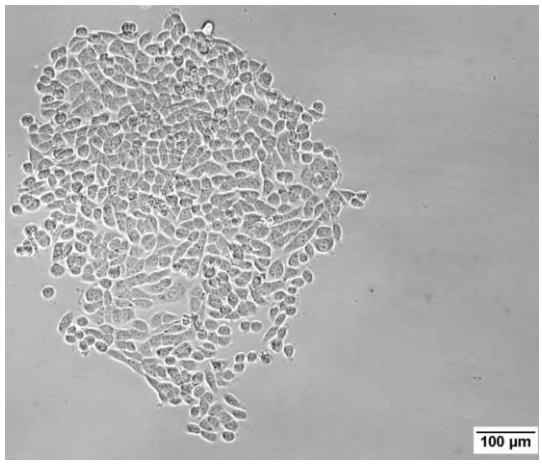
3.5.3 Cell morphology during *in vitro* culture

Post-sorting, HK1 SP and NSP cells exhibited different growth patterns in culture with complete medium. There were three different morphologies observed after the cells attached and started to grow: cells growing into a tight cluster or colony with an overall well-defined clone border which are known as “holoclones”, cells with minimal growth and were widely scattered or growing into small fragmented colonies known as “paraclones”, and cells growing slightly scattered with an overall irregular clone border and showing intermediate structure between holoclones and paraclones which are known as “meroclones” (Figure 3.6) (Harper et al., 2007).

i.



ii.



iii.

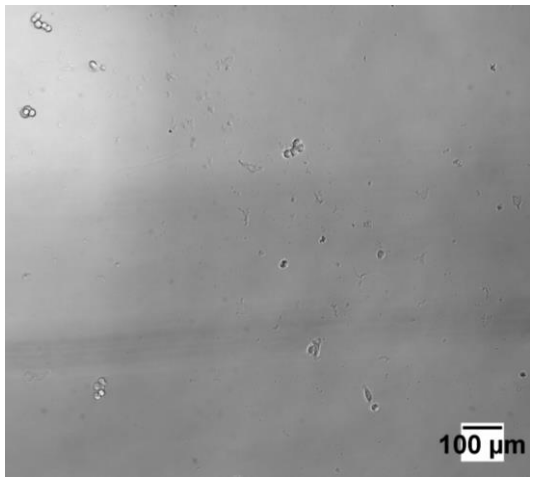


Figure 3.6: Three different clone morphologies displayed by SP and NSP cells *in vitro*. Holoclone (**i**), meroclone (**ii**) and paraclone (**iii**) morphologies in cultured SP and NSP cells of HK1. All photographs were taken on day 9 post-sorting.

After more than a week of *in vitro* culture, most of the SP cells grew into holoclones as compared to NSP cells which primarily formed meroclones and/or paraclones. Repeated experiments showed that SP cells formed more holoclones (6.11 ± 2.89) than NSP cells (2.34 ± 1.57) ($p < 0.0001$) (Figure 3.7).

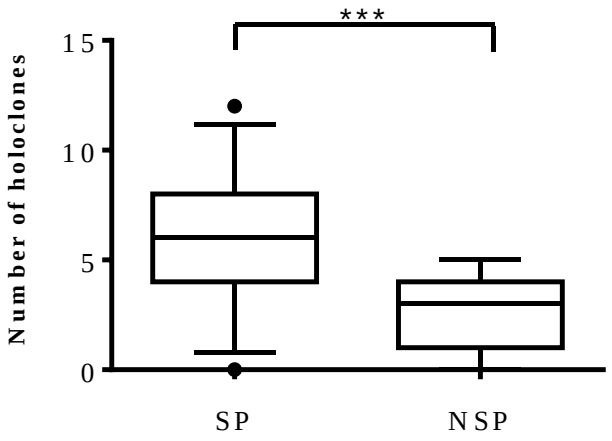


Figure 3.7: Number of holoclones in SP and NSP cells of HK1. SP cells formed holoclones markedly than NSP cells. Results, median (n = 3 sorting experiments with a total of 36 wells per SP or NSP cells). *** $p < 0.0001$

3.5.4 Expression of aldehyde dehydrogenase (ALDH) activity

The enzymatic activity of ALDH1, a marker of stemness, (Alison et al., 2010) can be detected using ALDEFLUOR staining kit. In the presence of ALDH1, BAAA substrate is converted into BODIPY-aminoacetate (BAA) which is retained intracellularly, leading to increased fluorescence which can then be detected by flow cytometry and cell imaging. $ALDH^{bright}$ region in flow cytometry analysis is defined with a "control" sample containing cells incubated with both BAAA and DEAB (an ALDH1 inhibitor).

Figure 3.8 shows that the staining condition for the experiment was initially optimized in SKBR3 breast cancer cell line which was reported to have high ALDH1 expression (Marcato et al., 2011).

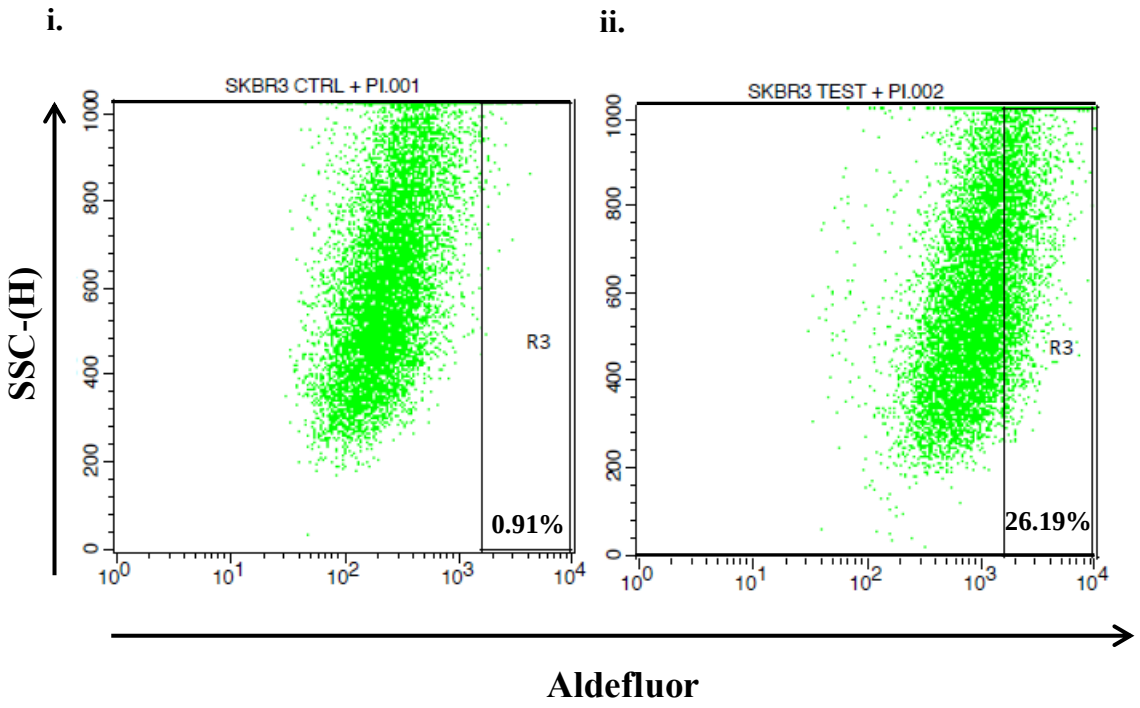


Figure 3.8: Optimization of ALDH staining in SKBR3 by flow cytometry. **i.** SKBR3 cells in ALDH assay buffer containing BAAA and DEAB inhibitor, and viability dye PI ("SKBR3 Ctrl + PI") showed less than 1% of ALDH^{bright} cells ("R3"). **ii.** SKBR3 cells in ALDH assay buffer containing only BAAA and viability dye PI ("SKBR3 Test + PI") showed approximately 26% of ALDH^{bright} cells ("R3"). Net percentage of ALDH^{bright} expression was given by a deduction of "SKBR3 Ctrl + PI" data from "SKBR3 Test + PI" data.

The ALDH^{bright} region for SP and NSP cells was defined with viable parental HK1 stained with both BAAA and DEAB reagents (representative dot plots as shown in Figure 3.9). The ALDH^{bright} gate (R6) was set using these cells and used for SP and NSP cells in the same experiment.

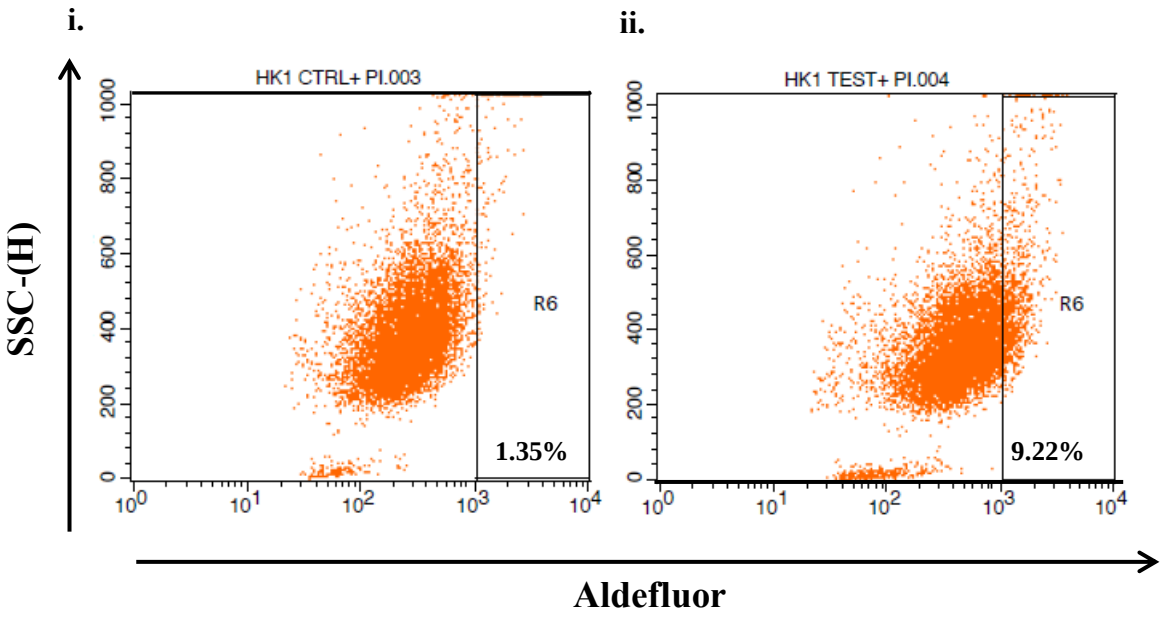


Figure 3.9: Detection of ALDH^{bright} cells in parental HK1 by flow cytometry. **i.** HK1 cells in ALDH assay buffer containing BAAA and DEAB inhibitor, and viability dye PI ("HK1 Ctrl + PI") showed less than 1.5% of ALDH^{bright} cells ("R6"). **ii.** HK1 cells in ALDH assay buffer containing only BAAA and viability dye PI ("HK1 Test + PI") showed approximately 9% of ALDH^{bright} cells ("R6"). Net percentage of ALDH^{bright} expression was given by a deduction of "HK1 Ctrl + PI" data from "HK1 Test + PI" data.

Expression of ALDH^{bright} cells for SP and NSP cells was then determined by each subpopulation individually stained with BAAA and viability dye PI only ("Ctrl +PI" tube), and BAAA, DEAB and PI ("Test + PI" tube) (representative dot plots as shown in Figure 3.10).

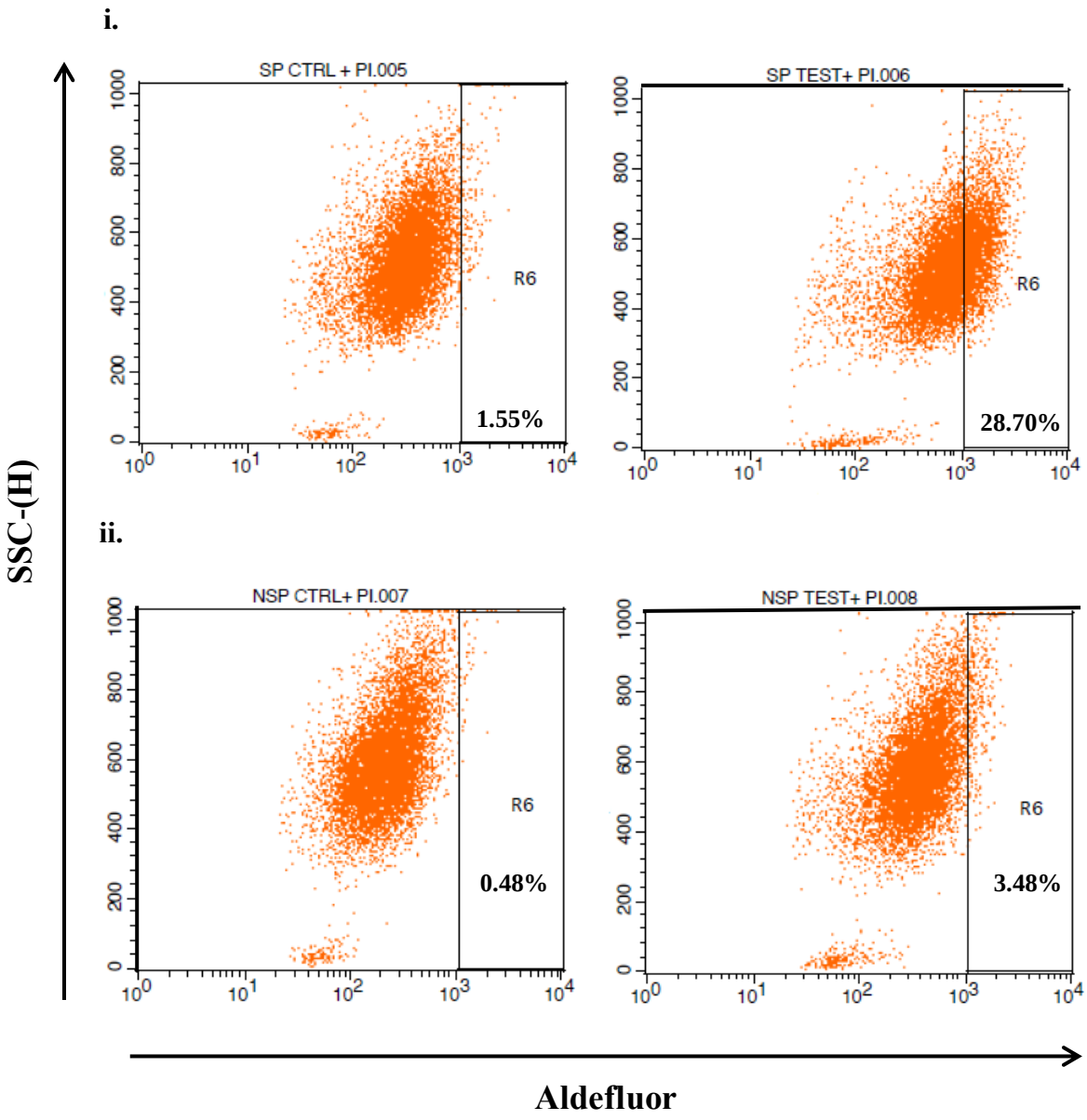


Figure 3.10: Detection of ALDH^{bright} cells in SP and NSP cells by flow cytometry. **i.** SP cells in ALDH assay buffer containing BAAA and DEAB inhibitor, and viability dye PI ("SP Ctrl + PI") showed approximately 1.6% of ALDH^{bright} cells ("R6"). There were approximately 29% of ALDH^{bright} cells in SP cells in ALDH assay buffer containing only BAAA and viability dye PI ("SP Test + PI"). Nett percentage of ALDH^{bright} expression in SP cells was given by a deduction of "SP Ctrl + PI" data from "SP Test + PI" data. **ii.** NSP cells in ALDH assay buffer containing BAAA and DEAB inhibitor, and viability dye PI ("NSP Ctrl + PI") showed approximately 0.5% of ALDH^{bright} cells ("R6"). There were approximately 3.5% of ALDH^{bright} cells in NSP cells in ALDH assay buffer containing only BAAA and viability dye PI ("NSP Test + PI"). Nett percentage of ALDH^{bright} expression in NSP cells was given by a deduction of "NSP Ctrl + PI" data from "NSP Test + PI" data.

The SP cells showed a significantly higher population of ALDH^{bright} cells (18.08 ± 11.46%) than NSP cells (5.10 ± 3.56%) ($p < 0.05$) (Figure 3.11).

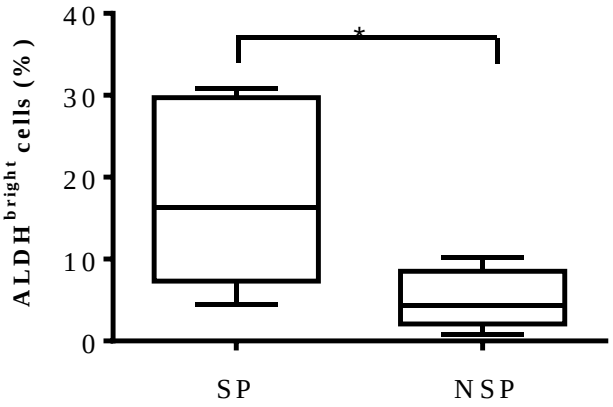


Figure 3.11: Expression of ALDH^{bright} cells in SP and NSP cells. Although SP cells contained significantly higher ALDH^{bright} cells than NSP cells, there was a greater variation amongst its replicates than in NSP cells. Results, median (n = 5 flow cytometry experiments). * $p < 0.05$

3.5.5 Mode of cell division during *in vitro* culture

In order to determine if the sorted cells were able to divide symmetrically or asymmetrically *in vitro*, freshly-sorted SP and NSP cells were cultured in complete medium for 21 days with fresh medium change every 2 - 3 days. The cultured cells were subjected to Hoechst 33342 staining and SP cell analysis as described earlier.

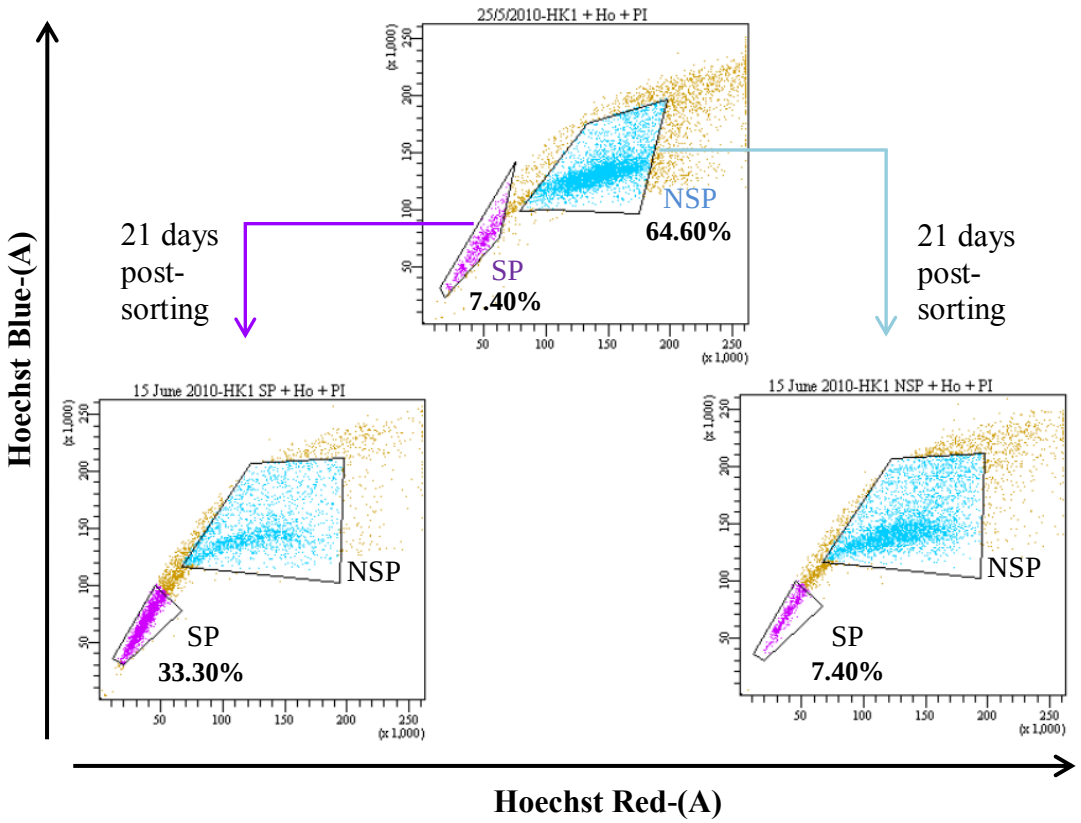


Figure 3.12: Analysis of SP and NSP fractions in sorted cells after short-term culture. SP cells (purple) were collected and cultured for 21 days, prior to re-staining with Hoechst 33342 dye and viability dye PI for an analysis of SP cell percentage. Similarly, NSP cells (light blue) were collected and cultured for 21 days prior to a repeat staining round with Hoechst 33342 dye and viability dye PI for an analysis of SP cell percentage. Results from 1 representative experiment are shown here.

Each subpopulation of sorted cells had divided into both SP and NSP fractions upon re-analysis (representative dot plots as shown in Figure 3.12). SP sorted cells had divided into 24.57 ± 7.97% SP fraction with remaining subpopulation comprising of NSP fraction. NSP sorted cells had largely remained as NSP fraction with only 6.07 ± 1.74% SP fraction.

3.5.6 Proliferation during *in vitro* growth

Proliferation or growth rates of SP and NSP cells *in vitro* were measured using an impedance-based experiment. SP cells had the slowest proliferation in all replicate

experiments with NSP cells having the highest growth rate (representative growth curves as shown in Figure 3.13).

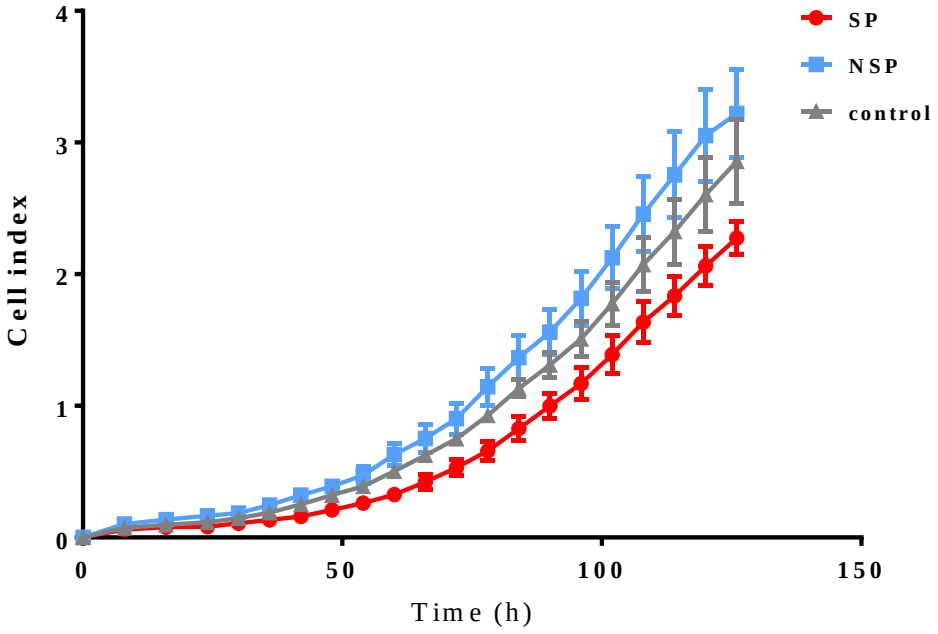


Figure 3.13: Growth curves of SP, NSP and control HK1 cells cultured in complete medium for 125 hours. SP cells proliferated slower than NSP and control subpopulations. Results, mean ± SD (n = 4 wells per group from 1 sorting experiment).

Normalized growth rates were calculated by dividing the cell index of SP or NSP cells at each recorded time point with the cell index of the control cells in order to rule out any bias due to technicality. There was a significant difference in the average normalized growth rates between SP and NSP cells ($p < 0.0001$) (Figure 3.14). SP cells grew at a slower rate as compared to NSP cells especially at the first 72 h of observation. At the end of the experiment, the growth rates of SP and NSP cells were getting similar to the growth rate of control cells.

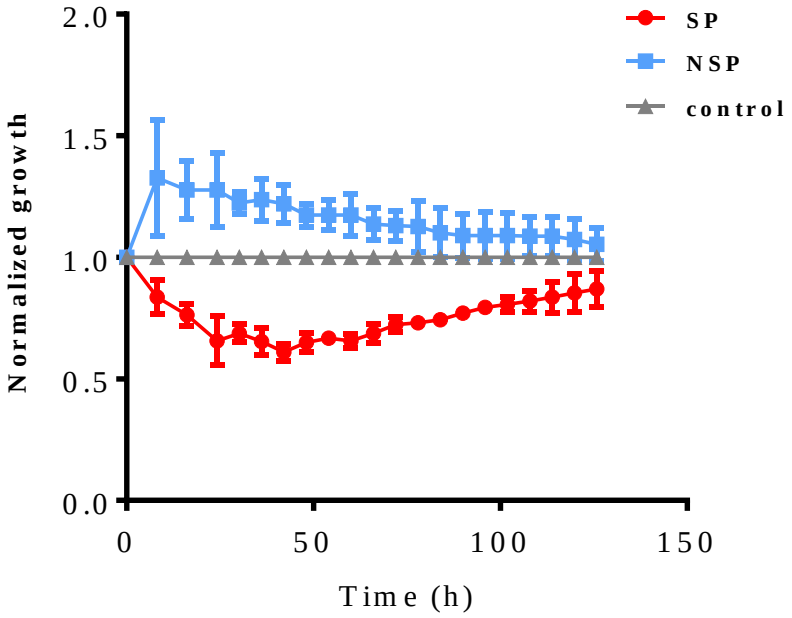


Figure 3.14: Average normalized growth curves of SP and NSP cells. Growth curves of SP and NSP cells were normalized against the growth curve of control HK1 cells in each experiment. Results, mean ± SD (n = 3 sorting experiments).

3.5.7 Relative quantification of stem cell and stem cell signalling genes

A total of 168 genes related to stem cell identification and/or involved in key stem cell pathways were analysed by quantitative RT-PCR. Fifty genes were significantly upregulated by at least 2 folds in SP cells compared to NSP cells ($p < 0.05$) (Appendix D). None of the genes tested was significantly downregulated in SP cells.

Stem cell genes which were upregulated in SP cells included *TERT* (5.44 folds), *MYC* (3.01 folds), *SOX2* (2.87 folds) and *ABCG2* (2.63 folds) ($p < 0.05$) (Figure 3.15).

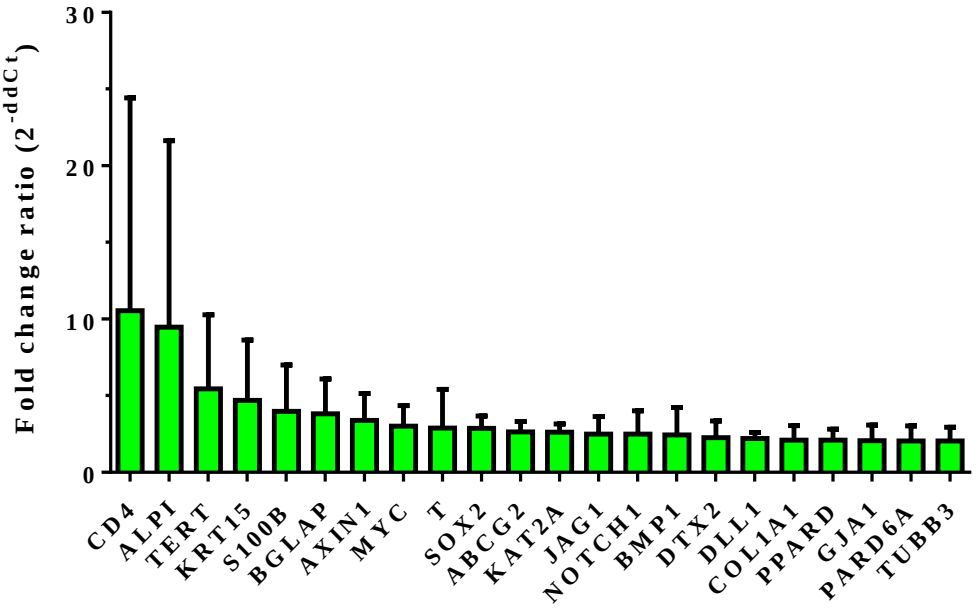


Figure 3.15: Relative expression of stem cell genes which were significantly deregulated and of more than 2-fold change in SP cells. Fold change ratio was expressed as normalized expression of SP cells versus normalized expression of NSP cells. Results, mean with 95% CI (n = 3 sorting experiments). $p < 0.05$

Members of commonly deregulated signalling pathways in cancer were also upregulated significantly more than 2 folds in SP cells: they included *GLI1*, *GLI2*, *GLI3*, *PTCHD2*, *SUFU* and *PTCH1*of the Hedgehog pathway, *NOTCH1* and *JAG1* of the Notch family, *LTBPs2-3*, *ACVR1B*, *SMAD1* and *SMAD7* of the TGFβ superfamily and *BCL9*, *FZD6*-7, *FZD2* and *BCL9L* of Wnt signalling pathway ($p < 0.05$) (Figure 3.16).

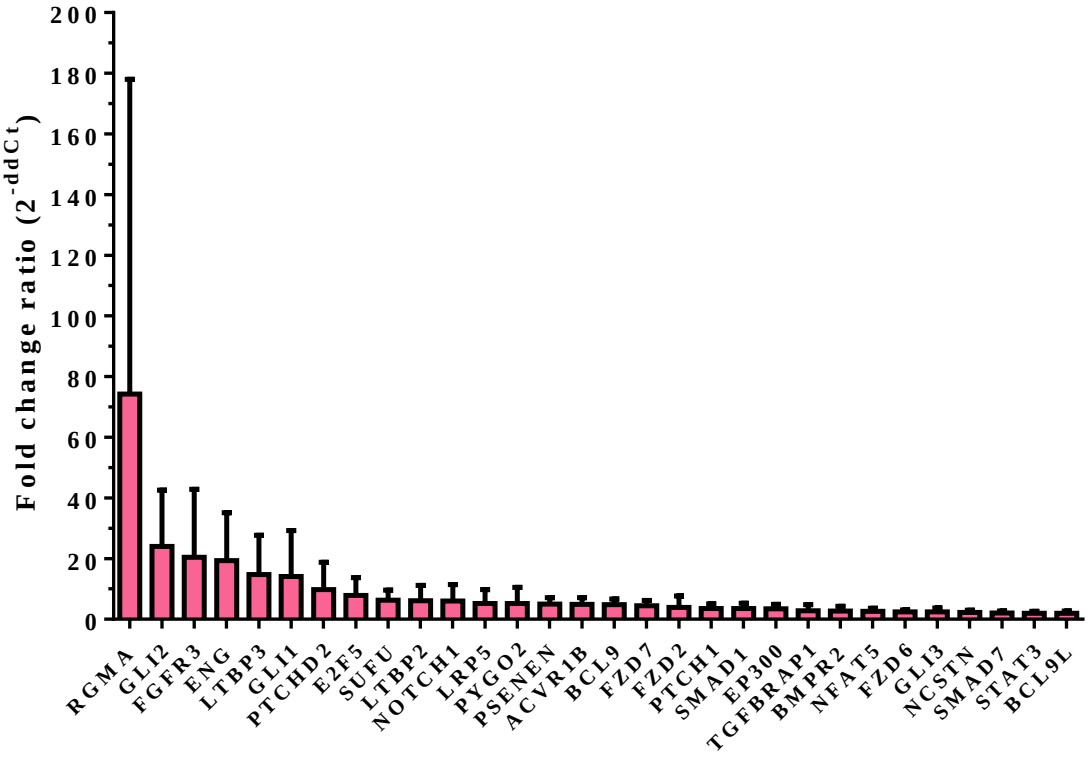


Figure 3.16: Relative expression of stem cell signaling genes which were significantly deregulated and of more than 2-fold change in SP cells. Fold change ratio was expressed as normalized expression of SP cells versus normalized expression of NSP cells. Results, mean with 95% CI (n = 3 sorting experiments). $p < 0.05$

3.5.8 Tumourigenicity of SP and NSP cells

Cells from SP subpopulation and NSP subpopulation were inoculated subcutaneously into nude mice to determine their tumour-forming ability. Figure 3.17 shows representative average growth curves of tumours initiated by 1,000, 100 and 10 SP and NSP cells. Overall, SP cells formed faster-growing and larger tumours than NSP cells.

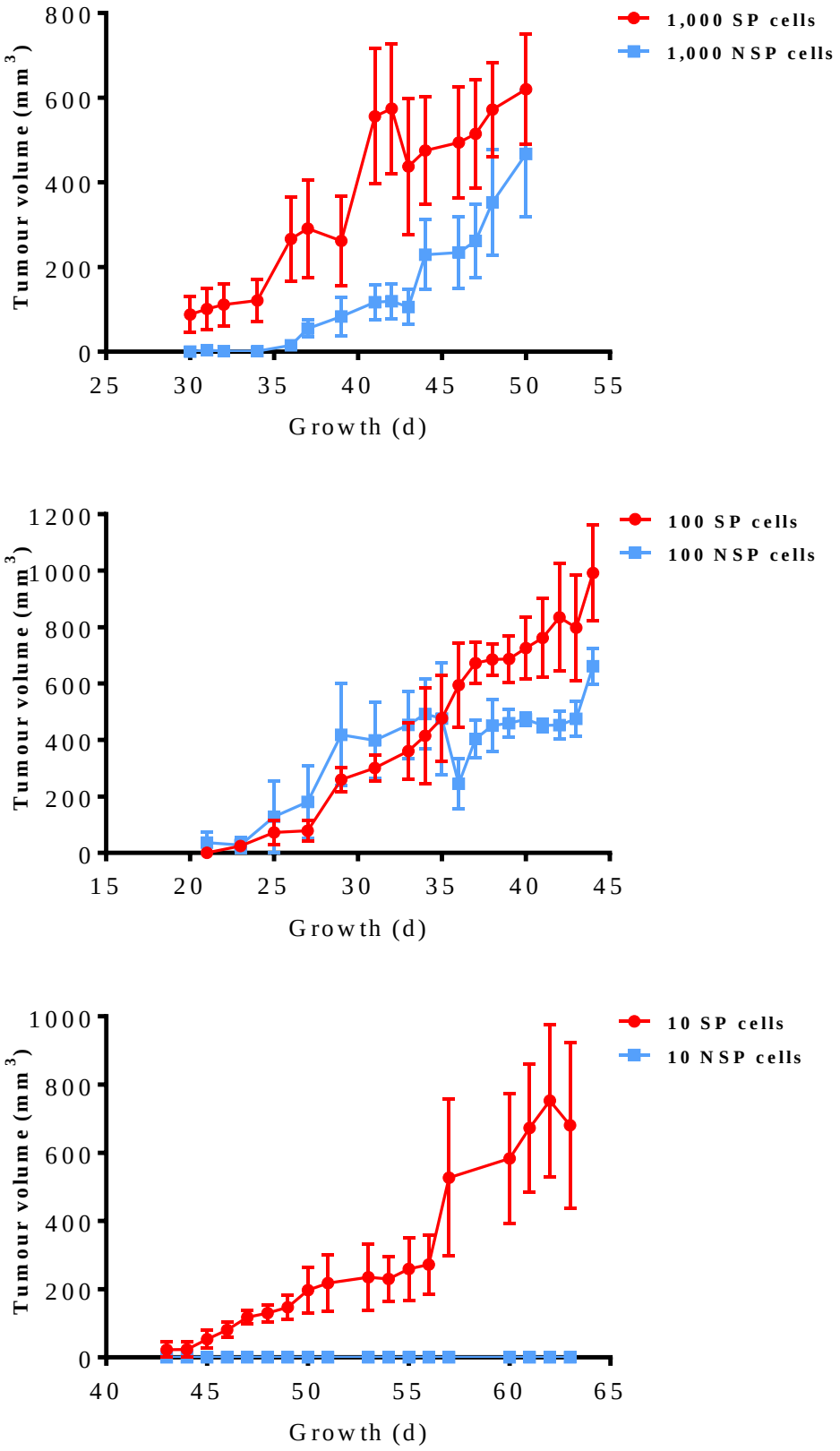


Figure 3.17: Growth curves of nude mice inoculated with 1,000, 100 and 10 SP and NSP cells. Tumour growths were measured every 1 - 3 days and all mice were terminated as described in Methods. Results, mean ± SEM (n = 3 or 4 mice per group from 1 sorting experiment for each inoculation).

Tumour latency data showed that 3 out of 4 tumours from the inoculation of 10 SP cells were detected earlier than the first tumour arising from the inoculation of NSP cells of the same number (Figure 3.18). This was suggestive of SP cells possessing a growth advantage over NSP cells; however, it was lost at higher inoculations of 100 and 1,000 cells.

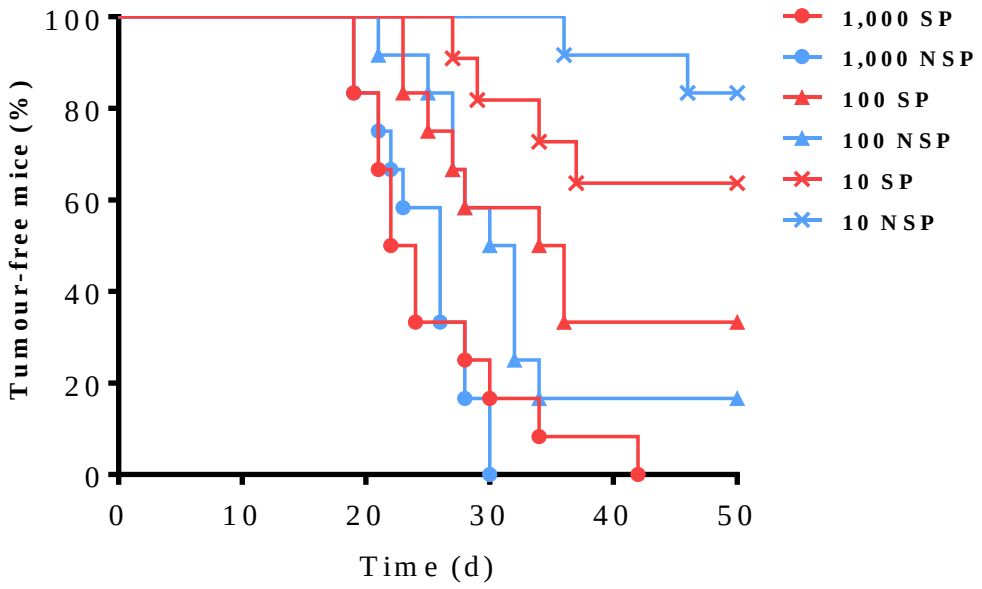


Figure 3.18: Tumour latency data of SP and NSP cells. Disparity in the latency data was most evidently seen at the lowest inoculation number of 10 cells, whereby 4 out of 11 SP mice had earlier tumour growths as compared to only 2 out of 12 NSP mice which had growth at a later period (n = 3 sorting experiments with a total of 12 mice per inoculation group, except for "10 SP" group which had 11 mice).

An estimation of number of TICs by ELDA software failed to reveal significant differences between the SP and NSP subpopulations, implying that the difference in the potential of SP and NSP cells to form tumours was not significant ($p > 0.05$) (Table 3.1).

Table 3.1: Tumour-initiating cell (TIC) frequency of SP and NSP mice inoculated with a limiting dilution of 1,000, 100 and 10 cells.

Phenotype	Number of tumours/number of mice			TIC frequency (95% CI)	p-value
	1,000 cells	100 cells	10 cells		
SP	12/12	8/12	4/11	1 in 66 (1/35 – 1/126)	0.7
NSP	12/12	10/12	2/12	1 in 56 (1/30 – 1/104)	

There was no visible difference between gross morphology of SP and NSP tumours (representative photographs as shown in Figure 3.19). Regardless of SP or NSP tumours, they were either mostly encapsulated or slightly attached to surrounding skin nearly all of them were reddish and contained necrosis of various degrees.

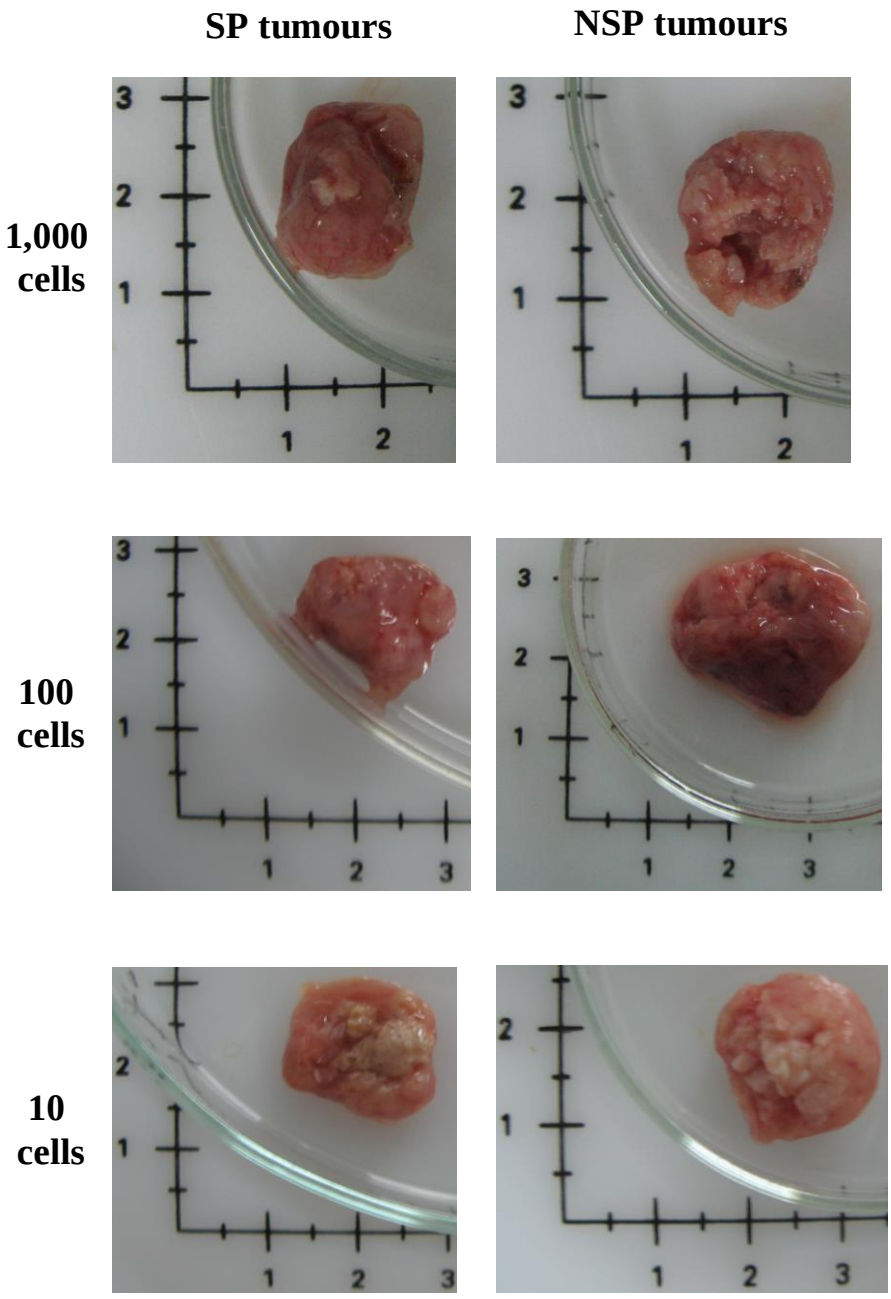


Figure 3.19: Photographs of excised tumours initiated from SP cells and NSP cells of 1,000, 100 and 10 cell inoculations.

Both SP and NSP tumours showed similar histomorphology as seen in unsorted HK1 cells grown in mice (Figure 3.20). There was also no substantial difference between SP and NSP tumours in the degree of differentiation, stromal reaction and cell

pleomorphism, except for one SP tumour which showed vascular invasion (Figure 3.20iv).

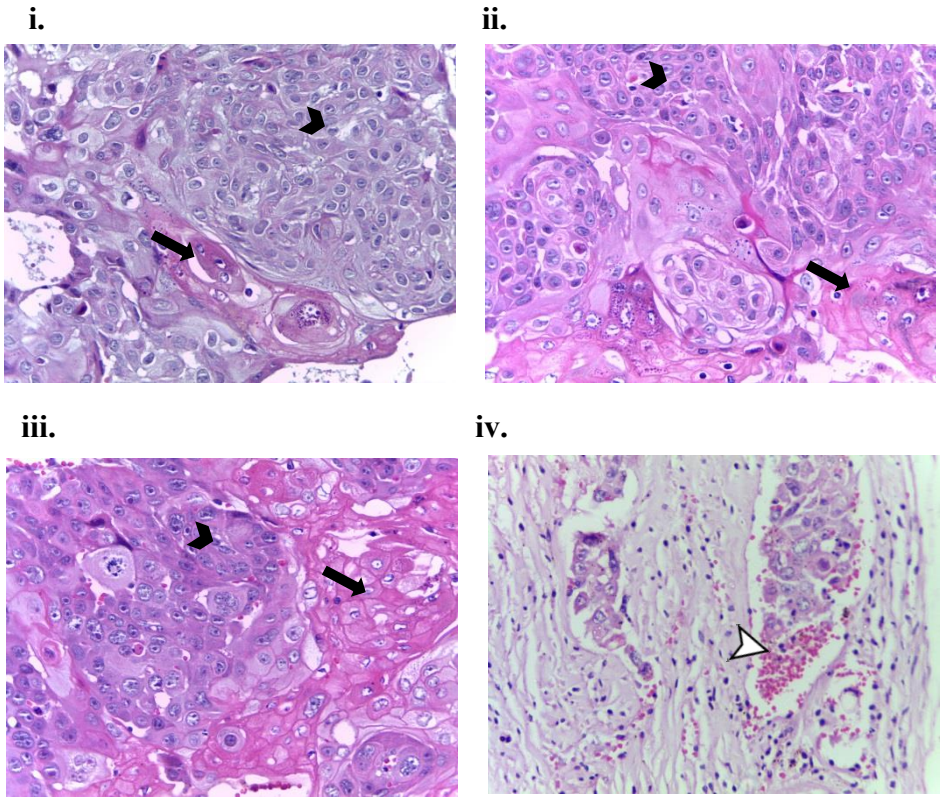


Figure 3.20: Photographs of H&E staining of resulting tumours after inoculation with **(i)** unsorted HK1 cells, **(ii)** SP cells and **(iii)** NSP cells. Tumours derived from SP and NSP cells showed similar histomorphology to unsorted HK1 cells (black arrowhead pointing to squamous carcinoma; long arrow pointing to keratinization). **iv.** A SP cell-derived tumour showed vascular invasion (white arrowhead pointing to the invasion of SP cells into a blood vessel) (all photos with objective 40X).

3.6 Discussion

With a growing plethora of antibodies and dyes to identify cancer cells as well as the advances in flow cytometry and cell sorting techniques to isolate and purify multiple subpopulations of cells from a single sample, cell biologists of present day are able to study the roles, functions and abilities of heterogeneous cancer cells ranging from tumour formation, cell division, disease metastasis to therapy resistance, with more ease and accuracy. SP assay which depends on the presence of ABC drug transporter(s) to

efflux fluorescent dyes is one of the approaches frequently adopted to identify and isolate cancer cells with putative stemness qualities. These cells are capable of initiating tumours at a higher rate, possess slower or faster proliferation, can form more colonies or spheroids *in vitro*, and are able to survive radiation and/or chemotherapeutic drugs (Golebiewska et al., 2011; Wu & Alman, 2008).

Both HK1 and xeno-284 cells were established from patients with recurrent NPC (Huang et al., 1980; manuscript in preparation). Cancer recurrence is believed to be associated with the resistance of cancer cells to treatment, of which overexpression of ABC drug transporters is one of the main mechanisms of evasion to chemotherapy (Kathawala et al., 2015). Hence, HK1 and xeno-284 cells were used for SP assay in this study to isolate and characterize biological differences of SP and NSP subpopulations.

Two main criticisms of SP assay, namely Hoechst 33342 toxicity and variations in data due to differences in gating strategy (Golebiewska et al., 2011; Montanaro et al., 2004; Wu & Alman, 2008), were also addressed in this study. Firstly, Hoechst 33342 was used at a final concentration of 5 μM which could be considered as the "standard" concentration in SP assay and had minimal, if any, toxicity effects on stained cells (Broadley et al., 2011; Goodell et al., 1996; Luo et al., 2012; Yi et al., 2015). Although HK1 NSP cells lacked of the ABC drug transporters which are responsible for pumping out Hoechst 33342 dye from the cells, they did not suffer from the retained dye as evidently seen in the *in vitro* cell morphology, cell division and proliferation experiments, in addition to the aptness of NSP cells to form tumours in nude mice. Secondly, the position of SP gate is largely dependent on the type of ABC drug transporter present in the sample; as such, both FTC and verapamil were initially tested at different concentrations in HK1. FTC at a final concentration of 1 μM was able to block nearly 100% of dye-pumping action of a subpopulation of HK1, as opposed to the

inability of verapamil to do so at 50 and 100 μM. Also, a conservative SP gating strategy was used in this study to ensure only cells with the lowest Hoechst intensity were gated as "SP" cells and cells with relatively high Hoechst intensity as "NSP" cells. This is to reduce the likelihood of including other subpopulations of cells with varying abilities to efflux the dye, as was reported in Montanaro et al. (2004).

The bulk of SP publications on identification of tumourigenic cells, and thereafter investigation of their biological properties, used cancer cell lines as their study models (Fukaya et al., 2010; Kato et al., 2010; Nakayama et al., 2014; Salcido et al., 2010). This may be primarily due to technical difficulties in establishing PDXs, leading to their rarity. Percentages of SP cells from PDXs are generally much lower than cell lines, as were similarly seen in prostate tumour and melanoma PDXs (Luo et al., 2012; Patrawala et al., 2005). Both studies used an enzyme digestion cocktail containing collagenase which was also used in this study. One of the prostate tumour PDXs had undetectable presence of SP with another registering a 0.07% SP presence (Patrawala et al., 2005). Although the melanoma PDX model had a slightly higher presence of SP cells (0.34%) (Luo et al., 2012), the content was still a few folds lower than cancer cell lines which typically had a reported range of >1 to 10% SP cells.

Due to the low and inconsistent presence of SP cells in xeno-284, HK1 was used as a study model to discern the biological properties of SP and NSP subpopulations in *in vitro* and *in vivo* experiments. The range of SP cells detected in the HK1 cells used in this study (5 - 10%) was similar to a later publication on SP cells in NPC cell lines which reported an approximate 5.20 ± 0.72% SP presence in their HK1 cells (Zheng et al., 2016). SP gate was however determined with the use of verapamil at a final concentration of 50 μM in the latter. One plausible explanation for the difference in dye efflux inhibition efficacy of verapamil in this and Zheng and colleagues' studies is the

different source of HK1 cells: University of Hong Kong (this study) and Chinese University of Hong Kong (Zheng et al., 2016). SP presence in HK1 cells was also comparable to those detected in the primary cultures of NPC cells from NPC patients (approximately 3.9%) (Wang et al., 2015).

Results from *in vitro* experiments indicated that the SP cells of HK1 were more stem-like than NSP cells. Both SP and NSP cells exhibited different clone morphology when grown in complete growth medium. SP cells tended to form holoclones as opposed to NSP cells which mostly grew as meroclones and/or paraclones. The association of holoclones and stemness properties, in particular self-renewal, has been widely accepted in CSC studies in which its formation was often regarded as a read-out of CSC phenotype (Li et al., 2008; Locke et al., 2005; Roudi et al., 2016; Shigeishi et al., 2013). One caveat of using *in vitro* approach to identify CSCs is the sub-optimal physiological culture conditions do not fully support or provide factors such as three-dimensional structure and environment for *in vivo* tumour formation and self-renewal. Hence, holoclones need to be further assayed for their ability to induce tumour formation followed by serial-passaging of the induced tumours for a thorough validation of CSC identity.

ALDEFLUOR assay buffer used in the ALDH experiment contained an efflux inhibitor which prevented active efflux of the fluorescent BAA product from stained cells, thus allowing cells with high ALDH activity to be detected (Alison et al., 2010). As such, $ALDH^{bright}$ cells could be regarded synonymous with SP cells as the mechanism of dye pumping accorded by the presence of ABC drug transporters is the common backbone of both ALDH and SP assays. This was clearly evident in this study where HK1 SP cells were significantly enriched with ALDH activity. A later publication using RT-qPCR approach concurred with the data in this study as it also

found that SP cells had higher expression levels of *ALDH1A1* gene than NSP cells in HK1 ($p < 0.01$) (Zheng et al., 2016). Ishiguro et al. (2016) reported that $ALDH^{high}$ cells from ovarian cancer patients showed stem-like features as they contained higher levels of SOX2, OCT3/4 and NANOG proteins and were able to grow as spheroids more efficiently than $ALDH^{low}$ cells. Most importantly, their study showed that $ALDH^{high}$ cells were able to form xenografts but not $ALDH^{low}$ cells.

The type of ABC drug transporter present in HK1 was further validated with a more than 2-fold upregulation of *ABCG2* expression in SP cells. This was also similarly reported in the other HK1 publication on SP cells (Zheng et al., 2016). Although several types of transporter can be expressed within the same cell or within distinct SP subpopulations (Golebiewska et al., 2011), the possibility of a presence of ABCB1-expressing cells in HK1 remains remote as dye-pumping activity in these cells was not inhibited even at a high verapamil concentration of 100 μM.

Gene expression data indicated that SP cells of HK1 displayed stemness signature. These cells showed increased mRNA levels for *SOX2*, *MYC* and *TERT*, which were associated with CSC biology and maintenance of pluripotency state (Chen et al., 2012; Wu et al., 2013b). Over-expression of Hedgehog pathway genes such as *GLI1*, *GLI2*, *GLI3FL*, *PTCH1* and *PTCHD2* was in line with reports that Hedgehog pathway can induce *ABCG2* expression (Singh et al., 2011) and that members of the pathway such as GLI1, GLI2, GLI3 and PTCH1 had increased mRNA and protein expression in stem-like cells in EBV positive NPC (Lun et al., 2012; Port et al., 2013).

Nonetheless, there were other findings in this study which suggest that the stem-like characteristics of SP cells in HK1 may be transient. SP cells of HK1 grew slower in culture which concurred with other reports that CSCs from various tumours had slower proliferation rate, thus enabling these cells to evade chemotherapeutic agents targetting

fast growing cells (Buczacki et al., 2011; Dembinski & Krauss, 2009). However, the slower growth rate of SP cells of HK1 gradually diminished after nearly a week *in vitro* and would be similar to that of NSP cells if the growth observation had been prolonged (Figure 3.16). A separate study on SP cells from CNE-2 found that SP fraction grew faster than NSP fraction *in vitro* despite containing similar proliferation index and cell cycle content (Wang et al., 2007). However, the authors did not give an explanation for this observation. Also, NSP cells gave rise to SP cells in long term culture (Figure 3.13). Altogether, these two time course *in vitro* experiments on proliferation and cell division indicated that the difference between SP and NSP sorted cells may be transient, as both phenotypes could give rise to each other as a result of asymmetric division. Phenotype interconversion during culture was similarly reported in previous SP studies on glioma, NPC and renal cell carcinoma cell lines (Huang et al., 2013; Platet et al., 2007; Wang et al., 2007).

The transient stem-like phenotype displayed by SP cells of HK1 would explain for the discordance of *in vitro* results with *in vivo* data. Although SP cells were generally more tumourigenic (i.e. formed faster-growing tumours in addition to a growth advantage in the lowest inoculation of 10 cells), an overall comparable tumour-initiating potential of both SP and NSP cells was recorded after seven weeks of *in vivo* transplantation with similar TIC frequency for both SP and NSP cells. These two seemingly mismatched observations of growth rate and TIC frequency may be attributed to the presence of TICs with differential growth rates in both SP and NSP cells. The discordance of *in vitro* and *in vivo* results had been reported in studies utilizing SP approach to identify CSCs in thyroid and cervical cancers (Mitsutake et al., 2007; Qi et al., 2014) as well as studies using surface marker approach to identify CSC in NPC and colorectal cancer (Peickert et al., 2012; Zhang et al., 2012a). Moreover, a prior NPC study reported that ALDH positive cells of 5-8F and CNE-2 cell lines

showed stemness features including higher tumour-initiating ability than ALDH negative cells (Wu et al., 2013a). Based on this finding, the ability of NSP cells of HK1 to generate tumour growth in this study may be also partially contributed by the 5% presence of $ALDH^{bright}$ cells in them.

The negative *in vivo* findings suggested that SP approach in the study model of HK1 only enriched for a subtype of tumourigenic cells with stem-like or partial stemness features but not *bona fide* CSCs as there was no significant difference in tumour initiation ability between SP and NSP cells. This is in contrast to an earlier report that showed significant tumour growth difference of SP cells over NSP cells in NPC (Wang et al., 2007). The disparity may be related to firstly, the type of NPC cell line being studied. Wang et al. (2007) utilized CNE-2 cells which had been reported to be contain a partial genome of HeLa, whereas HK1 NPC cells used in this study was verified by STR profiling to be free from HeLa cell contamination (Chan et al., 2008). Secondly, FTC was used in this study as a dye efflux inhibitor to identify ABCG2-specific SP cells from HK1 and xeno-284 cells, instead of verapamil, a classic inhibitor for ABCB1-expressing cells (Duan et al., 2004; Robey et al., 2007) which showed no inhibition in this study but was used in Wang and colleagues' study. Lastly, the length of tumour assessment time was different. Mice inoculated with CNE-2 SP and NSP cells were euthanized four weeks after inoculation, whereas mice were observed for seven weeks in this study. Although both studies did not employ long period of observation, preferably for more than eight weeks to prevent under-estimation (Quintana et al., 2008), there is a higher likelihood of under-reporting for CNE-2 NSP cells with an observation period of only four weeks in Wang and colleagues' study than in this study.

3.7 Conclusion

This study shows the presence of dye-extruding SP subpopulation, a characteristic of stem cells, in HK1 (a cell line generated from well-differentiated recurrent NPC) but they were minimal in xeno-284 cells (a PDX derived from poorly differentiated recurrent NPC in our lab). Additionally, HK1 SP cells were ABCG2-specific. In spite of SP cells growing slower and possessing more traits of stemness *in vitro* (holoclone formation, ALDH activity, asymmetric division and upregulation of stem cell genes and associated pathways) than NSP cells, *in vivo* validation study showed that both SP and NSP cells of HK1 had similar tumourigenic potential. Therefore, it is concluded that the SP approach alone cannot identify for CSCs accurately in the study model of HK1.

CHAPTER 4: CD24, CD44 AND EPCAM ENRICH FOR TUMOUR-INITIATING CELLS IN A NEWLY ESTABLISHED PATIENT-DERIVED XENOGRAFT OF PRIMARY NASOPHARYNGEAL CARCINOMA

4.1 Introduction

The use of CSC markers as my second approach to identify, isolate and study NPC tumourigenic and non-tumourigenic cells is presented and discussed in this chapter. HK1, xeno-284 cells, C666-1 cell line derived from *in vitro*-passaged xenograft cells of an undifferentiated NPC biopsy (Hui et al., 1998) and xeno-B110 established from a non-keratinizing differentiated NPC biopsy were first immunophenotyped for CD24, CD44 and EpCAM expression levels using flow cytometry. Isolation of CD24, CD44, EpCAM, and combination of EpCAM and CD44 bright and dim cells from C666-1 and xeno-B110 was carried out for tumour initiation (incidence and latency), growth and mitotic activity assessments. Due to marked growth differences seen in xeno-B110 marker-selected cells, cell cycle and gene expression analyses were also performed in these cells. Finally, self-renewal (a determinant of stemness) was further evaluated in xeno-B110 marker-selected cells by conducting serial transplantation passages *in vivo* for three cycles.

4.2 Brief Literature Review

Cancer cells are known to be made up of heterogeneous cell types with unequal ability to form tumours and it is strongly believed that tumour formation is associated to the presence of a subpopulation of cancer cells called cancer stem cells (CSCs) which lie at the apex of tumour cell hierarchy (De Sousa E Melo et al., 2013; Shackleton et al., 2009). Besides having the common characteristics of tumourigenic cells such as the abilities to initiate and drive tumour formation and propagation, to evade immune surveillance, to resist therapy and having altered genetic makeup (Chang, 2016; Hanahan & Weinberg, 2011; Krampitz et al., 2016), CSCs have self-renewal ability

which ensures the perpetuality of tumour generation, and the ability to differentiate and give rise to heterogeneous lineages of cells which constitute the tumour (Clarke et al., 2006; Hanahan & Weinberg, 2011).

The use of CSC markers has aided in the isolation of tumourigenic and non-tumourigenic cells directly from clinical samples for the study of their roles in tumourigenesis. For example, $EpCAM^{high}CD44^{+}$ marker identified cells which were responsible for successful engraftments of six human colorectal carcinoma samples in immunodeficient mice and these patient-derived xenografts (PDXs) had similar heterogeneity to their original patient tumours (Dalerba et al., 2007). In non-small cell lung carcinoma samples, $CD166^{+}$ marker identified tumour-initiating cells (TICs) which consistently form tumours *in vivo* and $CD166^{+}$ cells were able to self-renew to initiate subsequent PDXs for at least two generations (Zhang et al., 2012b). However, there is a dearth of such studies using clinical samples in NPC due to sample size limitation as surgery is not being the mainstay treatment modality (Chan, 2010; Lee et al., 2015). Hence, NPC biological-based studies were mostly performed on NPC cell lines which have been in passage for many generations (Lun et al., 2012; Su et al., 2011; Yang et al., 2014b) or primary culture-derived xenograft (Yang et al., 2013, 2015), some with questionable authenticity and/or origin as were reported for CNE-1, CNE-2 and HONE-1 (Chan et al., 2008; Strong et al., 2014). Cell line misidentification and/or cross-contamination are universal issues which also occur in other cancers (MacLeod et al., 1999). To circumvent this, early-passage PDXs which have close resemblance to clinical samples, are advocated to be used as study models for tumour heterogeneity, CSCs and therapy-related studies, instead of cell lines (Clevers, 2011; Dodbiba et al., 2015; Julien et al., 2012; Visvader & Lindeman, 2012).

The use of CD24 marker to identify and to isolate subpopulations of NPC for biological-based investigations especially on stem-like characteristics has been minimally reported. Kondo et al. (2011) reported that $CD44^{high}CD24^{low}$ subpopulation from AdAH cells (an EBV-negative nasopharyngeal epithelial cell line) transfected with LMP1 showed the ability to form more spheres *in vitro* than untransfected control cells. As the study focused on the role of LMP1 in inducing CSC phenotype in nasopharyngeal epithelial cells, there was no comparative analysis on the efficiency of $CD24^{low}$ and $CD24^{high}$, and $CD44^{low}$ and $CD44^{high}$ cells in sphere formation, a read-out of self-renewal ability *in vitro* (Pastrana et al., 2011). $CD24^{+}$ cells showed stem-like phenotype in NPC cell lines TW02 and TW04 on the basis of having higher levels of pluripotent markers of *OCT4* and *SOX2* as well as of stemness markers of *NANOG*, *BMI1* and *REX1* (Yang et al., 2014b). Functionally, $CD24^{+}$ cells from both cell lines were also able to form tumours at low inoculations of 500 and 1,000 cells (overall TIC frequency of 1 in 435 for TW02 and 1 in 540 for TW04) as compared to no tumour growth of same number of $CD24^{-}$ cells, besides showing more enhanced *in vitro* proliferation and higher resistance to cisplatin or docetaxel. Serial transplantation of $CD24^{+}$ cells was not performed to determine self-renewal ability. In a separate NPC study using hybrid HONE1 cells (a NPC cell line), the presence of CD24 was also associated with stem-like features as there was an enhanced presence of $CD24^{+}$ cells and expression of pluripotency genes such as *SOX2*, *KLF4*, *OCT4* and *NANOG* (Cheng et al., 2013).

Based on the latest reviews on CSC markers in NPC cell lines, CD44, an extracellular receptor for hyaluronan, seems to be the most widely studied marker with roles ranging from tumour initiation, cell proliferation and differentiation to 5-fluorouracil (5-FU) treatment resistance (Lun et al., 2014; Yu & Loh, 2014). $CD44^{+}$ cells from NPC cell line, SUNE-1 5-8F were presumed to be CSCs due to having

significantly increased *BMI1* and *OCT4* transcript levels, higher proliferation rate as well as resistance to radiation and chemotherapy treatments *in vitro*; however, the tumour-initiating ability of $CD44^+$ cells was not verified in this study (Su et al., 2011). A separate study using C666-1 cell line found that $CD44^+$ cells at 100,000-cell inoculation formed faster-growing tumours in NSG mice despite not having a proliferation advantage over $CD44^-$ cells *in vitro* and had similar *LMP1*, *EBNA*, *BamH1* and *VIM* gene expressions (Janisiewicz et al., 2012). Lun et al. (2012) used sphere-forming assay to enrich for self-renewing cells in C666-1 and demonstrated that CD44 was highly expressed in these spheres. $CD44^+$ cells from parental C666-1 were resistant to 5-FU treatment and had higher sphere formation efficiency than $CD44^-$ cells. Importantly, the ability of $CD44^+$ cells to form spheres was related to higher expression of *CCR7*, a gene which significantly correlated with disease recurrence or metastasis in a cohort of 39 primary NPC samples in the same study. Nonetheless, a comparison of tumour-initiating ability of C666-1 spheroids, $CD44^+$ and $CD44^-$ cells was not performed. Although CD44 is the most studied marker in NPC, little has been revealed regarding its TIC frequency and self-renewal ability *in vivo* in this malignancy. As reviewed in Chapter 2, CD44 exists in more than one (iso-) form due to alternative splicing. Functionally, CD44 variant 3 (CD44v3) showed higher resistance against cisplatin than CD44v6, CD44v10 and CD44s in a HNSCC cell line (Wang et al., 2009), CD44v8-10 was found to be responsible for tumour initiation instead of CD44s in an adenocarcinoma cell line (Lau et al., 2014), and high CD44v3 and high CD44v6 expression levels were individually related to lower survival rate and treatment resistance at primary sampling site in 42 NPC patients (Sagawa et al., 2016). Most studies, however, looked at both CD44s and CD44v isoforms and utilized panCD44 antibodies (unless specified) as described in Chapter 2.

To the best of my knowledge, there is only a previous NPC study which used EpCAM as a surface biomarker to identify NPC cells with stem-like features. Jiang et al. (2016) found that expression of β-catenin was related to a stem-like phenotype by virtue of increased EpCAM, c-myc, Nanog, Oct3/4 and Sox2 expressions in NPC cell line, CNE2. A suppressed β-catenin presence in modulated CNE-2 cell line led to a reduction of protein expression levels of EpCAM and the other CSC markers, and also smaller tumour volumes in nude mice.

Tumourigenic cells and their ability to induce growth are partly driven by the presence of stem cells which are capable of self-renewal to ensure long-term maintenance of the growth (Clarke et al., 2006; Clevers, 2011; Hanahan & Weinberg, 2011). Currently available "CSC" reports on defining the presence of such cells in NPC (aforementioned studies) do not thoroughly assess this ability with *in vivo* serial transplantation experiment, a gold standard to functionally define normal stem cells and CSCs (Clarke et al., 2006; Tang, 2012; Visvader & Lindeman, 2008). As such, these studies could only be deemed to be looking at CSC-like or stem-like cells, instead of *bona fide* CSCs.

4.3 Materials

4.3.1 Cell lines

HK1 is a NPC cell line established from an EBV negative recurrent NPC sample. The patient was diagnosed with well-differentiated squamous carcinoma with prior radiation treatment (Huang et al., 1980). HK1 was obtained as a gift from Professor George SW Tsao (University of Hong Kong).

C666-1 is a clone of parental NPC cell line, C666 which was derived from xeno-666, a NPC PDX. Xeno-666 was established from an EBV positive primary NPC sample.

The patient was diagnosed with undifferentiated NPC (Cheung et al., 1999). C666-1 was obtained as a gift from Professor KW Lo (Chinese University of Hong Kong).

4.3.2 NPC patient-derived xenografts (PDXs)

Xeno-284 is a NPC PDX established from an EBV negative recurrent NPC sample. The patient was diagnosed with poorly differentiating squamous cell carcinoma with prior concurrent chemoradiotherapy.

Xeno-B110 is a NPC PDX established from an EBV positive primary NPC sample. The patient was diagnosed with non-keratinizing differentiated carcinoma. Xeno-B110-gfp-luc2 was established from xeno-B110 transduced with lentivirus carrying GFP-Luc2 DNA transfer plasmid.

4.3.3 Cell lines and culture reagents

4.3.3.1 Complete medium for HK1 cells

Total volume of 100 mL

RPMI-1640 medium (with phenol red)	89.5 mL
Fetal bovine serum (FBS)	10 mL
10,000 U/mL Penicillin-Streptomycin	500 μL

All items were obtained from Thermo Fisher Scientific Inc., MA, USA.

4.3.3.2 Complete medium for C666-1 cells

Total volume of 100 mL

RPMI-1640 medium (with phenol red)	88.5 mL
FBS	10 mL

100X GlutaMAX™	1 mL
10,000 U/mL Penicillin-Streptomycin	500 µL

All items were obtained from Thermo Fisher Scientific Inc., MA, USA.

4.3.3.3 Freezing medium for HK1 and C666-1 cells

Total volume of 10 mL

Complete medium	5 mL
Dimethyl sulfoxide Hybri-Max (DMSO)	1 mL
FBS	4 mL

DMSO was purchased from Sigma-Aldrich, MO, USA.

4.3.3.4 1X phosphate-buffered saline (PBS)

Total volume of 100 mL

Deionized water (dH_2O)	100 mL
Phosphate-buffered salt	1 tablet

Phosphate-buffered salt was obtained from Takara Biotechnology, Shiga, Japan.

4.3.3.5 1X PBS supplemented with 1X antibiotic/antimycotic

Total volume of 100 mL

1X PBS	99 mL
100X antibiotic-antimycotic	1 mL

Antibiotic-antimycotic was purchased from Thermo Fisher Scientific Inc., MA, USA.

4.3.4 HBSS+ buffer

Total volume of 100 mL

Hanks' balanced salt solution (HBSS)	98 mL
FBS	2 mL
HEPES	0.2383 g

HBSS and HEPES were obtained from Thermo Fisher Scientific Inc., MA, USA and Sigma-Aldrich, MO, USA, respectively.

4.3.5 Reagents for cell cycle analysis

4.3.5.1 1X PBS supplemented with 50% FBS

Total volume of 10 mL

1X PBS	5 mL
FBS	5 mL

4.3.5.2 70% ethanol

Total volume of 10 mL

Absolute ethanol	7 mL
dH_2O	3 mL

Absolute ethanol (analytical grade) was purchased from Merck Millipore, MA, USA.

4.3.6 Reagents for RT-qPCR

4.3.6.1 70% ethanol

Total volume of 10 mL

Absolute ethanol	7 mL
dH_2O	3 mL

Absolute ethanol (molecular biology grade) was purchased from Merck Millipore, MA, USA.

4.3.6.2 80% ethanol

Total volume of 10 mL

Absolute ethanol	8 mL
dH_2O	2 mL

Absolute ethanol (molecular biology grade) was purchased from Merck Millipore, MA, USA.

4.3.6.3 Reverse transcription (RT) mix

Total volume of 5 μL (equivalent to 1 reaction of RT)

Reverse Transcription Master Mix	1 μL
Nuclease-free water	top up to 5 μL (after the addition of RNA)

Both items were components from Fluidigm® Reverse Transcription Master Mix kit (Fluidigm, CA, USA).

4.3.6.4 Preamplification pre-mix

Total volume of 3.75 μL (equivalent of 1 reaction of preamplification)

PreAmp Master Mix	1 μL
0.2X pooled TaqMan® assay mix	1.25 μL
Nuclease-free water	1.50 uL

PreAmp Master Mix and nuclease-free water were components from Fluidigm® PreAmp Master Mix kit (Fluidigm, CA, USA). Individual 20X TaqMan® gene expression assays (primers) were purchased from Thermo Fisher Scientific Inc., MA, USA.

4.3.6.5 quantitative PCR (qPCR)

Total volume of 5 μL of 10X assay

20X TaqMan® assay	2.5 μL
2X assay loading reagent	2.5 μL

Individual TaqMan® gene expression assays were purchased from Thermo Fisher Scientific Inc., MA, USA. Assay loading reagent was obtained from Fluidigm, CA, USA.

Total volume of 5 μL of sample

2X TaqMan® Universal PCR Master Mix	2.5 μL
20X GE sample loading reagent	0.25 μL

preamplified cDNA	2.25 μL

TaqMan® Universal PCR Master Mix was obtained from Thermo Fisher Scientific Inc., MA, USA. GE sample loading reagent was purchased from Fluidigm, CA, USA.

4.4 Methods

This study was performed according to the protocols approved by Research Management Committee of Institute for Medical Research (project code: JPP-IMR 11-026) and Animal Care and Use Committee of the Ministry of Health, Malaysia (approval code: ACUC/KKM/02(3/2013) and ACUC/KKM/02(05/2016)).

4.4.1 Cell culture

Both HK1 and C666-1 were adherent cells cultured in respective complete medium. The cells were maintained in a 5% CO_2-humidified incubator at 37 °C. Cell culture experiments were carried out in a biosafety cabinet class II and culture media and reagents were warmed to 37 °C in a water bath prior to use.

4.4.1.1 Thawing of frozen culture and culturing of revived cells

A cryovial containing cryopreserved cells was removed from the liquid nitrogen tank and immediately thawed in a water bath at 37 °C with gentle swirling. Thawed content was pipetted dropwise into a 15-mL centrifuge tube containing 5 mL of cold complete medium. The cell suspension was centrifuged at 1,000 rpm for 5 min and the supernatant was subsequently discarded. The cell pellet was resuspended with 5 mL of complete medium and transferred to a 6-cm culture dish. The cells were cultured in the incubator. Medium was changed every two to three days.

4.4.1.2 Passaging of cells

HK1 or C666-1 cells were passaged when they were 80 - 90% confluent (i.e. logarithmic growth phase). The spent medium was discarded and the cells were gently rinsed with 3 mL of 1X PBS. The cells were trypsinized with 2 mL of TrypLE™ Express Enzyme (1X) (HK1 cells) or trypsin-EDTA (0.05%) (C666-1 cells) (Thermo Fisher Scientific Inc., MA, USA) at 37 °C for 5 - 8 min. Cells were viewed under a phase contrast inverted microscope for detachment. Trypsinization process was stopped with the addition of 2 - 3 mL of complete medium. The cell suspension was transferred to a 15-mL centrifuge tube and centrifuged at 1,000 rpm for 5 min. The supernatant was discarded and the cell pellet was gently resuspended with 2 mL of complete medium prior to transfer to the required numbers of new 10-cm culture dishes (BD Falcon, NJ, USA brand for C666-1 cells). Six mL of complete medium were then added with gentle swirling to ensure an even distribution of cells in each dish. Cell detachment for immunophenotyping experiments was performed using Accutase (Millipore, MA, USA) instead of TrypLE™ Express Enzyme (1X) or trypsin-EDTA as described above.

4.4.1.3 Cryopreservation

Freezing medium was freshly prepared and chilled for each cryopreservation process. After the cells had been detached, centrifuged and the supernatant was discarded, the cell pellet was gently resuspended with chilled freezing medium. Cryovials were prelabelled and chilled prior to filling with 1 - 2 X 10^6 cells per vial. The cryovials were then stored at -20 °C for 2 - 3 h, followed by -80 °C storage for overnight prior to long term storage in liquid nitrogen.

4.4.1.4 Cell counting by automated cell counter

Following cell detachment and resuspension in complete medium, the cells were counted prior to use. Ten μL of cell suspension were combined with 10 μL of trypan blue solution for cell viability and counting checks. Ten μL of the mixture were then loaded into a Countess® cell counting chamber slide. The loaded slide was then inserted into the Countess® automated cell counter (Thermo Fisher Scientific Inc., MA, USA). Image quality was first adjusted to ensure a good contrast between live and dead cells before cells were counted automatically.

4.4.1.5 Mycoplasma detection

Mycoplasma contamination in cultured cells was periodically checked with Venor GeM Mycoplasma Detection Kit for Conventional PCR (Minerva Biolabs, Berlin, Germany) following the manufacturer's protocols.

4.4.2 NPC patient-derived xenografts (PDXs)

Xeno-284 and xeno-B110 xenografts were maintained and passaged *in vivo* in NOD-scid gamma (NSG) mice (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/ SzJ; The Jackson Laboratory, ME, USA) as subcutaneous growth in the SPF Animal Facility (Institute for Medical Research, Malaysia). Only PDXs of passage 5 - 9 were used in this study.

4.4.2.1 Harvesting of xenograft

Mouse bearing a tumour was checked for its correct identity via its ear tag. The mouse was humanely euthanized using CO_2 and/or cervical dislocation method. The exterior of the subcutaneous xenograft and its surrounding area was first disinfected with 10% w/v Povidone iodine solution (Polylab Biotech Sdn. Bhd., Malaysia). The xenograft was gently excised out and transferred to a sterile glass petri dish. Visible blood clots, blood capillaries and/or fat were removed. It was then cut length-wise to

inspect for the presence of necrotic tissue which would be removed. A cross section of the xenograft tissue was preserved in 10% neutral buffered formalin (Leica Biosystems, IL, USA).

4.4.2.2 Digestion of xenograft

Freshly harvested xenograft was rinsed several times with sterile cold 1X PBS buffer supplemented with 1X antibiotic/antimycotic (Thermo Fisher Scientific Inc., MA, USA). It was minced into fine pieces in the presence of RPMI-1640 supplemented with 10% fetal calf serum and 1X penicillin/streptomycin (all from Thermo Fisher Scientific Inc., MA, USA) ("RPMI-10 medium") in a sterile 6-cm glass petri dish. The mixture was transferred to a 50-mL Falcon tube (BD Biosciences, MA, USA) and allowed to stand for 8 min at room temperature. Dead and dying cells present in the supernatant were removed and the remaining clumps of tissue pieces were transferred back to the glass petri dish. The tissue was digested in the presence of 1:1 ratio of RPMI-10 and collagenase/dispase (Roche Life Science, IN, USA) for xeno-284 and collagenase type II (Sigma-Aldrich, MO, USA) for xeno-B110 in the presence of 2 U per mL of DNase I (Thermo Fisher Scientific Inc., MA, USA) to prevent clumping of undigested tissue pieces during digestion. Digestion was performed at 37 °C in the 5% CO_2-humidified incubator for 1 h at medium speed on a rotating belly dancer. The slurry of cell suspension mixed with undigested pieces was filtered with a 40-µm cell strainer. The cell suspension was centrifuged at 900 rpm for 6 min at room temperature and the cell pellet was washed once in sterile 1X PBS. The clean cell pellet was resuspended in 1 mL of RBC lysis buffer (Qiagen, Hilden, Germany) and incubated for 3 - 5 min at room temperature to lyse red blood cells. After centrifugation at 900 rpm for 6 min, the cell pellet was washed once in sterile 1X PBS and spun again. The final cell pellet was resuspended in HBSS+ buffer for cell and viability counts with trypan blue exclusion method.

4.4.3 Staining for immunophenotyping analysis of CSC markers or cell sorting

C666-1 or HK1 cells were detached using Accutase (in the presence of DNase I for C666-1 cells to minimise cell clumping) and resuspended in HBSS+ buffer. Cell suspension was stained with CD24 antibody conjugated to phycoerythrin (PE) (clone ML5; BD Biosciences, CA, USA), CD44 antibody conjugated to PE (clone G44-26; BD Biosciences, CA, USA), or EpCAM antibody conjugated to allophycocyanin (APC) (clone HEA-125; Miltenyi Biotec, Bergisch Gladbach, Germany) each at 10 μL per 10^6 cells, for 10 min at room temperature in the dark.

Prior to antibody-staining, xeno-284 and xeno-B110 cell suspension in HBSS+ buffer was blocked with Fc receptor blocking reagents for mouse and human antibodies at a 1:10 ratio (Miltenyi Biotec, Bergisch Gladbach, Germany) for 10 min at room temperature. Fc receptor-blocked xenograft cell suspension was then stained with mouse H2Kd antibody conjugated to fluorescein isothiocyanate (FITC) (clone SF1-1.1; BD Biosciences, CA, USA) and, CD24 antibody conjugated to PE, CD44 antibody conjugated to PE, or EpCAM antibody conjugated to APC, each at 10 μL per 10^6 cells, for another 10 min at room temperature in the dark.

Stained cells were washed once with sterile 1X PBS and centrifuged at 900 rpm for 6 min at room temperature before being resuspended in HBSS+ buffer at a concentration of 1 X 10^6 cells per 500 μL for flow analysis or 5 - 8 X 10^6 cells per mL for cell sorting. The cells were then filtered with a 40-μm cell strainer before the addition of 0.2 μg/mL of DAPI as a viability dye as well as 10 U per mL of DNase I to prevent cell clumping into the stained cell suspension. Unstained cells and cells stained with respective isotype control antibody (mouse IgG2a,κ-FITC/PE clone G155-178, mouse IgG2b,κ-PE clone 27-35, mouse IgG1-PE, mouse IgG2b-FITC clone GC198 or mouse IgG1-APC) were used as negative and gating controls, respectively.

4.4.4 Flow cytometry

4.4.4.1 Data acquisition and cell sorting

Data acquisition and cell sorting were performed on FACSAria SORP sorter equipped with 488-, 561-, 640- and 355-nm lasers (BD Biosciences, CA, USA). Fluorescence emission was measured with 530/30 (FITC), 582/15 (PE), 670/30 (APC) and 450/50 (DAPI) optical filters. Compensation was set up with CompBead Plus particles (BD Biosciences, CA, USA) singly-stained with the antibody being used. A small aliquot of cells (0.5 - 1 X 10^6 cells) heated at 95 °C for 10 min were stained with DAPI and used to compensate for DAPI spillover. Cells stained with isotype-matched antibodies were used as gating controls. Data acquisition, data analysis and cell sorting were performed using BD FACSDiva software (version 6.1.3; BD Biosciences, CA, USA).

4.4.4.2 Gating strategies for immunophenotyping and cell sorting experiments

For the immunophenotyping of CD24, CD44 and EpCAM markers, a hierarchical gating strategy was employed to identify the positive and negative marker expressions. Doublet discrimination gating was applied to remove doublets, followed by viable cell gating. Single viable cells were then gated for CD24, CD44 or EpCAM expression. The xenograft cells had an additional mouse cell exclusion gate (H2Kd negative) performed prior to gating for the individual surface marker. The expression levels for C666-1 and HK1 were reported as percentage of single viable cells, whereas the expression levels for xeno-284 and xeno-B110 were given as percentage of single viable non-H2Kd cells. The experiment was repeated three times.

Hierarchical gating strategy was also employed in the cell sorting experiment to identify the bright and dim phenotypes of CD24, CD44, EpCAM, and combination of EpCAM and CD44 (“EpCAM/CD44”) cells. Similarly, doublet discrimination gating

was first applied followed by viable cell gating. Single live cells were then gated for H2Kd negative CD24, CD44, EpCAM or EpCAM/CD44 positive cells. The bright and dim phenotypes of each marker were gated from the marker positive cells and sorted or collected for downstream experiments. Only the top total 5% of brightly-stained cells or the bottom total 5% of dimly-stained cells were regarded as "markerbr" or "markerdim" cells, respectively. Collection of sorted cells was performed at a sample flow rate of 1,000 - 1,500 cells per second at 4 °C. Sorted cells were collected into 1.5-mL or 15-mL polypropylene tube containing approximately 400 μL or 3 mL of complete medium, respectively. Typical collection time ranged from 5 min to slightly more than an hour, depending on the number of sorted cells needed for a downstream experiment.

4.4.4.3 Post-sort analyses

Approximately 2,000 - 5,000 sorted cells were collected for post-sort analyses. They were re-analysed by flow cytometry and found to be typically above 90% purity. Their morphology was also assessed microscopically. The majority of sorted cells were single, round and healthy-looking.

4.4.5 *In vitro* growth of marker-selected xeno-B110 cells

Xeno-B110-gfp-luc2 was previously established by a lentiviral transduction of GFP-Luc2 plasmid into xeno-B110 cells for a separate project (patent in application). The GFP-transduced cells were successfully and could be continuously propagated *in vivo*. Xeno-B110-gfp-luc2 cells were digested and stained with antibodies for cell sorting as described above. Twenty thousand freshly-sorted cells each of GFP positive non-mouse, GFP positive EpCAMbr and EpCAMdim phenotypes were seeded in 50 μL of complete medium per well for a total of three wells per phenotype in a 96-well ViewPlate-96 Black plate (PerkinElmer Inc., MA, USA). Luminescent signal generated

with the addition of 2X D-Luciferin was regarded as a read-out of growth of viable xenograft cells. Signals were read on day 0, 2, 4, 6 and 8.

2X D-Luciferin diluted in RPMI-10 medium was added into each well of xenograft cells containing RPMI-10 in a 1:1 ratio. The plate was gently agitated in the dark for 3 min and luminescent signal was read by EnVision multi-label plate reader (PerkinElmer Inc., MA, USA). Cell medium containing D-Luciferin was gently removed and the cells were gently washed twice with 200 μL of RPMI-10. The cells were replenished with 100 μL of fresh complete medium and returned to the incubator until the next reading.

4.4.6 *In vivo* tumourigenicity

Data for tumour incidence and latency, growth curve and mitotic figures were obtained from five mice replicates per markerbr or markerdim group, whereas serial transplantation experiment was performed on five to six mice replicates per cell inoculation in the second generation (2,500 and 5,000 cells), five to six mice replicates per cell inoculation in the third generation (500 and 1,000 cells) and three to six mice replicates per cell inoculation in the fourth generation (10, 100 and 500 cells). The exact number of mice used in each experiment is as indicated in the figures/tables.

4.4.6.1 Animal strain, husbandry and endpoint

Four to 6 weeks old female NOD-scid gamma (NSG) mice (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/ SzJ; The Jackson Laboratory, ME, USA) were used and housed in individually vented cages GM500 DGM (Tecniplast, PA, USA) with Pure-o'Cel bedding and Enrich-n'Nest paper chips (The Andersons Inc., OH, USA) in the SPF Animal Facility with standard environmental conditions of temperature at 20 - 24 °C, a relative humidity of 45 - 65% and a 12-h dark-light cycle. They had *ad libitum* access

to sterile pelleted mouse feed and sterile acidic water (pH2.8 - 3.1) supplemented with co-trimoxazole (1:20 dilution).

A mouse would be sacrificed immediately if it showed signs of being unfit or moribund (even before the experiment endpoint was reached) such as big visible tumour (1,000 - 1,500 mm^3), significant weight loss (more than 15% body weight), any impairment of basic body functions (feeding, walking, etc.) and signs of continuous distress which could not be alleviated by other means.

4.4.6.2 Preparation of sorted cells for subcutaneous inoculation

Sorted cells in 1.5-mL or 15-mL polypropylene tube were centrifuged at 1,200 rpm for 6 min at room temperature. Supernatant was carefully aspirated out and the cell pellet was homogeneously mixed with 100 μL cold RPMI-1640 medium (Thermo Fisher Scientific Inc., MA, USA) and 100 μL cold BD Matrigel™ basement membrane matrix (BD Biosciences, CA, USA). Cold RPMI-1640 medium was then aspirated into a pre-chilled 1-mL syringe until the "0" graduated line, followed by the cell-Matrigel™ mixture. Presence of air bubble was removed by gently tapping the syringe to get the air bubble to the top of the syringe prior to pushing the bubble out with the plunger. The syringe containing cell suspension-Matrigel™ mixture was then placed in a near upright position (needle facing down) in an ice bucket and immediately transported to the SPF Animal Facility for mouse inoculation.

4.4.6.3 Tumour incidence, latency and growth curve

An endpoint experiment to measure tumour incidence, latency, growth curve and mitotic figures was performed with CD24, CD44, EpCAM and EpCAM/CD44 cells from C666-1 and xeno-B110. Two thousand markerbr/double markerbr, or markerdim/double markerdim cells were inoculated into five mice per group (a total of

seven groups for C666-1 and eight groups for B110). Tumour volume measurements were recorded every two to three days using the formulae of (width^2 X length)/2. All animals (with and without tumour burden) were terminated at day 52 (C666-1 animals) or day 69 (xeno-B110 animals) post-inoculation when the majority of xenografts were below 800 mm^3. The rate of xenograft volume increase (m) was calculated from the slope of the best-fit line of a plot of volume versus time generated by a simple linear regression (modified from Blankenberg et al., 1995).

Tumourigenicity was measured by tumour incidence (i.e. number of tumours/number of inoculated mice) and latency (i.e. time from inoculation to time of first tumour measurement). In the event of no measurable growth, the length of experiment was taken as the period of latency for the calculation of mean latency per phenotype for the animal, i.e. 52 days for C666-1 experiment, 69 days for xeno-B110 experiment and 150 days for serial transplantation experiment. Mean tumour volume per group was calculated as total volume of all xenografts divided by total inoculated mice and used to plot for *in vivo* growth curves. Frequencies of TICs were calculated using Extreme Limiting Dilution Analysis (ELDA) software by Hu & Smyth (2009). The basis of and logics employed in ELDA have been described in Chapter 3.

4.4.6.4 Serial transplantation

Serial transplantation was performed for CD24, CD44 and EpCAM cells from xeno-B110 to assay for their self-renewal ability. The experiment was not performed on EpCAM/CD44 cells because they did not enhance *in vivo* growth or show higher TIC frequency than single CD44br cells. Markerbr xenografts from the first generation were digested, stained and re-sorted for markerbr cells according to the above mentioned xenograft processing and flow cytometry procedures. The cells were then re-inoculated into recipient mice in a limiting dilution manner. Markerdim xenografts were also re-

sorted for markerdim cells prior to re-inoculation in a limiting dilution manner. This was performed for three successive passages of xenografts. In order to avoid underestimating the frequencies of TICs (Quintana et al., 2008), mice with no tumour burden were kept for 150 days of observation prior to humane termination.

4.4.6.5 Harvesting of xenograft

Mouse bearing a xenograft was checked for its correct identity via its ear tag. The mouse was humanely euthanized using CO_2 and/or cervical dislocation method. The exterior of the subcutaneous xenograft and its surrounding area was disinfected with 10% w/v Povidone iodine solution (Polylab, Malaysia). The xenograft was gently excised out and transferred to a sterile glass petri dish. Visible blood clots, blood capillaries and/or fat were removed. It was then cut length-wise to inspect for the presence of necrotic tissue which would be removed. A cross section of the xenograft was preserved in 10% neutral buffered formalin (Leica Biosystems, Newcastle, UK). The remaining xenograft was digested into single cell suspension for downstream experiments as previously described.

4.4.7 Histology and scoring of mitotic figures

Xenograft tissues were formalin-fixed for 24 h, processed in an automated tissue processor Leica ASP300 S (Leica Biosystems, Melbourne, Australia), embedded in paraffin and sectioned at 3-μm thickness. Haematoxylin and eosin (H&E) staining was performed using an autostainer Leica XL (Leica Biosystems, Melbourne, Australia) and evaluated by a histopathologist. Mitotic figures were counted in seven to 10 high power fields (HPFs; 40X objective) per section (i.e. per xenograft tissue) using Leica DM1000 LED light microscope (Leica Microsystems, Singapore) by the histopathologist.

There was presence of necrosis and/or stromal cells in the majority of harvested xenografts from C666-1 and xeno-B110 experiments. As such, mitotic activity index

(MAI) could not be accurately evaluated (Medri, 2003). To circumvent this problem, the presence of necrosis and/or stromal cells were first evaluated by the histopathologist. Following this, MAI was adjusted to the presence of 100% viable tumour cells in a HPF ("adjusted MAI"), with the following formulae:

In a HPF, *x*% was necrotic and/or had stromal infiltration and there was a presence of *n* mitotic figures. The presence of viable tumour cells (i.e. non-necrotic cells) *y* equals (100 – *x*)%. Therefore, adjusted MAI for a HPF = [100% / *y*%] X *n*.

Mean adjusted MAI per section was calculated as total adjusted MAI from all scored HPFs divided by total number of HPFs in the section (Figure 4.1). Mean adjusted MAI per group of xenografts was calculated as total of mean adjusted MAI per section from all sections divided by total number of sections in the group.

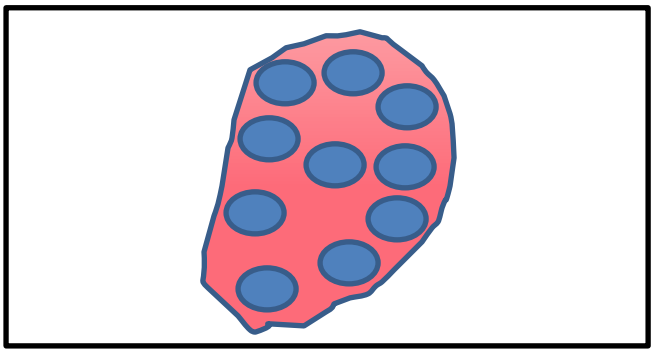


Figure 4.1: Schematic diagram of a H&E-stained xenograft tissue section used for scoring of mean adjusted MAI per section. In this illustrated xenograft tissue section, mitotic figures were scored in 10 available high power fields (HPFs, blue circles) under 40X objective. Mean adjusted MAI per section was calculated as total adjusted MAI from all 10 HPFs divided by 10 HPFs.

4.4.8 EBER-*in situ* hybridization (ISH) staining

Xenograft tissues which had been processed into formalin-fixed paraffin-embedded tissue blocks were sectioned at 3-µm thickness and prepared for EBER-ISH and cytokeratin IHC staining. Pre-treatment steps and staining were performed on the

BondMax^{TM} fully automated IHC and ISH immunostainer (Leica Biosystems, Melbourne, Australia).

Prior to hybridization with specific RNA probes (EBER, RNA positive control and RNA negative control), the tissue sections underwent pre-treatment steps which included baking, dewaxing, rehydration and proteinase K enzymatic retrieval with Bond Enzyme Pretreatment Kit (Leica Biosystems, Newcastle, UK). Subsequently, the tissue sections were hybridized with EBER probe (Bond ready-to-use ISH EBER Probe, Leica Biosystems, Newcastle, UK), RNA positive control probe (Bond ready-to-use ISH RNA Positive Control Probe, Leica Biosystems, Newcastle, UK) or RNA negative control probe (Bond ready-to-use ISH RNA Negative Control Probe, Leica Biosystems, Newcastle, UK) for 2 h at 37 °C, followed by post-hybridization washing steps at ambient temperature. The RNA positive control probe is proprietary and used to indicate correct tissue preparation with good quality RNA and a proper staining technique. The RNA negative control probe is also proprietary and used to verify the specificity of the labelling of probe to target RNAs in the tissue. After hybridization, the tissue sections underwent further staining processes which included peroxidase blocking (Bond Polymer Refine Detection, Leica Biosystems, Newcastle, UK) for 5 min at ambient temperature and incubation with anti-FITC/horseradish peroxidase (HRP) antibody (1:50 dilution, Dako, CA, USA) for 20 min at ambient temperature. Visualization of HRP activity was carried out by incubation with DAB chromogen (Bond Polymer Refine Detection, Leica Biosystems, Newcastle, UK) for 8 min at ambient temperature. Finally, they were counterstained using haematoxylin for 5 min. All staining steps were followed by washing steps using washing buffer or deionized water.

4.4.9 Cytokeratin immunohistochemistry (IHC) staining

Cytokeratin IHC staining was performed with Bond Polymer Refine Detection System (Leica Biosystems, Newcastle, United Kingdom). The sections underwent pre-treatment steps which included baking for 30 min at 60 °C, dewaxing using Bond Dewax Solution (Leica Biosystems, Newcastle, UK) and rehydration with absolute ethanol (Merck Millipore, MA, USA). They were then treated with antigen retrieval solution (Bond ER2, Leica Biosystems, Newcastle, UK) for 20 min at 98 °C. Next, the sections were incubated with peroxidase blocking solution for 8 min at ambient temperature and followed by incubation with Rodent Block M blocking solution (Biocare Medical, CA, USA) for 15 min at ambient temperature. Subsequently, the sections were incubated with mouse anti-human cytokeratin (clone AE1/AE3; 1:100 dilution, Dako, CA, USA) for 15 min and followed by a second incubation with Rodent Block M blocking solution for 15 min at ambient temperature. This followed with incubation with post primary rabbit anti-mouse IgG for 8 min and incubation with the polymer of anti-rabbit poly-HRP-IgG for 8 min at ambient temperature. Finally, the expression of cytokeratin was visualized using DAB chromogen (staining of 10 min at ambient temperature), and haematoxylin was used as a counterstain (5 min at ambient temperature). All staining steps were followed by washing steps using either washing buffer or deionized water. After the final staining step, tissue samples were dehydrated by absolute alcohol (1 min) and clearing in xylene (3 min).

4.4.10 Cell cycle analysis

Markerbr or markerdim xeno-B110 cells (CD24, CD44, EpCAM and EpCAM/CD44) were centrifuged at 1,200 rpm for 6 min at room temperature. Supernatant was discarded until approximately 100 μL remained in the tube. This volume was used to resuspend the cell pellet and the cell suspension was transferred to a 1.5-mL tube for a repeat centrifugation step. Supernatant was carefully aspirated

without disturbing the cell pellet. Three hundred microlitres of cold 1X PBS supplemented with 50% FBS was used to thoroughly resuspend the pellet. By using dropwise dispensing method, 900 μL of cold 70% ethanol was slowly added to the suspension with gentle mixing. The fixed cell suspension was incubated at -20 °C until analysis. On analysis day, the fixed cells were washed with cold 1X PBS twice to remove the ethanol and precipitated protein. Centrifugation was performed at 1,200 rpm for 6 min. The pellet was resuspended in cold 1X PBS and filtered with a 40-μm cell strainer before the addition of 1 μg per mL of DAPI. Cell cycle analysis was performed on the FACSAria SORP sorter and fluorescence emission for DAPI was measured with 450/50 optical filter with a 355-nm laser. Percentage of cells in each cell cycle phase was analysed using ModFit LT software version 6 (Verity Software House, Inc., ME, USA). The experiment was repeated at least three times.

4.4.11 RT-qPCR experiment

4.4.11.1 RNA extraction of marker-selected xeno-B110 cells

Total RNA from markerbr or markerdim xeno-B110 cells (CD24, CD44, EpCAM and EpCAM/CD44) were extracted using miRNeasy Micro kit (Qiagen, Hilden, Germany). Freshly-sorted cell pellet was lysed with 700 μL of QIAzol lysis buffer and vortexed for 1 min to ensure complete cell lysis. The homogenate was placed at room temperature for 5 min to promote dissociation of nucleoprotein complexes. For phase separation, 140 μL of chloroform was added to the homogenate and the tube was shaken vigorously for 15 s, prior to incubation at room temperature for 5 min. The chloroform-homogenate mixture was then centrifuged at 12,000 x g for 15 min at 4 °C. The clear upper aqueous phase was carefully aspirated to a new tube and 525 μL of 100% ethanol was added. The mixture was pipetted up and down several times, and transferred to an RNeasy MinElute spin column in a 2-mL collection tube. The spin column was spun at

10,000 x g for 15 s at room temperature. After discarding the flow-through and a wash with Buffer RWT, an on-column DNase digestion was performed with an addition of 80 μL of DNase I solution directly onto the spin column membrane, followed by incubation for 15 min at room temperature. Buffer RWT was added into the column, prior to centrifugation at 10,000 x g for 15 s. The flow-through was reapplied to the spin column and spun as before. The flow-through was discarded and 500 μL of Buffer RPE was added, followed by centrifugation at 10,000 x g for 15 s. Prior to elution, the total RNA was washed with 500 μL of 80% ethanol. This was followed with two centrifugation steps to ensure total removal of the ethanol. Finally, total RNA was eluted from the membrane of the column with the addition of 14 μL of nuclease-free water and a full speed centrifugation for 2 min. All RNA samples were stored at -80 °C until used.

4.4.11.2 RNA extraction of RNA positive control cell lines

Total RNA from C666-1, HK1, HONE1, B95.8, SW480, Namalwa and MDA231 were extracted using RNeasy® Plus Mini kit (Qiagen, Hilden, Germany). Buffer RLT Plus was added to the cell pellet and the lysate was homogenized with a QIAshredder spin column following the manufacturer's protocol. The homogenized lysate was then transferred to a gDNA Eliminator spin column and centrifuged at 10,000 x g for 30 s at room temperature to remove genomic DNA. To ensure appropriate binding conditions for RNA, the flow-through was mixed with 350 μL of 70% ethanol and transferred to an RNeasy spin column for centrifugation at 10,000 x g for 15 s at room temperature. Removal of contaminants from the membrane-bound RNA was performed with Buffer RW1 and Buffer RPE as per manufacturer's protocol. After the final wash with Buffer RPE, the column was centrifuged at full speed for 1 min at room temperature to ensure complete removal of Buffer RPE. Total RNA was eluted from the membrane of the

column with 40 μL of nuclease-free water. All RNA samples were stored at -80 °C until used. They were equally pooled to generate RNA positive control. Starting from 500 ng/μL of total RNA, a four-fold serial dilution of the RNA positive control sample was carried out to generate 16 standard points to be used in quality assessment of assays for gene expression study.

4.4.11.3 Quantification of total RNA

Quantitation of RNA was performed using NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific Inc., MA, USA). Nuclease-free water was used as a "blank". Samples with contamination at 230- and/or 270-nm were cleaned using RNeasy® Plus Micro kit (Qiagen, Hilden, Germany) and re-quantitated.

4.4.11.4 Reverse transcription (RT)

RT was performed using Fluidigm® Reverse Transcription Master Mix (Fluidigm, CA, USA) with 2 μL of each standard point, 4 ng of total RNA each from CD24br and CD24dim-selected cells, and 8 ng of total RNA each from CD44br, CD44dim, EpCAMbr, EpCAMdim, EpCAM/CD44dbr and EpCAM/CD44dim-selected cells in a PCR plate. Nuclease-free water was used as a RT negative control. Each 5-μL RT reaction was made up of 1 μL of RT Master Mix, the specified amount of total RNA with the remaining volume made up by nuclease-free water. The plate was then incubated at 25 °C for 5 min, 42 °C for 30 min and 85 °C for 5 min.

4.4.11.5 Preamplification

Preamplification was performed using Fluidigm® PreAmp Master Mix (Fluidigm, CA, USA). Each 5-μL preamplification reaction consisted of 1.25 μL of cDNA sample and 3.75 μL of pre-mix. Nuclease-free water was used as preamplification negative control. The preamplification conditions were 95 °C for 2 min followed by 14 cycles of

denaturation at 95 °C for 15 s and annealing/extension at 60 °C for 4 min. Preamplified cDNA products were diluted 1:5 in TE buffer prior to qPCR.

4.4.11.6 quantitative PCR (qPCR) and data analysis

qPCR was performed using Fluidigm® 48.48 Dynamic Array integrated fluidic circuits (IFC) in the Biomark™ System (Fluidigm, CA, USA, Figure 4.2). Briefly, an IFC is a chip made up of fabricated rubber as a base material and contains a series of channels transected by control lines. When pressurised, the channels deflect to form an effective seal. Tens of thousands of tiny structures can be integrated into a dense network of channels for regulating micro-, nano- or picolitre-scaled samples and reagents. The features and benefits of using an IFC include its precision and efficiency despite needing only ≤ 5 μL of cDNA per sample (48.48 Dynamic Array), versatility/scalability to high-throughput qPCR (for example, 48.48 Dynamic Array enables a maximum of qPCR of 48 samples against 48 genes/assays simultaneously), and cost-effectiveness due to less sample and reagents being used per qPCR.

RT negative control, preamplification negative control and a no-template qPCR control were included in each IFC so that non-specific amplification (if any) can be ruled out from further analysis. Data from standard points were used to construct standard curves for the evaluation of linear amplification in each assay. Three sorting replicates each for CD24, CD44 and EpCAM, and two sorting replicates for EpCAM/CD44 were analysed in this study. All qPCR reactions were performed in triplicate wells.

Prior to loading, the IFC was primed by injecting control line fluid into the accumulator on both sides of the array (Figure 4.2) and loaded into the IFC Controller MX for a Prime (113x) script run. This was followed by loading of 5 μL of each assay

and 5 μL of each diluted preamplified cDNA sample into their respective inlets on the array as shown in the figure. The filled array was loaded into the IFC Controller MX for a Load Mix (113x) script run. The primed and loaded IFC was then loaded into Biomark™ System with the following qPCR cycling conditions: 50 °C for 2 min (UNG step to protect against carryover contamination of dUTP-containing DNAs), 95 °C for 10 min (Hot Start step to activate the *Taq* DNA polymerase) followed by 40 cycles of denaturation at 95 °C for 15 s and annealing at 60 °C for 60 s.

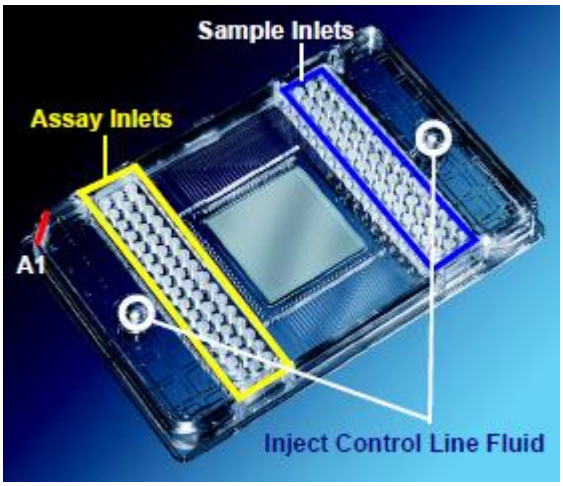


Figure 4.2: Overview of Fluidigm® 48.48 Dynamic Array IFC. (adapted from Fluidigm® 48.48 Real-Time PCR Workflow Quick Reference, BioMark Real-Time PCR Analysis Software User Guide, PN 68000088)

Fluidigm Real-Time PCR Analysis software was used to generate raw Ct values. Wells with undetected amplification (Ct = 999) or higher than 25 were excluded from calculation of average Ct values. Ct > 25 is regarded as low expression in Biomark™ (Tan & Tan, 2017). Microsoft Excel was used to calculate average Ct values from duplicate or triplicate wells. Assays were excluded from data analysis when one of the followings was identified: (a) non-linearity was observed in the standard curves of

serially-diluted pooled RNA positive controls ($R^2 \leq 0.9$), (b) PCR efficiency outside the range of 90 - 110% or (c) undetected amplification in more than 33% of all samples. Geometric mean of three reference genes (RGs) *UBC, HPRT1* and *RPL13A* was used for data normalization. List of assays analysed is in Appendix E. Gene expression was calculated with the formulae $2^{(-dCt)}$ whereby $dCt = Ct_{assay} - Ct_{geo\ mean\ RGs}$. The normalized values of markerbr were divided by the normalised markerdim to calculate fold change.

4.4.12 Statistical analysis

Data are reported as mean ± SD or SEM as indicated in each figure description. Unpaired t-test was applied for mean latency and adjusted MAI, and paired t-test for cell cycle and RT-qPCR data using GraphPad Prism 6 (GraphPad Software, Inc., CA, USA). Significance was defined at the $p < 0.05$ or $p < 0.01$ level as indicated in each figure description. TIC frequency was analysed according to Extreme Limiting Dilution Analysis (ELDA) (Hu & Smyth, 2009).

4.5 Results

4.5.1 Immunophenotyping of CD24, CD44 and EpCAM expression in NPC cell lines and NPC PDXs

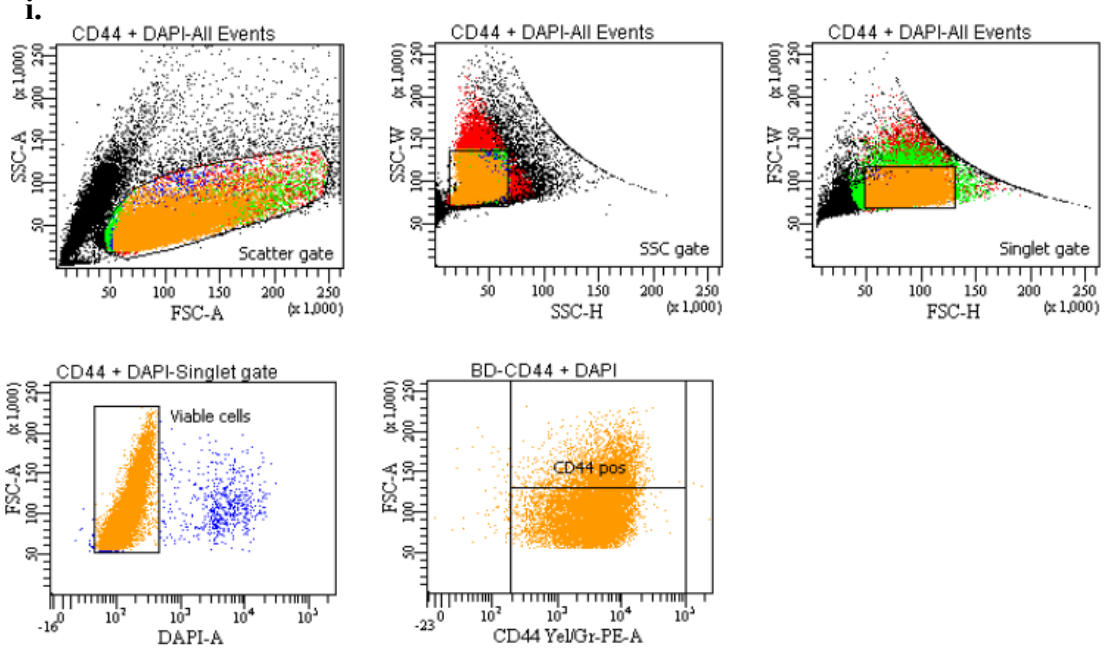
Prior to use, HK1 and C666-1 cells were authenticated by STR profiling and found to be identical and closely related, respectively, to the ones used by other NPC researchers (Chan et al., 2008) (Appendix C). Routine tests also showed that both cell lines were mycoplasma-free. Xeno-284 and xeno-B110 showed a high concordance to their respective original patient sample used for their engraftment in NSG mice (Appendix C).

Slightly different gating strategies were employed to immunophenotype the expression of CD24, CD44 and EpCAM in NPC cell lines and NPC PDXs due to the

presence of mouse cells in the latter. After doublet and viability gating on both types of samples, PDX samples were subjected to a mouse cell delineation gating prior to analyses of CSC marker expressions.

As shown in Figure 4.3, doublet gating was initially performed on C666-1 cells to remove doublets and/or cell clumps from the analysis. A viability gate was applied to single cells ("singlet gate") prior to determining the expression level of each CSC marker. Stained cells from HK1 were similarly analysed. Single cells from both cell lines were typically above 90% viability after detachment and antibody staining (representative data as shown in Figure 4.3).

i.



ii.

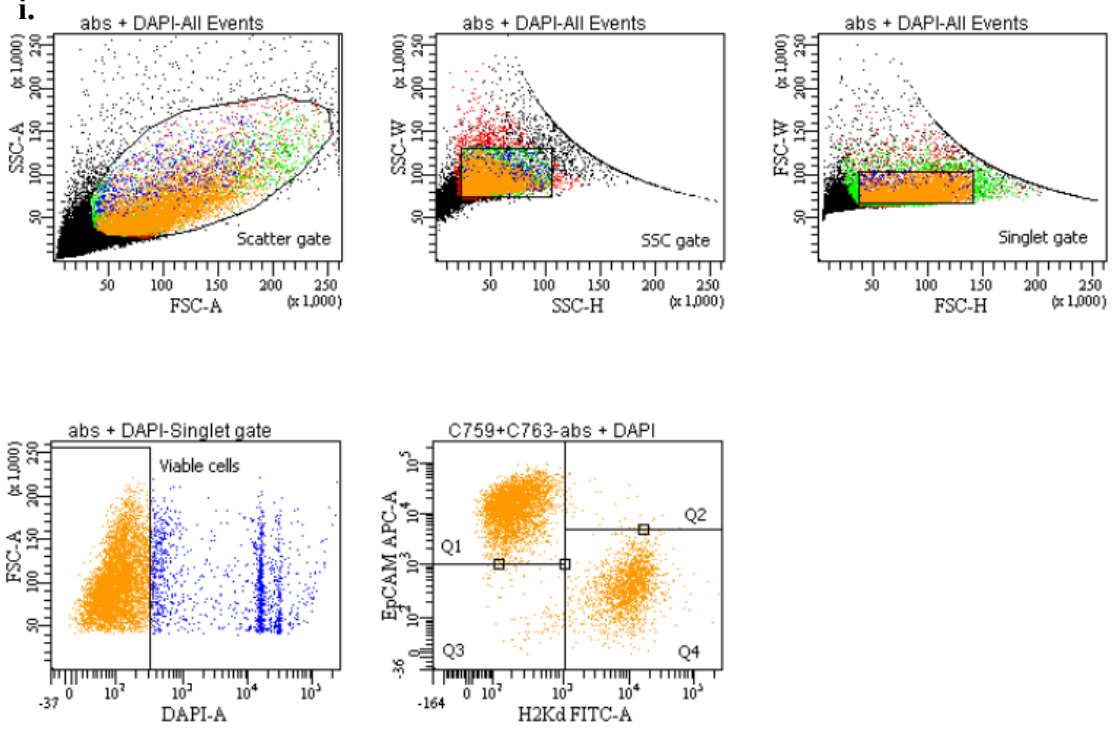
Tube: CD44 + DAPI

Population	#Events	%Parent	%Total
All Events	35,524	####	100.0
Scatter gate	19,383	54.6	54.6
SSC gate	17,520	90.4	49.3
Singlet gate	15,088	86.1	42.5
Viable cells	14,414	95.5	40.6
CD44 pos	14,338	99.5	40.4

Figure 4.3: Flow cytometry dot plots detailing the gating strategy employed for immunophenotyping C666-1 cells. **i.** C666-1 cells were stained with CD44 antibody conjugated to PE and viability dye DAPI and subsequently analysed in a hierarchical manner as shown in the statistics box **(ii)**. In the sample depicted here, 99.5% of single, viable C666-1 cells are CD44 positive.

As PDX cells contained a mixture of NPC (human cells) and host mouse cells, single, viable PDX cells (typically above 80% viability) were further stratified into mouse and non-mouse cells by H2Kd antibody (a mouse cell marker) conjugated to FITC (H2Kd-FITC) (Figure 4.4). The expression levels of CSC markers were then determined from non-mouse (H2Kd negative) cells.

i.



ii.

Tube: abs + DAPI

Population	#Events	%Parent	%Total
All Events	30,000	####	100.0
Scatter gate	11,297	37.7	37.7
SSC gate	9,919	87.8	33.1
Singlet gate	7,279	73.4	24.3
Viable cells	5,840	80.2	19.5
Q1	3,886	66.5	13.0
Q2	41	0.7	0.1
Q3	100	1.7	0.3
Q4	1,813	31.0	6.0

Figure 4.4: Flow cytometry dot plots detailing the gating strategy employed for immunophenotyping xeno-B110 cells. **i.** Xeno-B110 cells were stained with H2Kd antibody conjugated to FITC, EpCAM antibody conjugated to APC and viability dye DAPI and subsequently analysed in a hierarchical manner as shown in the statistics box **(ii)**. In the sample depicted here, 66.5% of single, viable, non-mouse xeno-B110 cells are EpCAM positive.

The percentage of CD24-expressing cells was the highest in xeno-B110 (85.37 ± 10.51% positive cells), moderate in xeno-284 and HK1 (55.33 ± 14.17% and 62.77 ± 14.63%, respectively), and nearly absent in C666-1 (0.00 ± 0.06%) (Figure 4.5). The marker was similarly barely detectable in C666-1 maintained and propagated as xenografts for three successive passages (Figure 4.6). CD44 was expressed in all NPC

samples, ranging from a moderate proportion in xeno-B110 (70.15 ± 3.23%) to the highest proportion in HK1 (99.47 ± 0.15%). There was an overall high percentage of EpCAM-expressing cells in all samples (more than 95% positivity), except for xeno-B110 (75.79 ± 12.45%). The negative subpopulations of CD24, CD44 and EpCAM were scarce in xeno-B110 (Figure 4.7); hence, bright and dim phenotypes of each marker were selected and studied for their biological properties instead of positive and negative subpopulations.

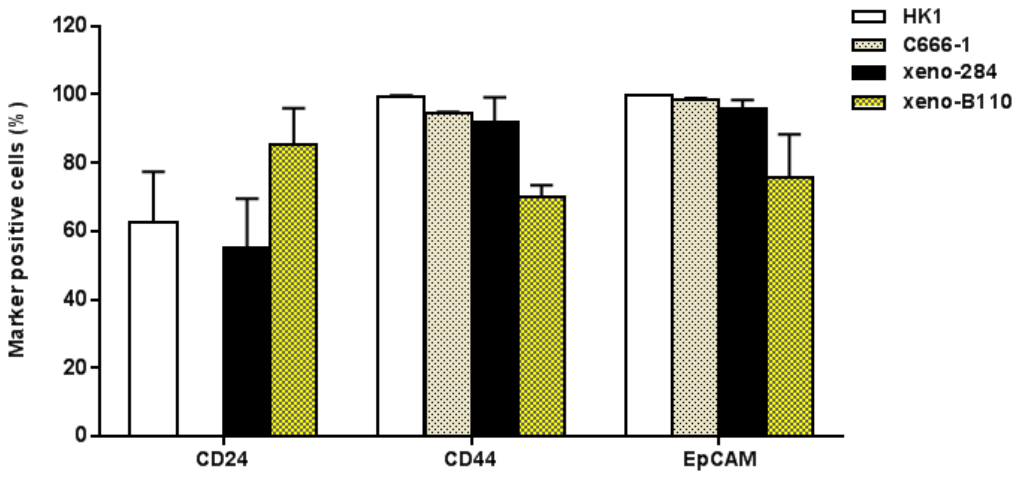


Figure 4.5: Immunophenotyping of CD24, CD44 and EpCAM in NPC cell lines and NPC PDXs. Three common CSC surface markers (CD24, CD44 and EpCAM) were differentially expressed in NPC cell lines and NPC PDXs. Percentage of marker positive cells from the cell lines was counted from the total number of single, viable cells. As for the PDXs, the denominator was total number of single, viable, non-mouse cells. Results, mean ± SD (n = 3 flow cytometry experiments).

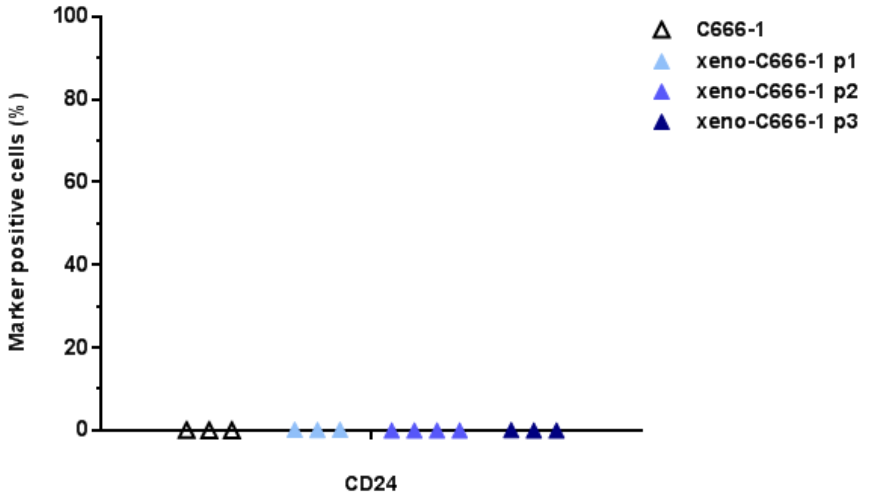


Figure 4.6: Immunophenotyping of CD24 in C666-1 maintained and propagated *in vitro* and *in vivo*. Expression of CD24 was measured by flow cytometry in *in vitro*-maintained C666-1 cells ("C666-1") as compared to *in vivo*-propagated C666-1 xenografts ("xeno-C666-1") for three successive passages (p1-p3). Percentage of CD24 positive cells was counted from total number of single, viable, non-mouse cells.

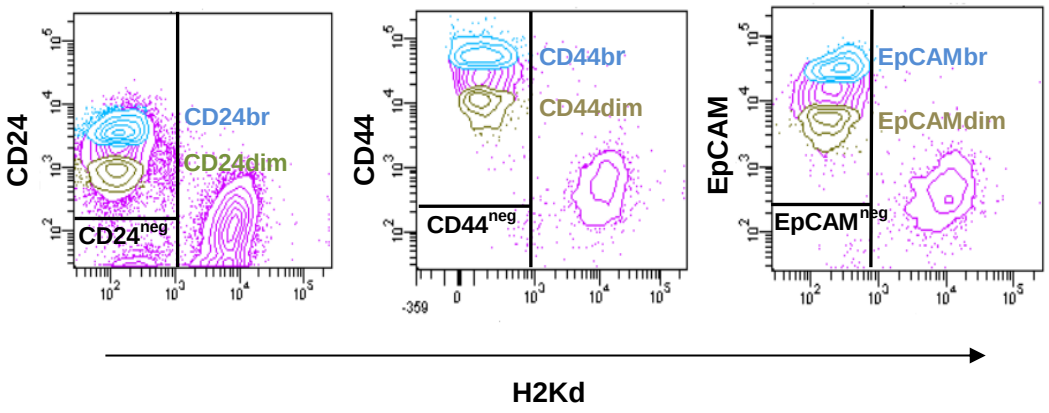


Figure 4.7: xeno-B110 contained differential expression levels of CD24, CD44 and EpCAM. Representative contour plots of immunophenotyping data for CD24, CD44 and EpCAM showed that non-mouse (H2Kd negative) cells with $CD24^{neg}$, $CD44^{neg}$ or $EpCAM^{neg}$ expression were too few to be sorted from xeno-B110.

4.5.2 Characterization of CD44, EpCAM and combination of EpCAM and CD44 marker-selected cells in C666-1

Majority of NPC are EBV-positive. Therefore, subsequent *in vitro* and/or *in vivo*-based experiments to characterize biological differences of marker-selected (marker bright or marker dim) NPC cells were performed using C666-1 as a representative NPC cell line and xeno-B110 as a representative early-passage NPC PDX (Figure 4.8) as both samples are EBV positive and derived from primary NPC specimens.

Besides CD24, CD44 and EpCAM, a combination of EpCAM and CD44 ("EpCAM/CD44") was also investigated to determine if it would further enrich for tumourigenic and/or faster-growing cells as both markers had been independently found to be associated with tumour initiation and proliferation in published reports.

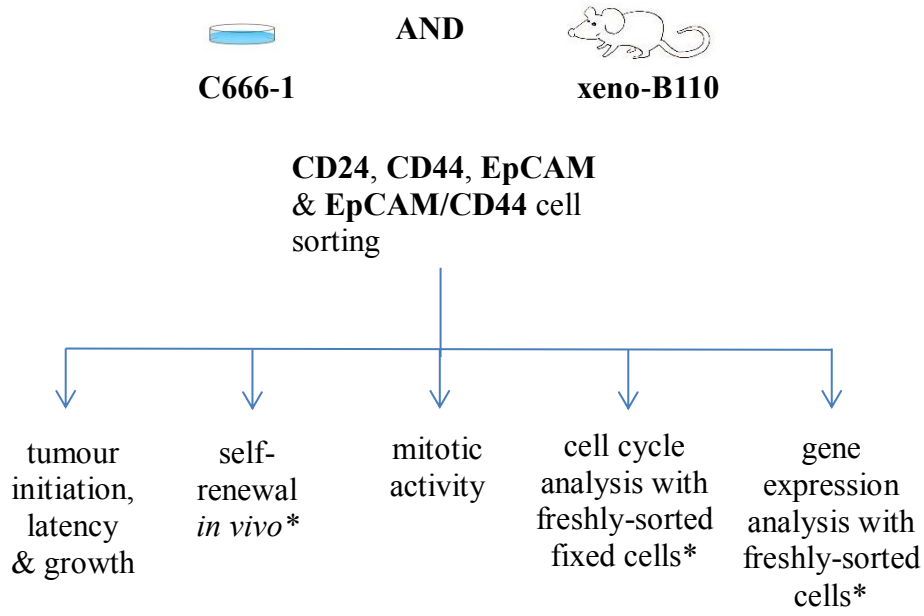


Figure 4.8: Schematic work flow of characterization studies in C666-1 and xeno-B110. Self-renewal *in vivo*, cell cycle and gene expression analyses were only performed in xeno-B110 cells (*).

Owing to an extremely low level of CD24 positive cells in C666-1 (Figure 4.5), only CD44, EpCAM and EpCAM/CD44-selected C666-1 cells were evaluated for their tumourigenicity in NSG mice. Table 4.1 showed the tumour formation ability of 2,000 marker-selected cells in a 52-day experiment. All cell phenotypes initiated 100% tumour formation (5/5 for each phenotype), except for CD44dim phenotype (80%, 4/5). Only CD44 group of xenografts showed a significant growth difference between the phenotypes: CD44br cells induced faster growth with a mean latency of 35.60 ± 1.50 days in contrast to CD44dim cells with a longer mean latency of 44.80 ± 6.85 days ($p = 0.03$).

Table 4.1: Tumour initiation and mean latency data of the xenografts arising from bright and dim phenotypes of CD44, EpCAM and EpCAM/CD44 cells from C666-1.

Sample	Phenotype	Number of tumours/Number of mice	Mean latency ± SD (days)	p value
		2,000 cells		
C666-1	viable	5/5	42.00 ± 5.55	
	CD44br	5/5	35.60 ± 1.50	0.03
	CD44dim	4/5[#]	44.80 ± 6.85	
	EpCAMbr	5/5	40.60 ± 3.07	0.14
	EpCAMdim	5/5	45.20 ± 4.62	
	EpCAM/CD44dbr	5/5	40.60 ± 5.85	0.21
	EpCAM/CD44ddim	5/5	46.00 ± 5.40	

[#] mouse with no tumour was given latency of 52 days (duration from inoculation until termination of experiment)

Growth curve of CD44br-induced xenografts was indicative of higher proliferation rate as shown by the higher rate of volume increase (steeper slope) as compared to the growth curve of CD44dim-induced xenografts (Figure 4.9i). The former also produced larger xenografts than the latter at the end of the experiment (Figure 4.9ii). Although most of the CD44br and CD44dim xenografts were less than 800 mm^3 on day 52, necrosis and/or stromal infiltration were observed in nearly all of them (Figure 4.10). Hence, the calculation of mitotic activity index (MAI) was adjusted to 100% presence of viable tumour cells in each xenograft section ("adjusted MAI"). CD44br xenografts had slightly more necrosis and/or stroma presence than CD44dim xenografts ($p > 0.05$) which may be attributed to their comparatively higher proliferation rate ($m = 26.85$) than the other xenografts. The mean adjusted MAI for CD44br xenografts was 13.00 ± 2.17 as compared to 8.50 ± 3.01 for CD44dim xenografts ($p > 0.05$) (Figure 4.9iii).

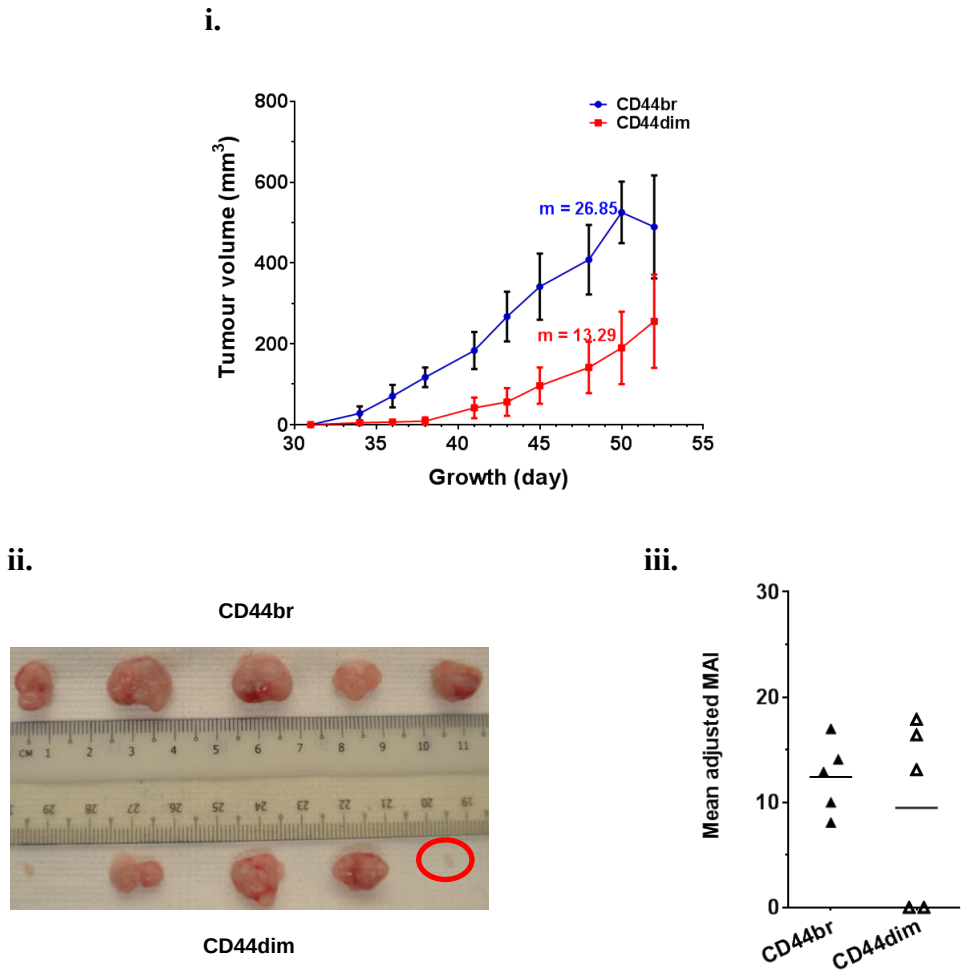


Figure 4.9: *In vivo* growth properties of CD44 cells from C666-1. Freshly-sorted CD44br and CD44dim cells were inoculated into NSG mice and monitored for 52 days. **i.** Growth curves of CD44br and CD44dim xenografts, **ii.** image of the harvested xenografts on day 52 and **iii.** mean adjusted MAI per group between CD44br and CD44dim xenografts. Results of growth curve, mean ± SEM of 4 or 5 xenograft replicates. m, rate of volume increase. Red circle in **(ii)** indicates a fat tissue and is considered as "no tumour".

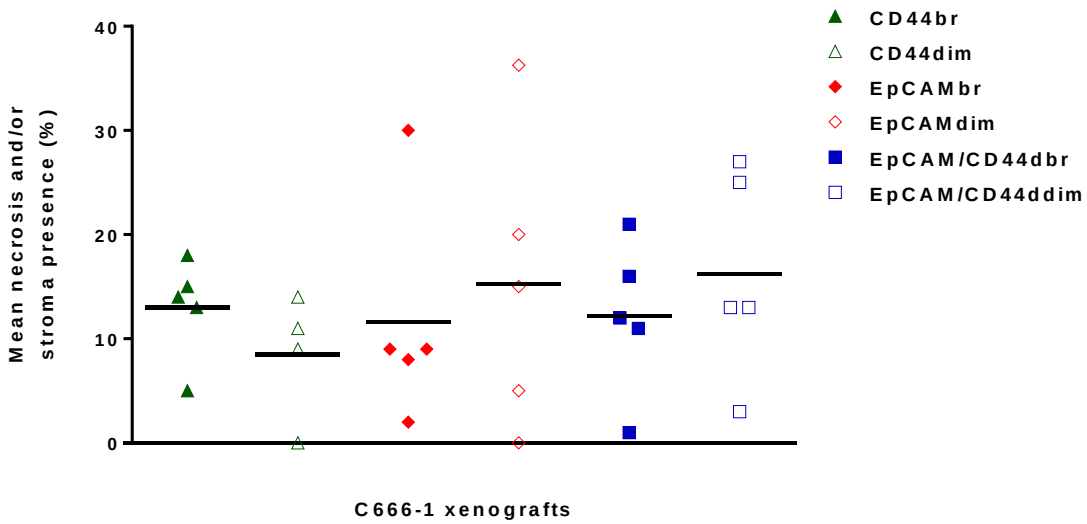


Figure 4.10: Presence of necrosis and/or stroma in analysed high power fields (HPFs) of FFPE sections of marker-selected cell-induced C666-1 xenografts. Majority of the marker-selected cell-induced C666-1 xenografts contained varying presence of necrosis and/or stroma. Results, mean (n = 4 or 5 xenografts per phenotype with 7 to 10 HPFs per xenograft).

There were less apparent differences between the growth curves and size of EpCAMbr and EpCAMdim xenografts ($p > 0.05$) (Figure 4.11i and ii) which was also reflected in the mean latency data (Table 4.1). However, the mean adjusted MAI for EpCAMbr was significantly higher than the mean adjusted MAI for EpCAMdim xenografts ($p = 0.03$) (Figure 4.11iii).

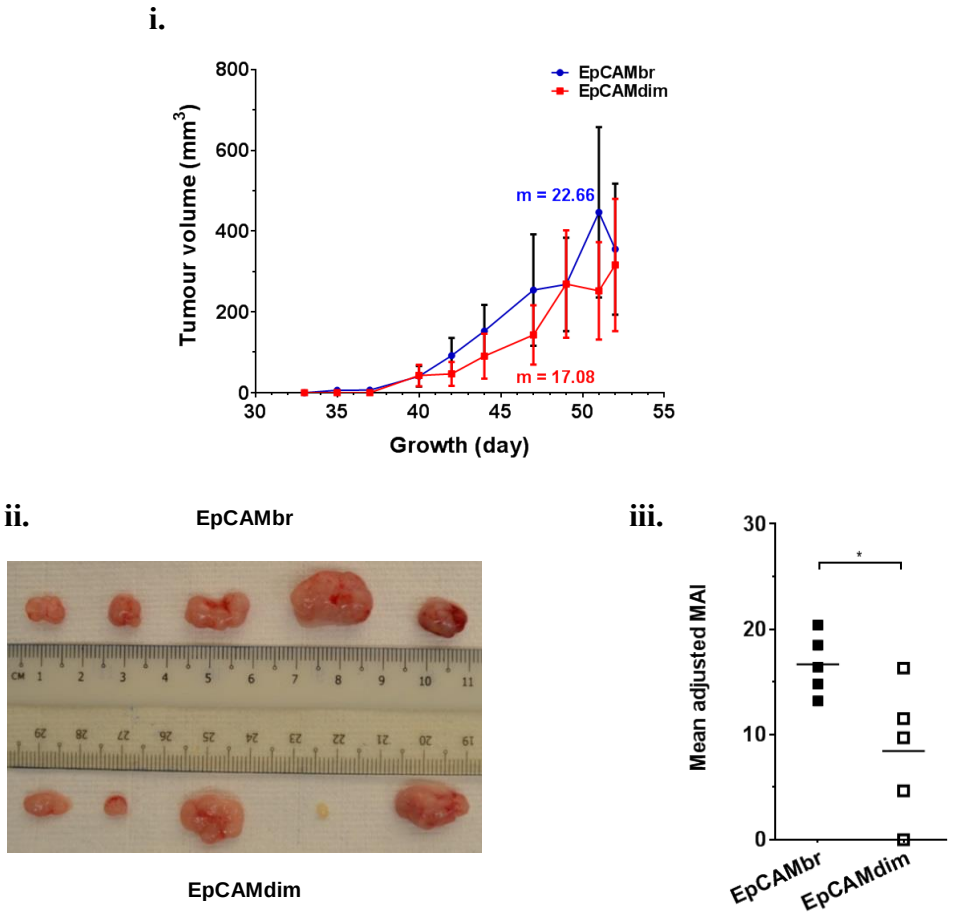


Figure 4.11: *In vivo* growth properties of EpCAM cells from C666-1. Freshly-sorted EpCAMbr and EpCAMdim cells were inoculated into NSG mice and monitored for 52 days. **i.** Growth curves of EpCAMbr and EpCAMdim xenografts, **ii.** image of the harvested xenografts on day 52 and **iii.** mean adjusted MAI per group between EpCAMbr and EpCAMdim xenografts. Results of growth curve, mean ± SEM of 5 xenograft replicates. m, rate of volume increase. * $p < 0.05$

Compared to single CD44 and EpCAM marker-induced xenografts, xenografts arising from cells co-expressing EpCAM and CD44 ("EpCAM/CD44") did not show any enrichment of growth abilities (Figure 4.12). EpCAM/CD44dbr and EpCAM/CD44ddim xenografts showed similar growth curves and mean adjusted MAI values.

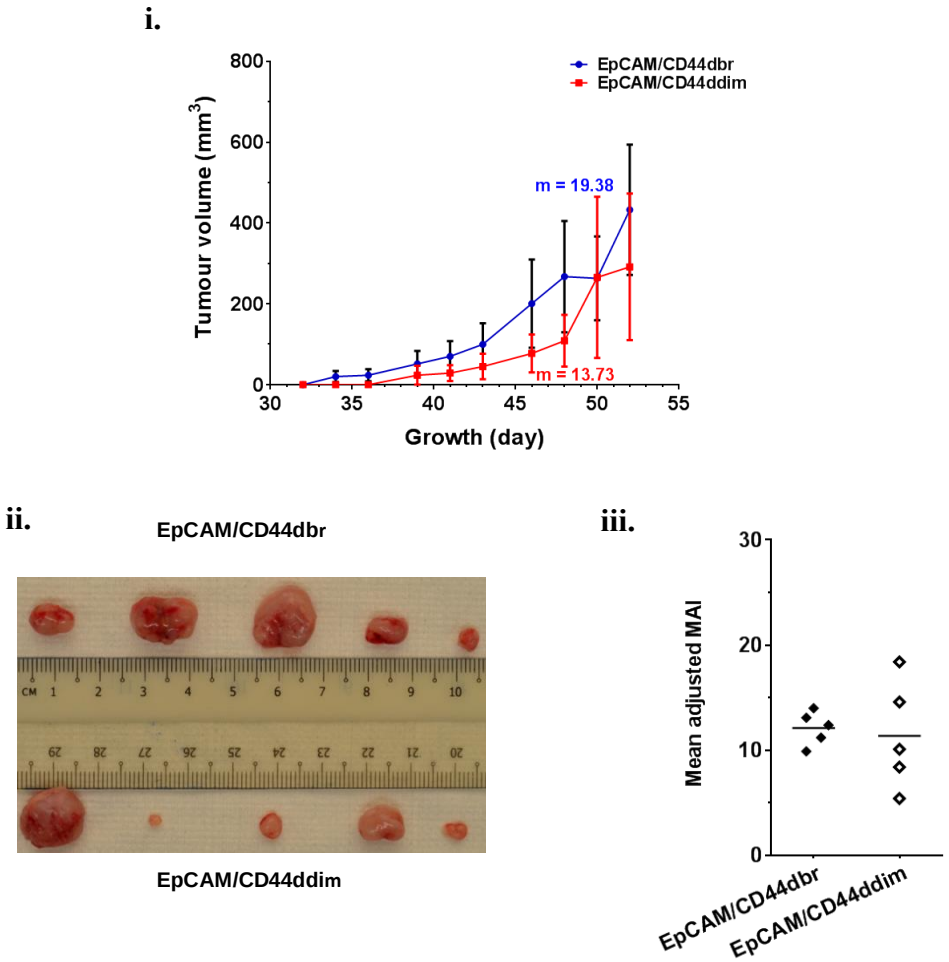


Figure 4.12: *In vivo* growth properties of EpCAM/CD44 cells from C666-1. Freshly-sorted EpCAM/CD44dbr and EpCAM/CD44ddim cells were inoculated into NSG mice and monitored for 52 days. **i.** Growth curves of EpCAM/CD44dbr and EpCAM/CD44ddim xenografts, **ii.** image of the harvested xenografts on day 52 and **iii.** mean adjusted MAI per group between EpCAM/CD44dbr and EpCAM/CD44ddim xenografts. Results of growth curve, mean ± SEM of 5 xenograft replicates. m, rate of volume increase.

4.5.3 Characterization of CD24, CD44, EpCAM and EpCAM/CD44 marker-selected cells in xeno-B110

Xeno-B110 is a NPC PDX newly established in the Institute for Medical Research, Malaysia. As part of its characterization work in a separate study, STR profiling shows that it is closely related to the original patient specimen (90.2%, Appendix C) and epithelial in nature (Figure 4.13ii). Xeno-B110 is also EBV positive with

heterogeneous staining intensities amongst its cells (Figure 4.14iii). Sample preparation and data analysis for STR profiling were performed by Norazlin Abd. Aziz and Dr. Kitson Liew (Institute for Medical Research, Malaysia). Sample preparation, staining of EBER-ISH and cytokeratin IHC, and imaging were performed by Norazlin Abd. Aziz (Institute for Medical Research, Malaysia).

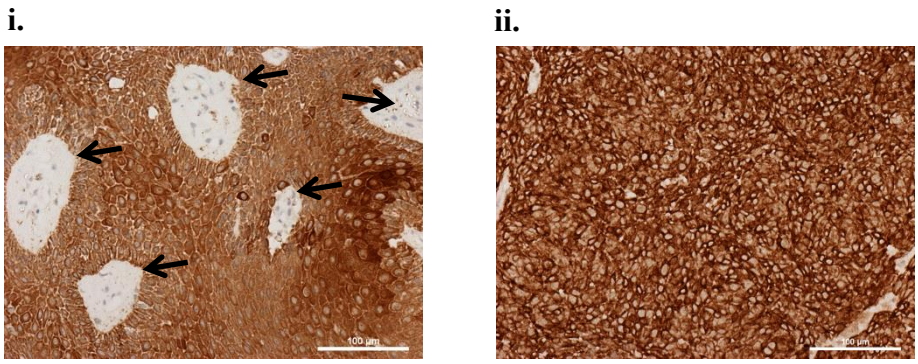


Figure 4.13: Representative cytokeratin IHC images of human skin tissue and xeno-B110. **i.** Human skin tissue section (as positive control) were stained with cytokeratin antibody and showed positive brown stain in epithelial cells and negative stain in non-epithelial cells (arrow). **ii.** Xeno-B110 tissue section was similarly stained and showed uniform positive brown stain. (20X objective; scale bar 100 μm).

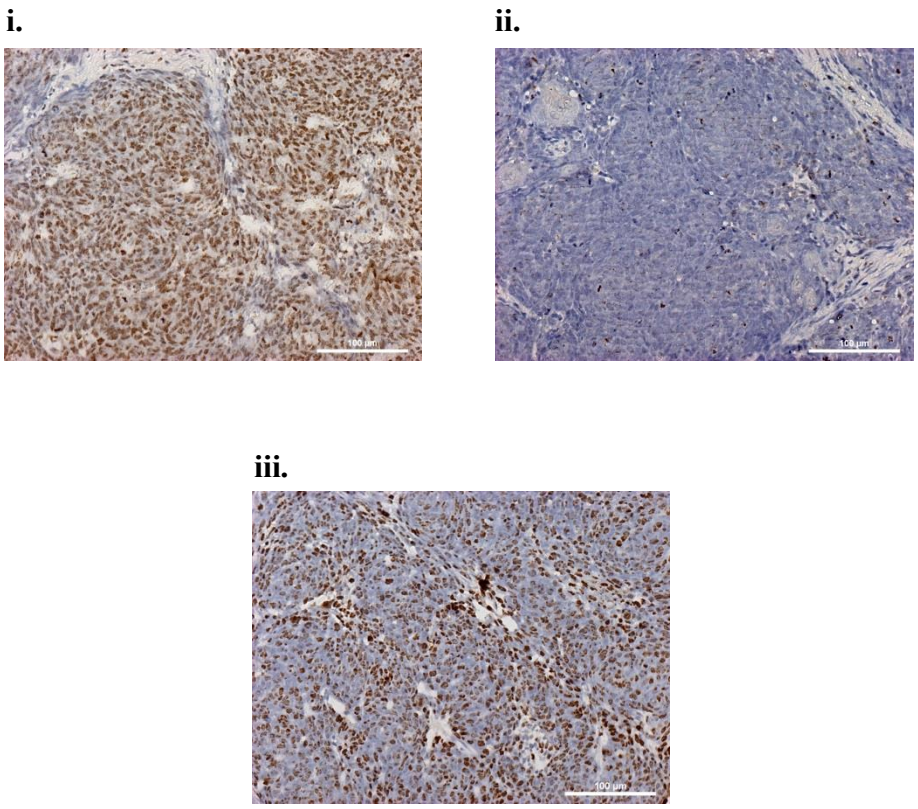


Figure 4.14: Representative ISH images of xeno-B110. Xeno-B110 tissue sections were stained with **i.** RNA positive control, **ii.** RNA negative control and **iii.** EBER probe. (20X objective; scale bar 100 μm).

A pilot experiment was first performed to determine its tumour-initiating ability with a titration of cell inoculation numbers (Table 4.2). Host mouse cells were removed by cell sorting and only viable non-mouse cells ("H2Kd negative") were inoculated into NSG mice. There was a 100% tumour formation from 100,000 to 500 cell inoculations (except for 5,000 cell inoculation at 83.33%). Tumour formation was greatly reduced at 100 cell inoculation (2/6; 33.33%) with no tumour formation with 10 cells (0/6; 0%).

Table 4.2: Tumour initiation and mean latency data of the xenografts arising from limiting dilutions of non-mouse cells (H2Kd neg) of xeno-B110.

Sample/ Phenotype	Number of cells inoculated	Number of tumours/Number of mice	Mean latency ± SD (days)
xeno-B110/ H2Kdneg	100,000	4/4	63.50 ± 1.00
	50,000	4/4	71.75 ± 13.70
	30,000	4/4	71.00 ± 6.93
	10,000	4/4	78.75 ± 11.06
	5,000	5/6*	74.83 ± 37.74
	500	3/3	70.33 ± 14.57
	100	2/6*	126.00 ± 37.71
	10	0/6*	150.00 ± 0.00

* mouse with no tumour was given latency of 150 days (duration from inoculation until termination of experiment)

A subsequent tumourigenicity experiment using 2,000 marker-selected cells for inoculation revealed that there was a significant difference in mean latency data between bright and dim phenotypes of each marker, with almost all phenotypes initiating a 100% tumour formation (5/5 or 4/4), except for CD44dim (60%; 3/5) and EpCAMdim (80%; 4/5) (Table 4.3).

Table 4.3: Tumour initiation and mean latency data of xenografts arising from H2Kd negative bright and dim phenotypes of CD24, CD44, EpCAM and EpCAM/CD44 cells in xeno-B110.

Sample	Phenotype	Number of tumours/Number of mice	Mean latency ± SD (days)	p value
		2,000 cells		
xeno-B110	CD24br	5/5	56.20 ± 2.68	0.03
	CD24dim	4/4^	60.25 ± 1.50	
	CD44br	5/5	57.80 ± 2.68	0.03
	CD44dim	3/5*	64.20 ± 4.55	
	EpCAMbr	5/5	54.40 ± 3.58	0.02
	EpCAMdim	4/5*	63.00 ± 5.61	
	EpCAM/CD44dbr	5/5	54.80 ± 2.68	<0.01
	EpCAM/CD44ddim	5/5	63.60 ± 4.93	

^ 1 mouse died after inoculation
* mouse with no tumour was given latency of 69 days (duration from inoculation until termination of experiment)

The percentage of S-phase cells was similar in freshly-sorted fixed CD24br (8.94 ± 1.18%) and CD24dim (7.93 ± 2.59%) subpopulations ($p > 0.05$) (Figure 4.15i). Despite this, there was a significant 4-day difference in mean latency data for CD24br and CD24dim xenografts (Table 4.3). Moreover, CD24br xenografts were more proliferative than CD24dim xenografts as indicated by more than 3-fold difference in rate of volume increase (Figure 4.15ii) which corresponded with a visible difference in harvested xenograft sizes (Figure 4.15iii) and mean adjusted MAI values (Figure 4.15iv).

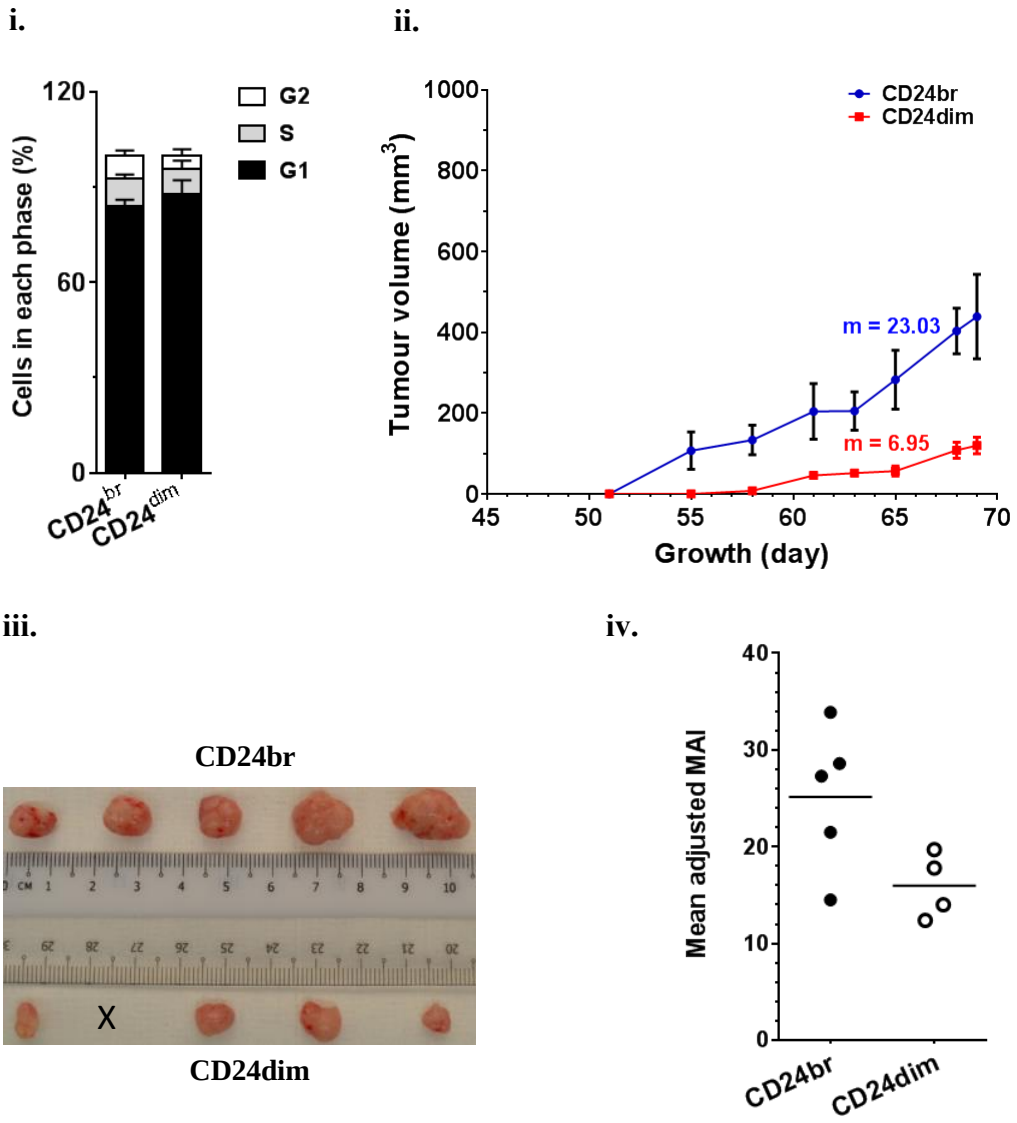


Figure 4.15: *In vivo* growth properties of CD24 cells from xeno-B110. **i.** Freshly-sorted and fixed CD24br and CD24dim cells were analysed for their cell cycle distribution. Another batch of freshly-sorted CD24br and CD24dim cells were inoculated into NSG mice and monitored for 69 days. **ii.** Growth curves of CD24br and CD24dim xenografts, **iii.** image of the harvested xenografts on day 69 and **iv.** mean adjusted MAI per group between CD24br and CD24dim xenografts. Results of cell cycle profile, mean ± SD of 3 flow cytometry experiment replicates. Results of growth curve, mean ± SEM of 4 or 5 xenograft replicates. m, rate of volume increase. X, mouse died immediately after inoculation.

On the other hand, CD44br cells contained 13.26 ± 1.56% of S-phase cells compared to 4.41 ± 0.47% in CD44dim cells ($p < 0.01$) (Figure 4.16i). As described earlier, there was a significant difference of mean latency between CD44br and CD44dim xenografts

(6.4 days; $p = 0.03$) (Table 4.3). CD44br cells produced faster-growing (more than 4-fold difference in rate of volume increase) (Figure 4.16ii) and larger xenografts (Figure 4.16iii) than CD44dim cells in NSG. Similar to C666-1 xenografts, a higher presence of necrosis and/or stroma was observed in CD44br xenografts as compared to CD44dim xenografts ($p < 0.0001$) (Figure 4.16) which may be caused by a relatively higher proliferation rate ($m = 38.27$) than the other xenografts. The mean adjusted MAI for CD44br xenografts was higher than the mean adjusted MAI for CD44dim xenografts ($p > 0.05$) (Figure 4.16iv).

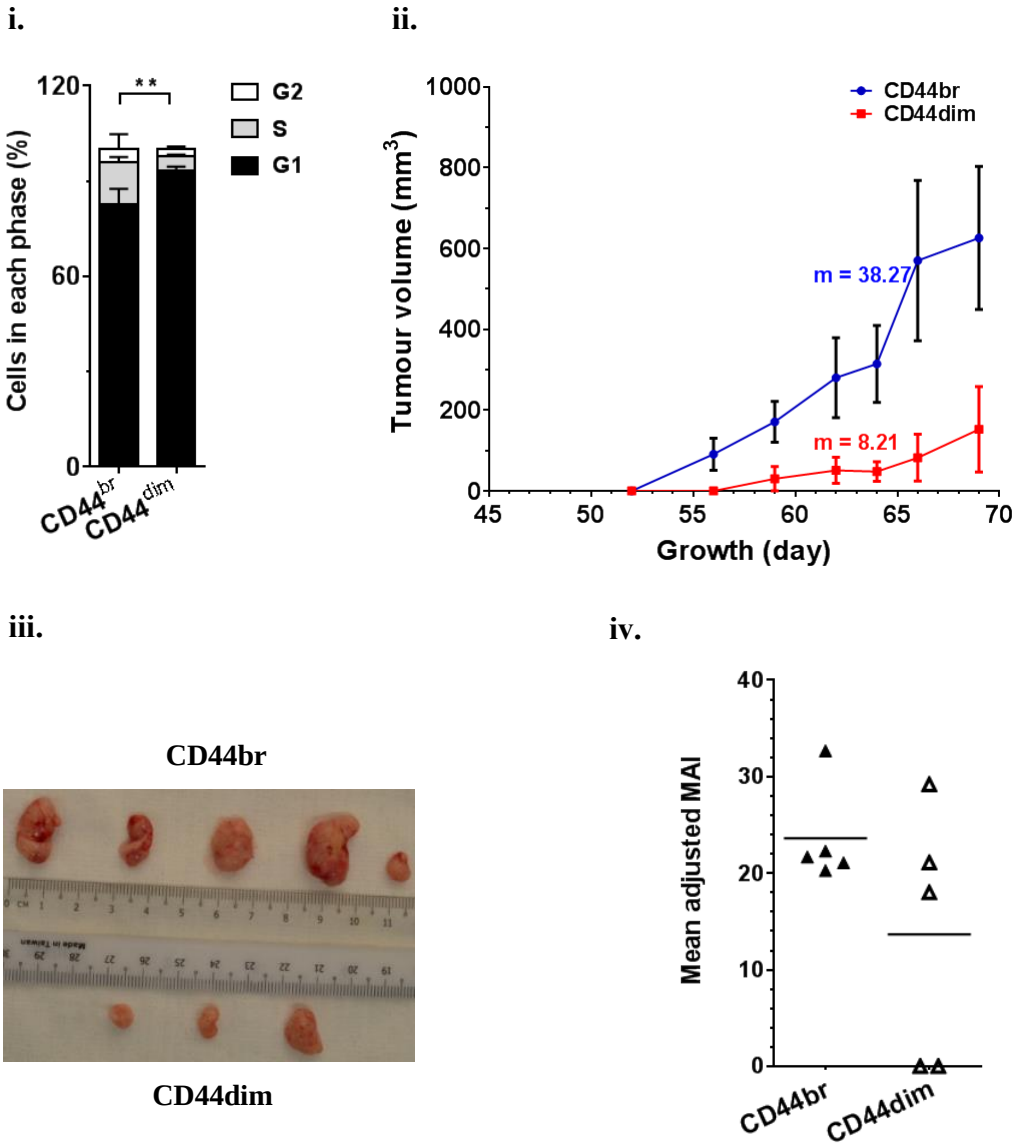


Figure 4.16: *In vivo* growth properties of CD44 cells from xeno-B110. **i.** Freshly-sorted and fixed CD44br and CD44dim cells were analysed for their cell cycle distribution. Another batch of freshly-sorted CD44br and CD44dim cells were inoculated into NSG mice and monitored for 69 days. **ii.** Growth curves of CD44br and CD44dim xenografts, **iii.** image of the harvested xenografts on day 69 and **iv.** mean adjusted MAI per group between CD44br and CD44dim xenografts. Results of cell cycle profile, mean ± SD of 4 flow cytometry experiment replicates. Results of growth curve, mean ± SEM of 5 xenograft replicates. m, rate of volume increase. ** $p < 0.01$ (S-phase)

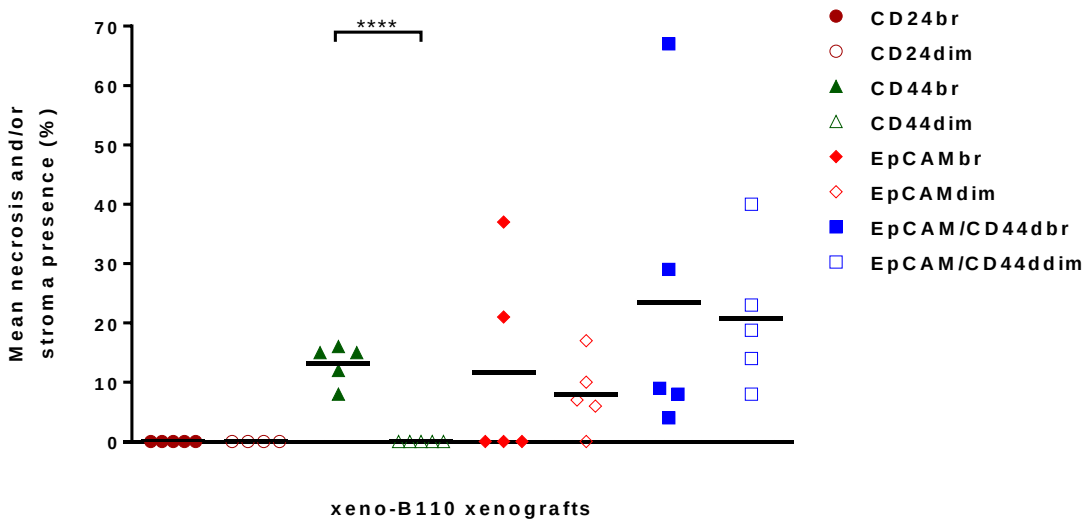


Figure 4.17: Presence of necrosis and/or stroma in analysed high power fields (HPFs) of FFPE sections of marker-selected cell-induced xeno-B110 xenografts. Many marker-selected cell-induced xeno-B110 xenografts contained varying presence of necrosis and/or stroma, except for CD24br, CD24dim and CD44dim xenografts. Results, mean (n = 4 or 5 xenografts per phenotype with 7 to 10 HPFs per xenograft). **** $p < 0.0001$

Akin to CD44br cells, EpCAMbr cells had significantly higher presence of S-phase cells (12.43 ± 2.77%) as compared to EpCAMdim cells (5.06 ± 0.33%) ($p = 0.01$) (Figure 4.18i). EpCAMbr xenografts also formed tumours much earlier with a significant 8.6-day difference (Table 4.3) and grew at a faster volume increase rate (Figure 4.18ii) which correlated with a higher mean adjusted MAI (Figure 4.18iv) than EpCAMdim xenografts.

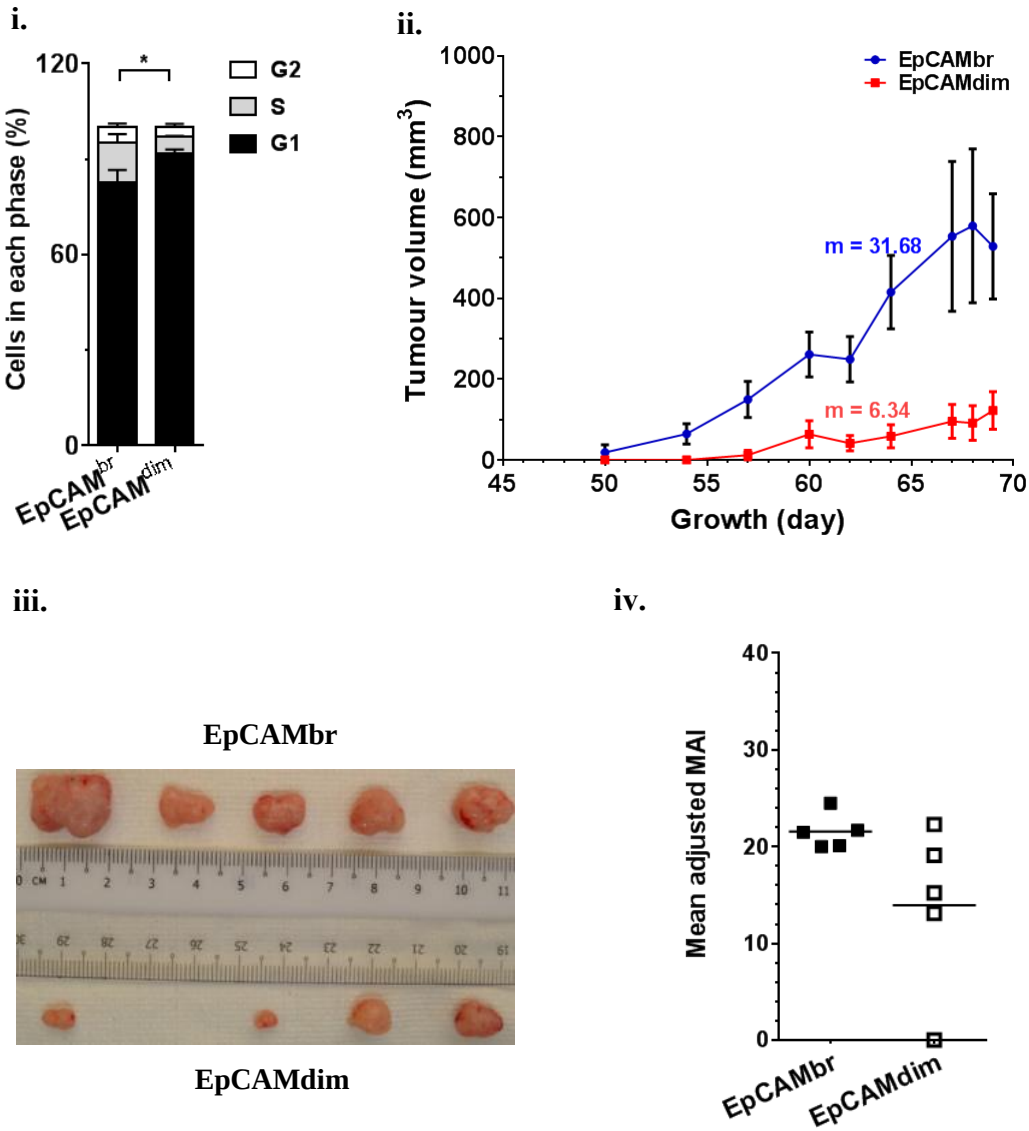


Figure 4.18: *In vivo* growth properties of EpCAM cells from xeno-B110. **i.** Freshly-sorted and fixed EpCAMbr and EpCAMdim cells were analysed for their cell cycle distribution. Another batch of freshly-sorted EpCAMbr and EpCAMdim cells were inoculated into NSG mice and monitored for 69 days. **ii.** Growth curves of EpCAMbr and EpCAMdim xenografts, **iii.** image of the harvested xenografts on day 69 and **iv.** mean adjusted MAI per group between EpCAMbr and EpCAMdim xenografts. Results of cell cycle profile, mean ± SD of 4 flow cytometry experiment replicates. Results of growth curve, mean ± SEM of 5 xenograft replicates. m, rate of volume increase. * $p < 0.05$ (S-phase)

S-phase cell distribution within the EpCAM/CD44 group was alike CD44 group and EpCAM group (Figure 4.19i). EpCAM/CD44dbr cells had 14.77 ± 4.15% of S-phase cells whereas there was 3.22 ± 0.47% of such cells in EpCAM/CD44ddim ($p = 0.04$). There was a significant 8.8-day difference in mean latency data between the double bright and double dim cells (Table 4.3). The growth curves of EpCAM/CD44dbr and EpCAM/CD44ddim xenografts also showed a very obvious difference in the rate of volume increase (Figure 4.18ii) which was seen in the sizes of harvested xenografts (Figure 4.18iii). These differences were not reflected in the mean adjusted MAI between EpCAM/CD44dbr (16.92 ± 2.42) and EpCAM/CD44ddim (16.32 ± 2.67) xenografts ($p > 0.05$) (Figure 4.19iv). Together, the cell cycle and *in vivo* growth data showed that EpCAM/CD44dbr marker did not identify for a substantial increase of tumourigenic cells than CD44br or EpCAMbr marker alone.

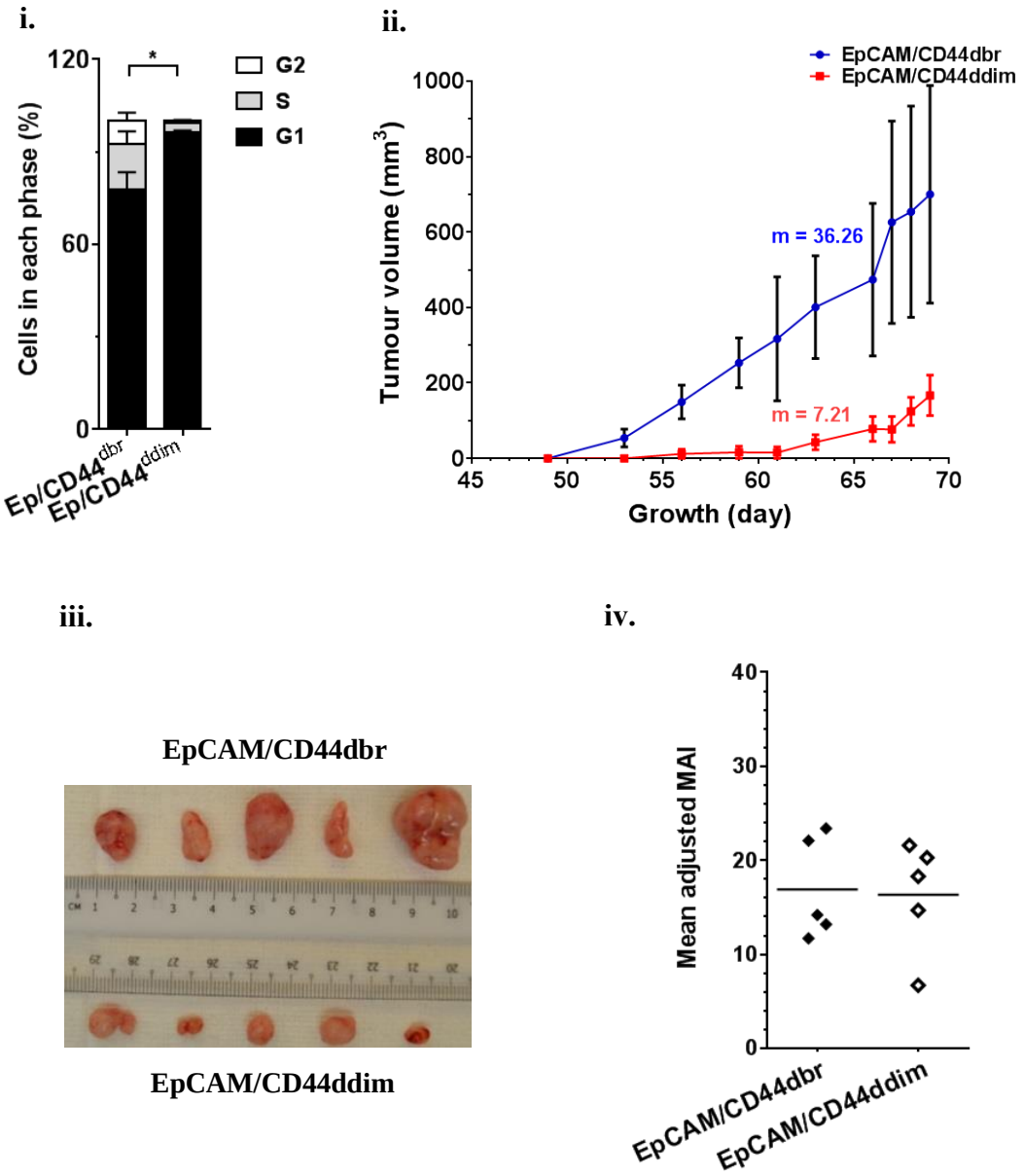


Figure 4.19: *In vivo* growth properties of EpCAM/CD44 cells from xeno-B110. **i.** Freshly-sorted and fixed EpCAM/CD44dbr and EpCAM/CD44ddim cells were analysed for their cell cycle distribution. Another batch of freshly-sorted EpCAM/CD44dbr and EpCAM/CD44ddim cells were inoculated into NSG mice and monitored for 69 days. **ii.** Growth curves of EpCAM/CD44dbr and EpCAM/CD44ddim xenografts, **iii.** image of the harvested xenografts on day 69 and **iv.** mean adjusted MAI per group between EpCAM/CD44dbr and EpCAM/CD44ddim xenografts. Results of cell cycle profile, mean ± SD of 3 flow cytometry experiment replicates. Results of growth curve, mean ± SEM of 5 xenograft replicates. m, rate of volume increase. * $p < 0.05$ (S-phase)

Overall, growth-related experiments assayed using xeno-B110 cells showed larger differences than C666-1 data. There are also numerous publications which demonstrated the advantages of using early-passage PDXs over long passaged cell lines in cancer studies (Chapters 4.2 and 4.6). Hence, subsequent characterization experiments were performed using xeno-B110 only.

4.5.4 *In vitro* growth of marker-selected xeno-B110 cells

Previous attempts to grow bulk/parental (no marker selection) xenograft cells from two other NPC PDXs C15 or C17 *in vitro*, either for short or long term, had been unsuccessful. Xeno-B110 cells transduced with green fluorescent (gfp) and luciferase (luc2) proteins, previously established for a separate project, were then used in this experiment to accurately determine *in vitro* growth ability of marker-selected xenograft cells. EpCAM-selected cells from this xenograft were employed as a study model.

Figure 4.20 describes the gating strategy used to sort for single, viable non-mouse GFP positive EpCAMbr and EpCAMdim cells. After removal of non-viable and non-mouse cells, the remaining cells ("COI" as named in the statistics box in Figure 4.20) were analysed for GFP and EpCAM expressions. EpCAMbr and EpCAMdim cells were sorted from the brighter GFP subpopulation.

Non-mouse ("mouse depleted") xeno-B110-gfp-luc2 cells were used as a growth control for the *in vitro* experiment (Figure 4.21). EpCAMbr, EpCAMdim and non-mouse xenograft cells showed positive growth during the first week of culture (up to day 6), with EpCAMbr cells having the fastest growth rate, followed by EpCAMdim and non-mouse cells. However, all xenograft cells were unable to sustain their growth under artificial conditions in the second week (after day 6). By day 11, luminescence levels for all three phenotypes had fallen to nearly the baseline level of day 0.

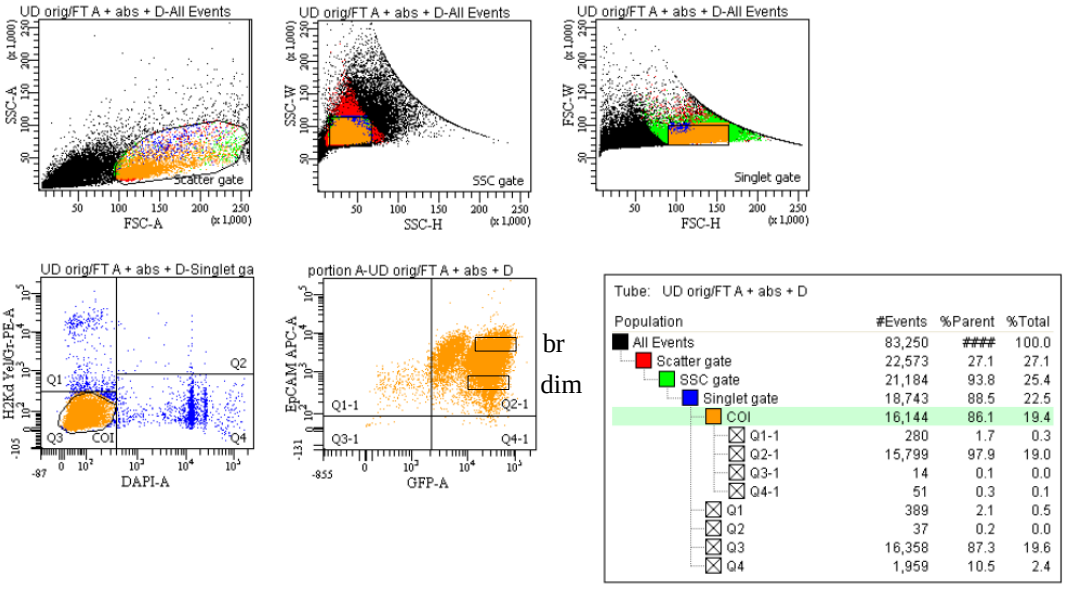


Figure 4.20: Flow cytometry dot plots detailing the gating strategy employed for sorting EpCAMbr and EpCAMdim cells from xeno-B110-gfp-luc2. Xeno-B110-gfp-luc2 cells were stained with H2Kd antibody conjugated to PE, EpCAM antibody conjugated to APC and viability dye DAPI and subsequently analysed in a hierarchical manner as depicted in the statistics box. Single, viable, non-mouse GFP positive EpCAMbr and EpCAMdim cells were collected for the *in vitro* growth experiment.

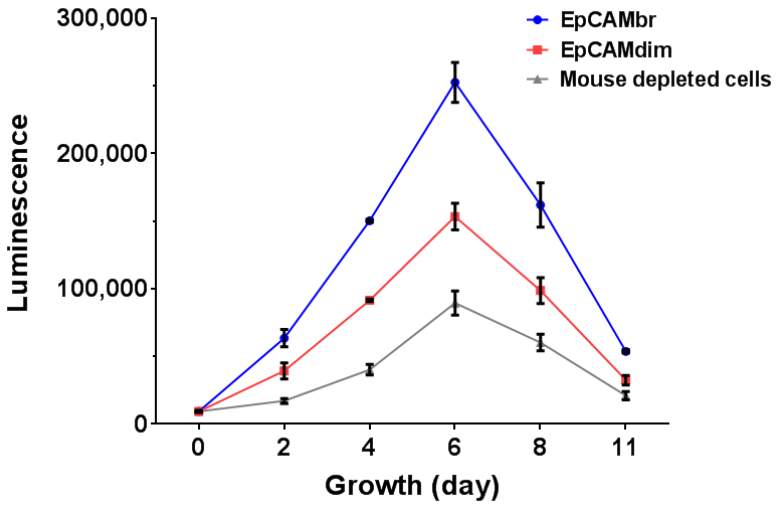


Figure 4.21: *In vitro* growth curves of non-mouse cells, EpCAMbr cells and EpCAMdim cells from xeno-B110-gfp-luc2. Non-mouse ("mouse depleted cells") and EpCAMbr and EpCAMdim cells were depleted or sorted from xeno-B110-gfp-luc2, respectively. The cells were seeded in complete medium and luminescence reading was taken after the addition of 2X D-Luciferin at different time points as shown above. Results, mean ± SEM of triplicate wells per phenotype from 1 representative sorting experiment.

4.5.5 Enrichment of tumour-initiating cells (TICs) by CD24, CD44 and EpCAM in the first generation of xeno-B110

Table 4.4 shows the results of a limiting dilution assay performed on CD24, CD44, EpCAM and EpCAM/CD44-selected cells from xeno-B110 from the highest inoculation of 30,000 cells to the lowest inoculation of 10 cells. It is observed that CD24br cells had approximately a 2-fold increase of TICs as compared to CD24dim cells ($p > 0.05$). CD44br cells were significantly 17.49-fold more enriched in TICs than CD44dim cells ($p < 0.001$). EpCAMbr cells had a 4.97-fold TIC enrichment over EpCAMdim cells ($p = 0.01$). Significant TIC frequency enrichment was also seen in EpCAM/CD44dbr cells (8.25 folds; $p < 0.01$). Nonetheless, it is noted that the enrichment fold of EpCAM/CD44dbr cells did not exceed that of CD44br cells.

Table 4.4: Limiting dilution assay for CD24, CD44, EpCAM and EpCAM/CD44 cells from xeno-B110 (first generation).

	Cell inoculation	**Number of tumours/Number of inoculated mice**								
		H2Kd neg	**CD24 br**	**CD24 dim**	**CD44 br**	**CD44 dim**	**EpCAM br**	**EpCAM dim**	**EpCAM/ CD44dbr**	**EpCAM/ CD44ddim**
xeno-B110	30,000	4/4	3/3	3/3	3/3	3/3	3/3	3/3	3/3	3/3
	10,000	4/4	6/6	6/6	5/5	5/5	5/5	5/5	6/6	6/6
	5,000	5/6	5/5	6/6	4/4	4/4	6/6	6/6	5/5	5/5
	2,000	ND	5/5	4/4	5/5	3/5	5/5	4/5	5/5	5/5
	500	3/3	3/3	3/3	3/3	3/3	3/3	2/3	3/3	2/3
	100	2/6	2/6	1/6	5/6	1/6	3/6	2/6	5/6	1/6
	10	0/6	1/6	0/6	2/6	2/6	0/6	0/6	1/6	0/6
Estimated TIC frequency (CI)		1 in 1177 (393-3526)	1 in 159 (63-404)	1 in 285 (108-754)	1 in 45 (18-110)	1 in 787 (333-1859)	1 in 147 (58-373)	1 in 730 (308-1728)	1 in 56 (23-135)	1 in 462 (181-1181)
p value			0.42		< 0.001		0.01		< 0.01	
Enrichment factor			1.79		17.49		4.97		8.25	

ND, not determined

4.5.6 Self-renewal property of CD24, CD44 and EpCAM marker-selected cells from xeno-B110

As CSCs are defined primarily by their ability to propagate long term, an *in vivo* serial transplantation experiment was performed using CD24, CD44 and EpCAM cells from xeno-B110 for three successive passages (up to the fourth generation). All groups of xenografts were re-sorted for respective phenotype of cells prior to re-inoculation into recipient NSG mice as secondary/tertiary/quartenary xenografts. Table 4.5 shows the TIC frequencies for all three markers at the fourth generation. All phenotypes (bright and dim from all markers) could self-renew by growing up to fourth generation although with different TIC frequencies. The histology of parental xeno-B110 cells was recapitulated throughout the four generations of xenografts (Figure 4.22).

Table 4.5: TIC frequency of CD24, CD44 and EpCAM cells from xeno-B110 (fourth generation).

	Cell inoculation	**Number of tumours/Number of inoculated mice**					
		CD24		**CD44**		**EpCAM**	
		br	**dim**	**br**	**dim**	**br**	**dim**
xeno-B110	500	3/3	1/3	3/3	2/3	3/3	2/3
	100	5/6	2/6	4/6	1/5	4/6	1/6
	10	1/6	0/6	2/6	0/6	0/6	0/6
Estimated TIC frequency (CI)		1 in 56 (23 – 135)	1 in 591 (177 – 1971)	1 in 67 (27 – 163)	1 in 474 (149 – 1511)	1 in 104 (42 – 261)	1 in 509 (162 – 1605)
p value		< 0.01		< 0.01		0.03	

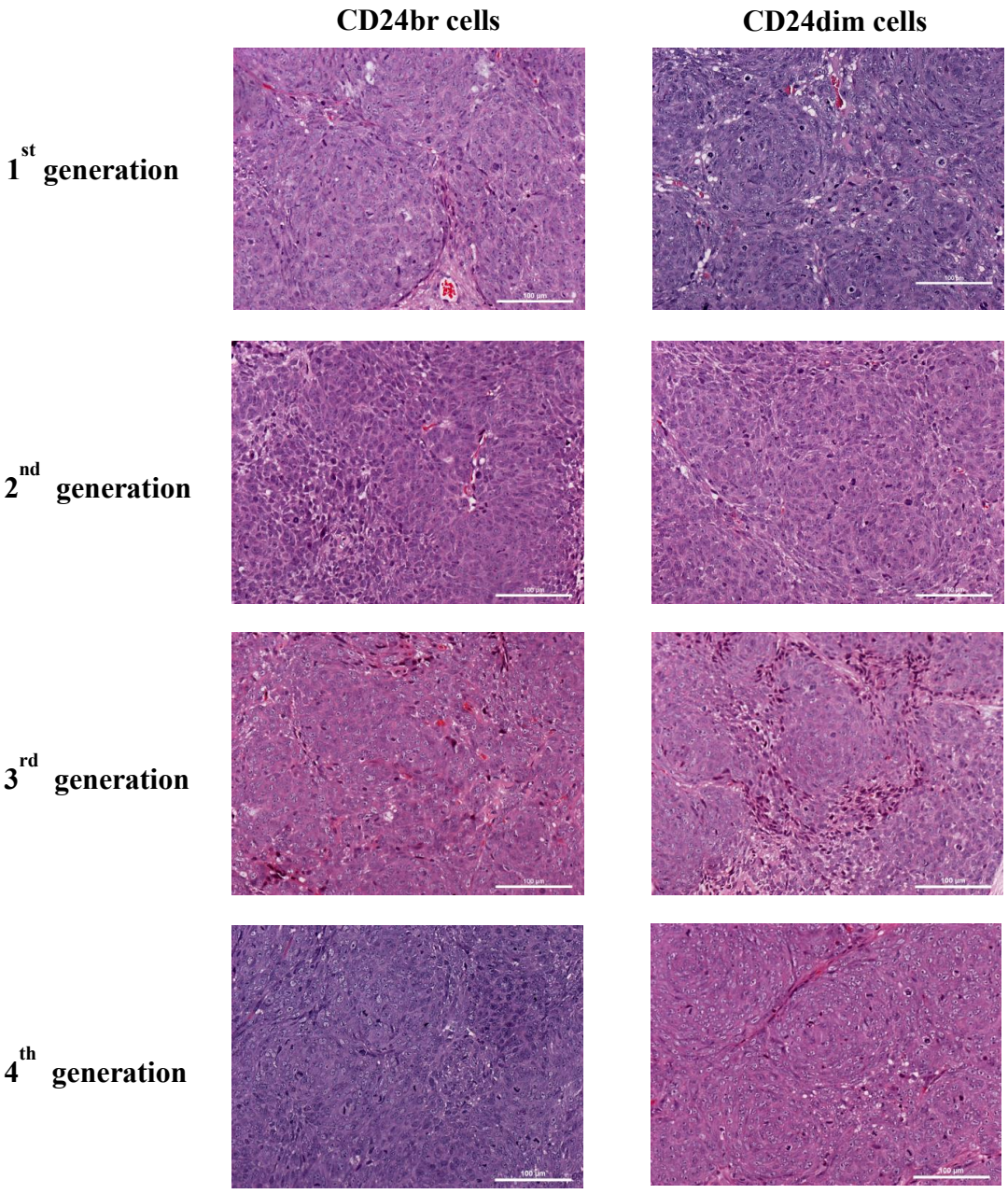


Figure 4.22: Representative H&E images of resulting xenografts from xeno-B110 CD24br and CD24dim cells (1st generation) and subsequent serially-passaged generations. CD24br and CD24dim cells from xeno-B110 were inoculated into recipient NSG mice and gave rise to the first generation of xenografts. The harvested CD24br xenografts were stained and re-sorted for CD24br cells prior to inoculation into NSG which formed the second generation. Serial passaging was thus continuously performed until the fourth generation for both cell phenotypes. All generations displayed similar histology of non-keratinizing differentiated NPC as the parental xeno-B110. (20X objective; scale bar 100 μm).

4.5.7 Relative quantification of cell cycle, proliferation, pluripotency and stemness-related genes in CD24, CD44, EpCAM and EpCAM/CD44 marker-selected cells from xeno-B110

The expression levels of 25 genes were measured in freshly-sorted marker-selected xeno-B110 cells (Appendix E). Seventeen assays (three reference genes, three genes coding for CSC markers used in this study and 11 genes associated with cell cycle, proliferation, pluripotency and stemness) passed quality control assessment while the remaining eight assays had poor PCR efficiency and/or showed no or nonlinear amplification. Dot plots showing the fold changes of 14 genes of interest for CD24, CD44, EpCAM and EpCAM/CD44-selected cells which passed quality control assessment are shown in Appendix F.

Although xeno-B110 is EBV positive as evident by positive EBER-ISH staining (Figure 4.14iii), *LMP1* and *LMP2A* mRNA transcripts from EBV were below the detection limit of the experiment (Appendix E). As expected, there was a more than 2-fold difference of *CD24* mRNA transcript between CD24br and CD24dim cells (Figure 4.23i). Similarly, *EPCAM* level was more than 2-fold enriched in EpCAMbr cells than EpCAMdim cells (Figure 4.23ii). The expression level of *CD44* was only slightly increased in CD44br cells as compared to CD44dim cells (Figure 4.23iii).

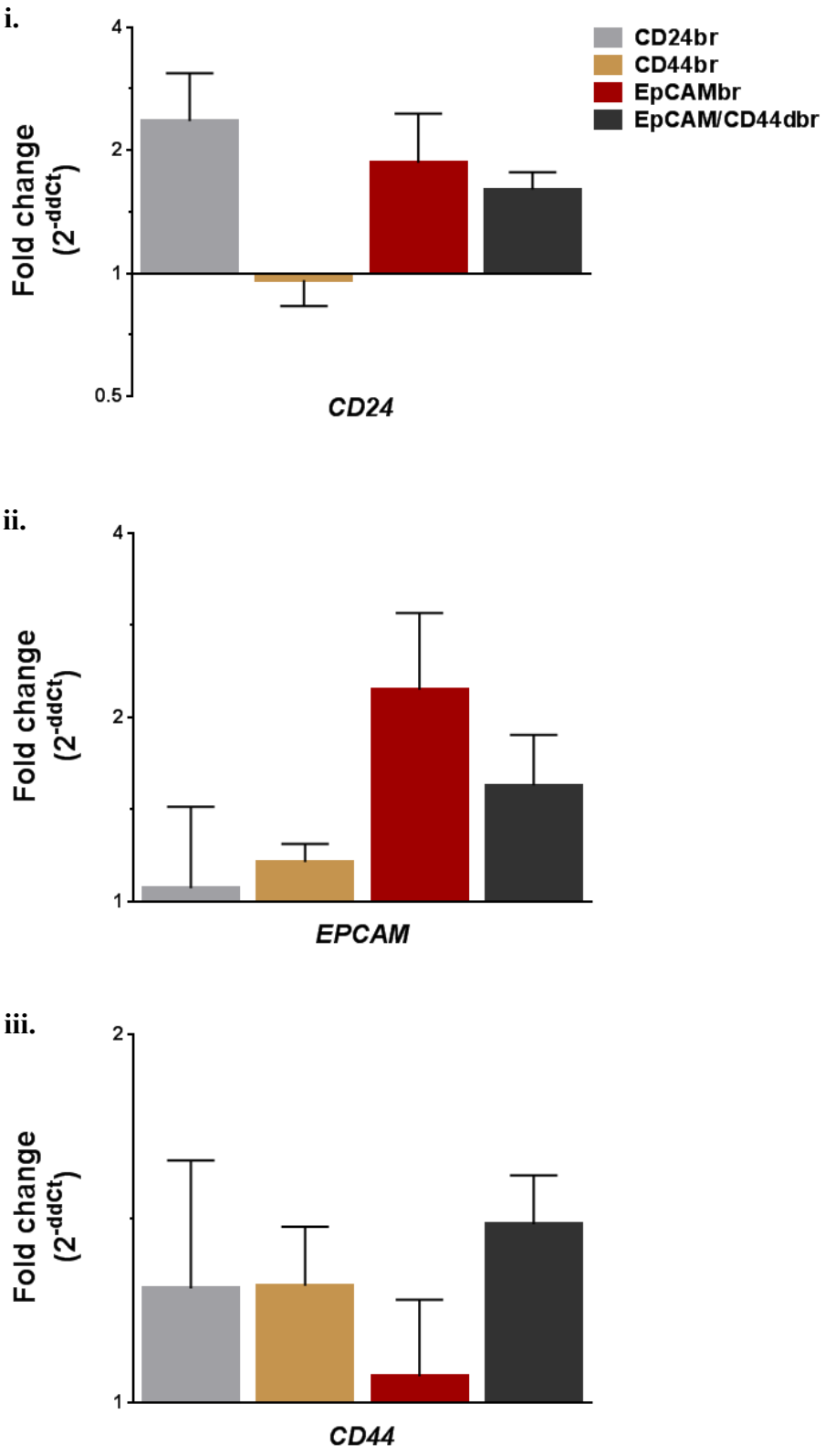


Figure 4.23: Gene expression levels of *CD24*, *EPCAM* and *CD44* in CD24br, CD44br, EpCAMbr and EpCAM/CD44dbr cells from xeno-B110. Fold changes of *CD24*, *EPCAM* and *CD44*genes were calculated as ratio of normalized gene expression of bright (dbright) cells to normalized gene expression of dim (ddim) cells of each marker. Results, mean ± SD of 2 or 3 sorted cell sample replicates.

Pluripotency gene, *KLF4* was significantly downregulated in CD24br ($p = 0.02$) and EpCAMbr cells ($p < 0.01$) compared to their respective dim phenotype (Figure 4.24i). EpCAM/CD44dbr cells also showed a downregulation trend for *KLF4* transcript ($p > 0.05$). *CDKN1A*, a cell cycle inhibitor as well as a transcriptional target of KLF4 protein was also downregulated in the bright phenotypes of EpCAM and EpCAM/CD44 ($p > 0.05$) (Figure 4.24ii). *VIM*, a marker for epithelial-mesenchymal transition (EMT), was moderately upregulated in CD24br cells ($p > 0.05$) (Figure 4.24iii).

Cell cycle regulators were also deregulated in the freshly-sorted marker-selected xeno-B110 cells. *CCND1* transcript was significantly downregulated in EpCAMbr cells ($p = 0.02$) (Figure 4.25i). There were more than 6 folds of *CCNE1* transcript in EpCAM/CD44dbr cells than EpCAM/CD44ddim cells ($p > 0.05$) (Figure 4.25ii). Higher *MKI67* transcript level was seen in CD44br cells than in CD44 dim cells ($p = 0.03$) which was also observed in EpCAM/CD44dbr cells as compared to EpCAM/CD44ddim cells ($p > 0.05$); however, it was downregulated more than 1.5 folds in CD24br cells ($p < 0.01$) (Figure 4.25iii).

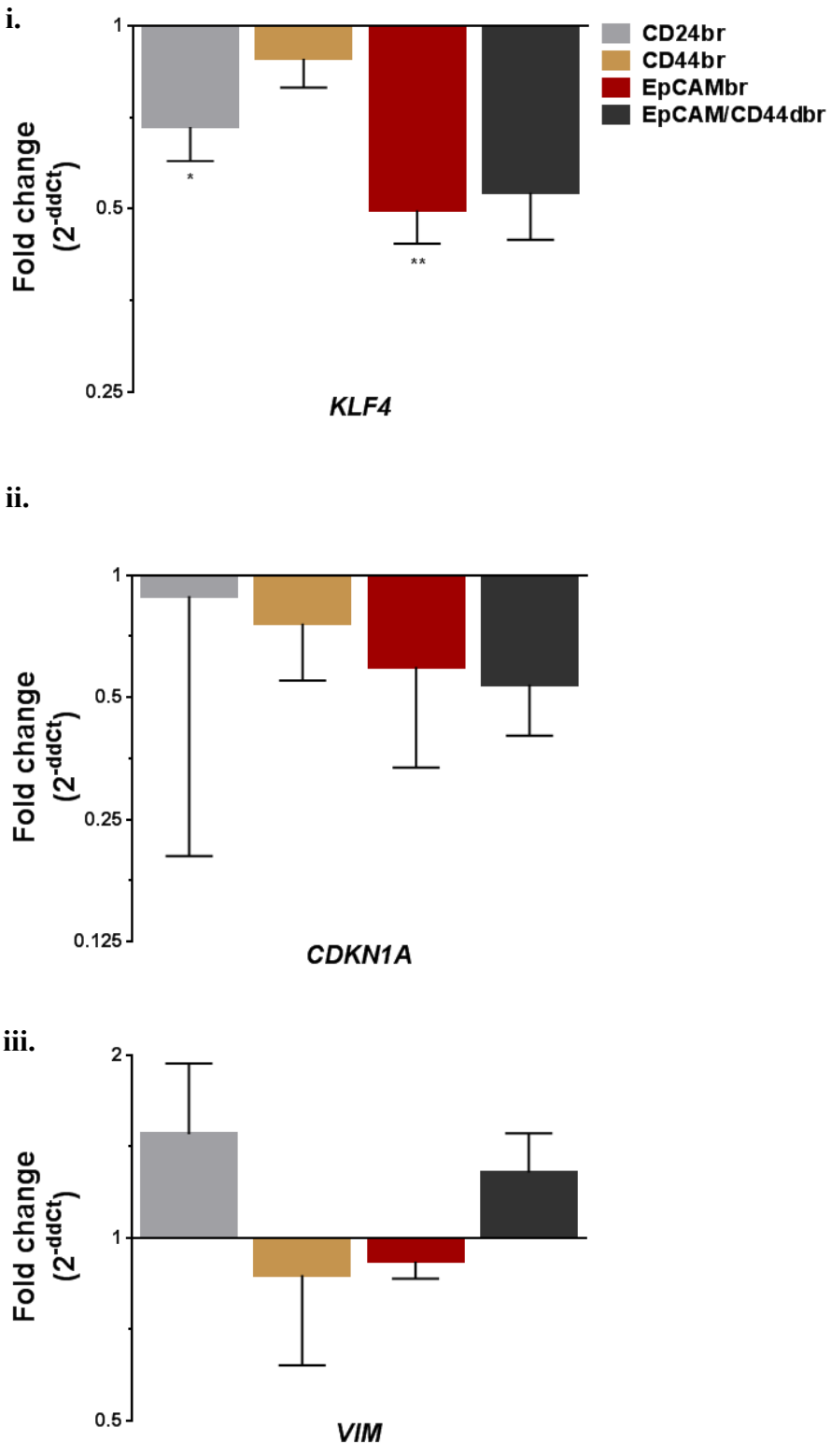


Figure 4.24: Gene expression levels of *KLF4*, *CDKN1A* and *VIM* in CD24br, CD44br, EpCAMbr and EpCAM/CD44dbr cells from xeno-B110. Fold changes of *KLF4*, *CDKN1A* and *VIM* genes were calculated as ratio of normalized gene expression of bright (dbright) cells to normalized gene expression of dim (ddim) cells of each marker. Results, mean ± SD of 2 or 3 sorted cell sample replicates.
* $p < 0.05$, ** $p < 0.01$

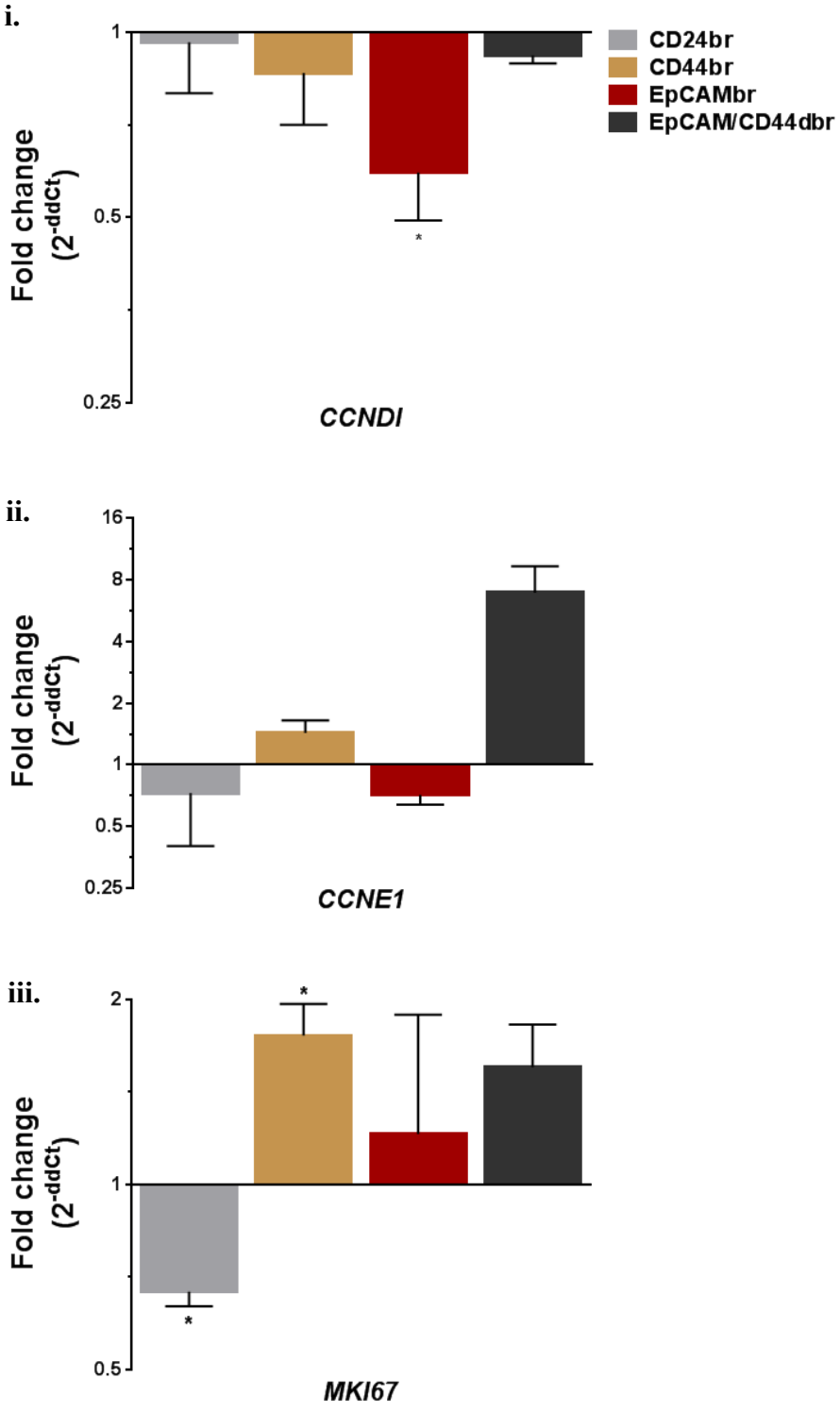


Figure 4.25: Gene expression levels of *CCND1*, *CCNE1* and *MKI67* in CD24br, CD44br, EpCAMbr and EpCAM/CD44dbr cells from xeno-B110. Fold changes of *CCND1*, *CCNE1* and *MKI67* genes were calculated as ratio of normalized gene expression of bright (dbright) cells to normalized gene expression of dim (ddim) cells of each marker. Results, mean ± SD of 2 or 3 sorted cell sample replicates. * $p < 0.05$

4.6 Discussion

This study attempts to illustrate the biological properties of NPC cells identified by three CSC markers concurrently using the same samples, namely CD24, CD44 and EpCAM which are widely reported markers for isolation of tumourigenic cells (Al-Hajj et al., 2003; Han et al., 2011; Hiraga et al., 2016; Prince et al., 2007; Visvader & Lindeman, 2012). Firstly, CD44br and EpCAMbr cells from both C666-1 and xeno-B110 consistently enriched for faster-growing tumourigenic cells, resulting in larger tumour growth. More notable growth differences within a group, regardless of markers, were seen in xeno-B110 cells. Secondly, higher TIC frequencies were observed in CD44br and EpCAMbr cells of xeno-B110 within the first generation of marker-induced growth. Thirdly, cells co-stained with a combination of CD44br and EpCAMbr markers ("EpCAM/CD44dbr marker") did not show additional enhancement of *in vivo* growth or had higher TIC frequency than cells stained with CD44br marker alone. Fourthly, CD24br, CD44br and EpCAMbr cells of xeno-B110 displayed self-renewal ability throughout three successive serial transplantation passages. Finally, there was an inverse relationship between *in vivo* tumour growth and the expression of *KLF4* gene, i.e. CD24br, CD44br, EpCAMbr and EpCAM/CD44dbr cells of xeno-B110 consistently induced larger tumour growth than their respective dim phenotypes with downregulated *KLF4* transcripts in the CD24br, EpCAMbr and EpCAM/CD44dbr cells.

Immunophenotyping of CD24 and CD44 in NPC had been performed in HK1 and C666-1 cell lines. Yang et al. (2014b) reported there were 0.86% of CD24 positive and 16.30% of CD44 positive cells in HK1, whereas there were 5 to 45% of CD44 positive cells in C666-1 as published in two independent studies (Janisiewicz et al., 2012; Lun et al., 2012). This study found that regardless of C666-1 maintained and propagated as a cell line or a xenograft, CD24 was barely detected. Also, there were slightly more than 60% of CD24 positive and nearly 100% of CD44 positive cells in HK1, and more than

90% of CD44 positive cells in C666-1. The variations in CD24 and CD44 expression levels in HK1 and C666-1 are believed to have arisen from technical differences such as culture conditions, the type of enzyme and the length of incubation used during detachment of the adherent cell lines as well as gating strategies used to derive the reported percentage of positive cells (Golebiewska et al., 2011; Greve et al., 2012; Khan et al., 2015). The overall high expression levels of CD24, CD44 and EpCAM presented in Figure 4.5 cannot be equated to *bona fide* presence of NPC cells with stemness properties in these samples without further empirical experiments, as can be clearly seen in the later part of this study. Numerous CSC meetings and review papers have pointed out that CSCs cannot be identified by mere phenotypic descriptions of presence/absence or high/low level of markers; they must be functionally defined by key abilities to self-renew, differentiate and induce tumours (Clarke et al., 2006; Dashzeveg et al., 2017; Valent et al., 2012).

An earlier study using C666-1 reported that the minimum number of unselected parental C666-1 cells able to initiate tumour growth in nude athymic mice was 500,000 cells (83%; 5/6 with latency of 28 days) (Lun et al., 2012). This study found that as few as 2,000 cells of C666-1 could form tumours *in vivo* (100%; 5/5 with latency of 42 days). There are two major differences between both these studies which attributed to the data disparity: 1) cell sorting technique was used to select for single viable parental C666-1 cells prior to inoculation in this study as compared to Lun *et al.*'s use of unselected C666-1 cells, and 2) NSG mice were used in this study as compared to inoculation into nude mice strain in Lun and colleagues' study.

Faced with an inherent lack of fresh NPC specimens and when available, limited NPC tissue size, most NPC investigations were and are still performed using NPC cell lines and/or archival NPC tissue sections. As pointed by Fernando et al. (2006) and

Cree et al. (2010), there are several disadvantages with using cancer cell lines. Cancer cell lines are grown in plastic dishes with serum or other growth factor supplements to enable them to survive in an artificial environment, unlike cancer cells which are naturally capable of independent growth with reduced attachment to substrate and other cells. Also, the former exhibit uniformity as they are largely consisting of a homogeneous population of rapidly-proliferating cells after numerous passaging cycles often in serum-supplemented media. On the other hand, a high level of concordance between PDXs (passages 6 to 12) and the corresponding patient's tumour was verified in a large scale PDX study on colorectal cancer (Julien et al., 2012). Their finding of a high genomic stability within the first 10 generations of PDXs was also shared by an evaluation of esophageal and gastro-esophageal junction cancer PDXs (Dodbiba et al., 2015). A recent comparative study on molecular similarities of patient tumours, PDXs and cancer cell lines reported highly similar differentially expressed genes between patient tumours and PDXs, in contrast to cancer cell lines which had much poorer expression similarity to patient tumours and PDXs (Guo et al., 2016). The relevance of PDX model as a surrogate for freshly-isolated cancer cells is reflected in the increasing numbers of publications, especially in CSC studies, which used freshly-isolated cancer cells and/or early-passage PDXs to obtain salient information, instead of solely deriving their data from cancer cell lines (Facompre et al., 2016; Guo et al., 2016; Samaeekia et al., 2017). In view of these evidences, early-passage PDX cells (i.e. passages 5 to 9) were used throughout this study.

The high *in vivo* growth seen in parental xeno-B110 cells (Table 4.2) may be attributed to the probability of xeno-B110 itself being highly tumourigenic and also to the use of NSG mice in this study. NSG mice are mutated in the severe combined immune deficiency (*scid*) gene and have a complete null allele of the IL2 receptor common gamma chain ($IL2rg^{null}$). Both mutations caused the mice to be deficient in

mature lymphocytes (B and T cells) and NK cells (Shultz et al., 2005). In addition, they are reported to have a median survival time of more than 89 weeks and were "non-leaky" at more than 1 year of age, both of which are crucial determinants for long term *in vivo* experiments. NSG mice can also provide highly efficient engraftment of exogenous cells as demonstrated by the seminal work of Quintana et al. (2008). The use of BD Matrigel™ basement membrane matrix extracted from the Engelbreth-Holm-Swarm mouse sarcoma as a co-inoculation agent in this study may have also improved tumour formation from as low as 100 cells. The ability of such low cell numbers to initiate tumours has been reported elsewhere. A 5-cell inoculation of melanoma cells mixed with Matrigel and injected into NSG mice had a tumour formation efficiency of 39% (7/18) (Quintana et al., 2008). One hundred cells each of Matrigel-mixed $CD44^+$ subpopulation and triple positive subpopulation of CD44/CD24/ESA combination from pancreatic cancer PDX formed tumours at an efficiency of 25% (4/16) and 50% (6/12), respectively (Li et al., 2007).

Biological differences in marker-selected cells were largely reflected in the growth and self-renewal data of this study. Compared to CD44br, EpCAMbr and EpCAM/CD44dbr, CD24br cells in xeno-B110 were not enriched with proliferative cells and had decreased expression for proliferation marker *MKI67*. Nonetheless, TIC frequency at the final serial transplantation cycle *in vivo* at the fourth generation signified that CD24br cells were enriched with TICs and could sustain long term self-renewal. *KLF4* transcript was previously reported to be inversely correlated with epithelial-mesenchymal transition (EMT) *SLUG* transcript in hepatocellular carcinoma (Lin et al., 2012). In this study, gene expression analysis of CD24br cells showed a lower presence of *KLF4* transcript with an upregulated *VIM* expression, another EMT marker. The interconnection of EMT, stem-like and tumour initiation had been highlighted before. Twist- or Snail-expressing non-tumourigenic human mammary

epithelial cells underwent EMT and generated progenies which managed to form mammospheres and tumours more efficiently (Mani et al., 2008). As such, self-renewing and EMT marker-expressing xeno-B110 CD24br cells could be CSC-like.

The abilities of $CD24^{+}$ cells to induce tumours and to form passageable spheres had been shown to be related to STAT3-mediated upregulation of *NANOG* expression in a study using chemoresistant hepatocellular carcinoma PDXs (Lee et al., 2011). CD24 as a potential CSC marker for NPC has received lesser attention than CD44. $CD24^{+}$ enriched for only stem-like properties in NPC cell lines TW02 and TW04 as CSC functionality i.e. *in vivo* self-renewal was not verified (Yang et al., 2014b). $CD44^{high}CD24^{high}$ cells from three NPC cell lines (TW01, TW06 and HONE1) displayed stem-like and EMT phenotypes as compared to parental unsorted cells (Shen et al., 2016). Upon knockdown of both CD44 and CD24 in the cells, sphere formation, chemoradioresistance, wound healing and invasion properties were suppressed. Overexpression of CD44 and CD24 in the NPC cell lines, which restored the CSC-like features, was found to be related to the activation of STAT3. Conversely, $CD24^{low}$ and $CD44^{high}$ LMP1-transfected nasopharyngeal epithelial cells produced spheres *in vitro*, an attribute associated with stemness (Kondo et al., 2011). As the above studies made use of cell lines with unknown authenticity, it is suggested that CD24br cells from ideally fresh NPC specimens or early-passage NPC PDXs to be evaluated for tumour initiation as well as *in vivo* self-renewal ability in more serial transplantation passages. This is to demonstrate with more conviction whether firstly, CD24 is a CSC marker for NPC and secondly, it is a better CSC marker than the currently regarded CD44 for NPC (Yang et al., 2014a).

CD44 is amongst the most investigated CSC markers in NPC (Lun et al., 2014); nonetheless, it was largely examined in NPC cell lines such as C666-1, SUNE-1 and

CNE-1 (Lun et al., 2012; Su et al., 2011; Yang et al., 2014a). The most striking differences between this study and current published reports are the use of an early-passage NPC PDX as a study model, and the verification of self-renewal ability, central to (cancer) stem cell's definition, using serial transplantation experiment following the recommendations of AACR Workshop on CSCs (Clarke et al., 2006), instead of using spheroid or colony-forming assays (Fillmore & Kuperwasser, 2008; Xu et al., 2015; Yang et al., 2014a). Tumour initiation and *in vivo* growth data from CD44 xenografts of C666-1 and xeno-B110 concurred with prior studies. CD44br cells were also proliferative in nature consistent with remarkably high S-phase cell content and increased *MKI67* mRNA transcripts. Additionally, they were able to self-renew *in vivo* for at least three passages.

Although this study did not find differential *BMI1* expression between CD44br and CD44dim cells in xeno-B110 (Appendix F), tumourigenicity ability of $CD44^+$ cells had been attributed to the functionality of self-renewal marker *BMI-1* (Xu et al., 2016b). Upon inoculation into nude mice, only $CD44^+$ cells from SUNE-1 5-8F had tumour formation (4/4) whereas $CD44^-$ and BMI-1-knockdown $CD44^+$ cells did not form any growth (0/4 for both groups). Two drawbacks of Xu and colleagues' report are the use of SUNE-1 5-8F cells with unknown authenticity, and the relatively short observation period of four weeks for a CSC *in vivo* study. Tumour generation *in vivo* generally takes longer to occur as reported by Quintana et al. (2008) who found that most human melanoma cell-induced xenografts in NOD/SCID developed after eight weeks of inoculation. Parental and marker-selected C666-1 cells, and non-mouse and marker-selected xeno-B110 cells in this study recorded mean latency of more than 35 and 60 days, respectively. Therefore, the possibility of underestimating no-xenograft growth effect of BMI-1-knockdown $CD44^+$ cells remains. Nonetheless, there was a

concurrence of *in vivo* data between the study by Xu and colleagues and this study, i.e. $CD44^+$ or CD44br marker enriched for TIC in NPC.

EpCAM has been in use as a CSC marker in breast, colon and pancreatic cancers since 2000s (Gires et al., 2009). Clinically, high expression of EpCAM was also notably associated with higher gastric carcinoma cell proliferation and disease progression (Kroepil et al., 2013). Concurring with these studies, EpCAMbr marker in xeno-B110 identified for fast-growing as well as serial-transplantable cells with stable TIC frequency at both first and fourth generations. In a study to evaluate the functions of β-catenin in stemness maintenance and tumourigenicity of NPC (Jiang et al., 2016), stable suppression of β-catenin in a CNE-2 cell line ("pLKO.1-sh-β-catenin-CNE2") led to reduction of EpCAM, vimentin, NANOG and SOX2 proteins in addition to SP cells. Amongst other findings, pLKO.1-sh-β-catenin-CNE2 cells were shown to contain lower percentage of S-phase cells and higher E-cadherin expression with arising tumours also significantly smaller than those of control and parental cells. Association of lower EpCAM level with fewer cells in S-phase as well as smaller and slower proliferating tumours *in vivo* are also seen in this study.

The role and links of EpCAM to cell proliferation were extensively investigated by Gires's group. EpCAM was proven to have directly induced proliferation by upregulating expression of MYC at both mRNA and protein levels as well as cyclin A and cyclin E proteins (Münz et al., 2004). EpCAMbr cells of xeno-B110 functionally showed a positive relationship with proliferation (i.e. faster xenograft growth and higher percentage of S-phase cells) although upregulation of *CCNE1* and *MYC* transcripts were not observed. This suggested that either upregulation of cyclin E and c-myc proteins in NPC cells is at the post-transcriptional level or proliferation in EpCAMbr NPC

xenografts may be via a different mechanism other than the involvement of cyclin E and c-myc.

As EpCAMbr cells and CD44br cells of xeno-B110 independently appeared to have the highest proliferation and TIC enrichment in the first generation of CSC marker-induced xenografts in this study, it was hypothesized that the combination of EpCAM/CD44dbr marker would identify for cells which were even more proliferative and/or enriching for TICs than these single markers. However, there was no difference in the proliferation of EpCAM/CD44dbr xenografts as compared with CD44br and EpCAMbr xenografts. Although TIC frequency of EpCAM/CD44dbr xenografts in the first generation was higher than that of EpCAMbr marker, it was comparable to CD44br marker. The failure of EpCAM/CD44dbr marker in this study to outperform CD44br marker in further enriching for proliferative cells or cells with higher TIC frequency may indicate that CD44 individually exerted a stronger influence than EpCAM in tumour growth.

Absent or weak *LMP1* expression in NPC tumours had been reported elsewhere (Bell et al., 2006; Lo, 2016). The expression of *LMP2A* was inconsistent from Ct 22 to Ct 999 (undetected) in the replicates of marker-selected xeno-B110 which led to a mean calculation of high Ct values in each marker. However, in the concurrently run of pooled RNA positive control comprising of C666-1, HK1, HONE1, B95.8, SW480, Namalwa and MDA231, *LMP1* transcript was detected with a mean Ct of 12.34 to 17.90 and *LMP2A* was also similarly detected with a mean Ct of 15.65 to 21.60. Hence, the overall low expression of both latent genes ("beyond the detection limit" of qPCR platform) in marker-selected xeno-B110 cells may be attributed to heterogeneous expression of the LMPs in xeno-B110 as seen in Figure 4.14iii. Sequences for both LMP1 and LMP2A assays (primers) had been reported in Wasil et al. (2013).

There are a few lines of evidence in this study which imply that KLF4 is anti-proliferation in xeno-B110. Bright phenotypes of CD24, CD44, EpCAM and EpCAM/CD44 caused larger xenograft growth, were more actively proliferating and had higher percentages of S-phase cells. CD24br, EpCAMbr and EpCAM/CD44dbr cells also expressed decreased *KLF4* transcripts than their respective dim cells, and *CDKN1A*, a target gene of KLF4, was also downregulated in EpCAMbr and EpCAM/CD44dbr cells. The data are in concordance with others. Overexpression of KLF4 in a mouse mammary tumour xenograft resulted in limited tumour formation (Yori et al., 2011). Liu et al. (2013) found that *KLF4* mRNA level was lower in NPC specimens than normal nasopharyngeal tissues, and NPC patients with larger tumours had slightly decreased *KLF4* level than those with smaller tumours. In pancreatic cancer, KLF4 knock-out mice grew larger tumours than mice with wild type KLF4 (Yan et al., 2016). The inverse relationship between *KLF4* level and cell proliferation is consistent with the report of KLF4 engaging p53 to activate the transcription of *CDKN1A* gene which encodes p21 (also known as WAF1 or CIP1), in turn leading to cell cycle arrest (McConnell & Yang, 2010).

4.7 Conclusion

This study has demonstrated that CD44br and EpCAMbr cells were fast-growing and more tumourigenic than their respective dim phenotype with resulting larger tumour growth in frequently used NPC cell line C666-1 and in early-passage NPC PDX xeno-B110. The combination of CD44br and EpCAMbr markers, however, did not further enrich for more fast-growing or tumourigenic cells. Larger growth differences were observed in xeno-B110 than in C666-1 marker-selected cells. CD24br, CD44br and EpCAMbr cells isolated from xeno-B110 were also enriched for TICs and they retained self-renewal property upon serial transplantation *in vivo*. *KLF4* was downregulated in CD24br, EpCAMbr and EpCAM/CD44dbr cells of xeno-B110. With observational

data that *KLF4* is anti-proliferation, this study provides preliminary evidence of an association between variations of *KLF4* and tumourigenicity of NPC subpopulations identified by CD24, CD44 and EpCAM.

CHAPTER 5: GENERAL DISCUSSION AND CONCLUSION

5.1 General discussion

This study evaluated the use of two approaches (SP and CSC markers) to identify tumourigenic NPC cells in two NPC study models (cell line and PDX). Using SP approach, both SP and NSP cells in HK1 cell line were similar in tumourigenic potential and TIC frequency, despite showing (significant) differences in *in vitro* growth-related experiments as well as gene expression data. On the other hand, CSC markers (CD24, CD44 and EpCAM) were able to differentiate C666-1 cell line and xeno-B110 PDX into subpopulations of unequal tumourigenicity abilities as well as different TIC frequencies. These *in vivo* growth differences were also reflected in cell cycle, mitotic activity and gene expression. Although all subpopulations were able to induce tumour formation, contrasting self-renewal capabilities between bright and dim phenotypes of each marker surfaced during long term propagation.

To date, there is no one single identification method or marker which is exclusively responsible for selecting tumourigenic cells from NPC cells. The lack of a universal and "failproof" identification method or marker to identify tumourigenic cells within a malignancy has been reported elsewhere. For example, Wu & Alman (2008) remarked that some cell lines lacking SP cells were equally capable of initiating tumours. Certain CSC markers such as CD133 reportedly showed discordance in tumour formation ability within the same malignancy. CD133 was first reported to be capable of distinguishing tumourigenic from non-tumourigenic brain tumour cells: as few as 100 $CD133^+$ cells produced a tumour but not from 10^5 $CD133^-$ cells (Singh et al., 2004). Subsequent independent studies revealed equal tumourigenic activities in $CD133^+$ and $CD133^-$ brain tumour cells (Beier et al., 2007; Chen et al., 2010; Joo et al., 2008).

In discussing tumourigenesis, it would be remiss not to briefly mention of tumour heterogeneity and cell plasticity here although both topics are out of the scope of this study. Tumour heterogeneity may be seen at both intra- and inter-tumour levels, whereby different subpopulations of tumour cells show distinctive morphology, growth as well as phenotype and genotype profiles (Jamal-Hanjani et al., 2015; Marusyk & Polyak, 2010). Hence, different identification methods, verification experiments and subsequent comparison of data may be required to prove the existence of different tumour clones. Cell plasticity, on the other hand, describes the ability of a cell to take on characteristics of another cell either in a uni- or bi-directional route (Meacham & Morrison, 2013). When and if a cell shows plasticity, the transition between tumourigenic and non-tumourigenic cells and vice-versa will have a profound consequence on the outcome of an *in vivo* transplantation experiment as the experiment may no longer be able to distinguish between these two categories of cells.

Although human tumourigenic cell and CSC are both defined functionally, preferably in animal models, there are some concerns regarding xenotransplantation. Firstly, most tumour cell transplantations are performed in immunocompromised mice which cause no or limited inflammation (van Staveren et al., 2009), as opposed to cancer-related inflammation (CRI) seen in patients (Colotta et al., 2009). CRI which includes infiltration of white blood cells and other immune cells such as tumour-associated macrophages and cytokines, is regarded as a key component of tumours. Secondly, the uptake and growth of implanted tumour cells or pieces are variable in mice and this may be due to a lack of or incompatible microenvironmental cues in mouse stroma (Williams et al., 2013). To circumvent the lack of human stromal and haematopoietic components seen in immunocompromised mice, genetically engineered mouse models (GEMMs) are established; however, generation of GEMMs with more than one driver mutation necessary to drive efficient tumourigenesis are very tedious

and laborious (Williams et al., 2013). Thirdly, processing of PDXs into single cell suspension prior to isolation of tumour cells can present a multitude of challenges. Commonly employed physical and enzymatic dissociation methods can be harmful to tumour cells if they are not first optimized in a pilot study. Certain CSC markers such as CD8, CD24 and CD44 may be more sensitive to enzymes resulting in under-representation in post-dissociation immunophenotyping (Autengruber et al., 2012; Quan et al., 2012). This study did not verify the effects of Accutase and collagenase type II on surface marker expression. However, independent comparative studies found that Accutase preserved the highest levels of CD24 and/or CD44 expression amongst the common types of cell detachment enzymes used in cell culture (Paebst et al., 2014; Quan et al., 2012). Collagenase type II had also been used successfully to digest solid tumours for CD44 and ALDH expression studies (Lau et al., 2014; Wang et al., 2011). Besides, post-dissociated and sieved cell suspension contains cell aggregates and dead cells which will make analysis and/or separation of subpopulations difficult. It is thus critical to ensure that stringent gating strategies be used to discriminate single viable cells from cell aggregates and dead or dying cells by flow cytometry techniques prior to analysis or cell sorting for downstream verification experiments.

5.2 Study limitations

This study has several limitations, namely:

i. a dearth of fresh and sizable tissue specimens which are confirmed to be NPC for immediate processing into cell suspension prior to immunophenotyping analysis and/or cell sorting for downstream experiments. All collaborating local hospitals within the vicinity of the laboratory do not offer surgery as a treatment regimen for NPC. Instead, patients with confirmed NPC are sent for radiotherapy which offers a more

promising treatment outcome as NPC is highly radiosensitive (Chan, 2010; Chua et al., 2016).

ii. as a result of the above, this study and most of previous studies on NPC elsewhere had to rely heavily on NPC cell lines as a study model for biological-based investigations. Yet, ever since the findings of HeLa partial genomes present in and possible mistaken identities of several frequently used NPC cell lines (Chan et al., 2008; Strong et al., 2014), there remains only two verified commonly used NPC cell lines, namely HK1 (EBV negative) and C666-1 (EBV positive).

iii. this study commenced with using SP approach to identify tumourigenic NPC cells in EBV negative samples (HK1 and xeno-284) as xeno-B110 (EBV positive) was yet to be established. C666-1 (EBV positive) could not be successfully analysed in FACSAria during the SP study, due to immense cell clumping and thereafter clogging of the sample line in the equipment. Upon the publication of HK1 SP study in Cancer Cell International, xeno-B110 was established and successfully maintained *in vivo*. Armed with better technical knowledge of sample preparation and an improved ability to use FACSAria for analysis and cell sorting of clumpy cells, biological properties of tumourigenic NPC cells were next examined using CSC marker approach in C666-1 and xeno-B110 samples as the majority of NPC cases are EBV positive.

iii. with the exception of limiting dilution (TIC frequency) and serial transplantation assays, *in vivo* growth endpoint experiments were conducted for 50 days (SP approach with HK1 cells; Table 3.1), 52 days (CSC marker approach with C666-1 cells; Table 4.1) and 69 days (CSC marker approach with xeno-B110 cells; Table 4.3). As the observation

period was kept short due to unequal tumour burden in the mice, there remains a small possibility of under-reporting in the growth data.

v. one of the biological characteristics investigated in this study is self-renewal ability for long term tumour propagation. Preferably, this ability is examined in patient specimens or in study models as similar as possible to the former. As discussed in Chapter 4, PDXs are the next-best study model for NPC due to insufficient fresh tissue specimens. Although there are at least three other known NPC PDXs in addition to xeno-284 and xeno-B110 (i.e. C15, C17 and C18), the former had been established for more than two decades with numerous *in vivo* passages which made them unsuitable as a study model for self-renewal ability.

vi. at the current stage of thesis-writing, xeno-B110 cells still cannot be adapted to grow well *in vitro*, as shown in Chapter 4. With a logarithmic growth phase lasting for less than a week, it poses tremendous challenges for drug inhibition or gene modulation experiments to study and/or verify the effects of *KLF4* and other genes of interest in tumourigenicity.

vii. a certain level of heterogeneity in xeno-B110 was seen in the RT-qPCR data, in which there was at least one sorting replicate having a different expression level than the other replicates for a particular transcript (Appendix F). The so-called outlier also differed from transcript to transcript. Also, there was a lack of available xeno-B110 at the time of experiment which led to a minimal number of replicates in the RT-qPCR experiment.

5.3 Future direction of study

This study provides a proof-of-concept research which highlights the potential of CD24, CD44 and EpCAM as markers to identify tumourigenic NPC cell with stemness-

related capabilities in a NPC PDX model. In order to establish which marker amongst CD24, CD44 and EpCAM is a *bona fide* CSC marker for NPC cells, extensive *in vivo* serial transplantation beyond three successive passages of marker-selected NPC cells from additional early-passage NPC PDX models is warranted. This is to rule out the possibility of CD24, CD44 or EpCAM being a marker for tumour progenitor cells, which are tumour cells with proliferative and dividing abilities in addition to short-term self-renewal propensity (Seaberg & Van Der Kooy, 2003). Furthermore, it will be interesting to perform gene expression profiling of these marker-selected NPC cells at a single-cell level using more sensitive detection platforms such as RNA-seq in order to verify the degree of transcriptional heterogeneity in the malignancy. The information will provide insights on transcript variation and regulation within and between different clones of NPC cells having differential expression of the markers for more in-depth mechanistic studies.

5.4 Conclusion

This study was carried out to characterize tumourigenic subpopulations of NPC cells isolated with SP and CSC marker approaches. In the first part of the study, SP cells from HK1 cell line displayed more stem-like properties *in vitro* than NSP cells such as an increased formation of holoclones, ability to undergo asymmetric cell division and slower proliferation. The former also expressed higher levels of ALDH activity and showed upregulation of genes from the Hedgehog, Notch, TGFβ and Wnt pathways as compared to NSP cells. Yet, both SP and NSP cells showed similar tumour formation ability and TIC frequency. CD24, CD44, EpCAM and EpCAM/CD44 markers were evaluated for their abilities to identify NPC tumourigenic cells in C666-1 cell line and xeno-B110 PDX in the second part of the study. CD44br, EpCAMbr and EpCAM/CD44dbr cells of C666-1 and xeno-B110 cells identified for faster-growing cells with higher tumourigenic potential and TIC frequency *in vivo* than their respective

dim phenotype. Subsequent work found that the faster-growing tumourigenic CD44br, EpCAMbr and EpCAM/CD44dbr cells of xeno-B110 had higher percentage of S-phase cells and mitotic activity than CD44dim, EpCAMdim and EpCAM/CD44ddim cells, respectively. However, EpCAM/CD44dbr marker was unable to further enrich for faster-growing tumourigenic cells than single CD44br marker. CD24br, CD44br and EpCAMbr markers could also isolate for self-renewing NPC cells with different efficiencies. *KLF4* transcript was downregulated in CD24br, EpCAMbr and EpCAM/CD44dbr cells which had also induced larger tumour growth, indicating an anti-proliferation role for KLF4 in NPC.

Overall, higher degree of and more convincing differential biological properties were obtained from the use of early-passage PDX samples rather than from long established cell lines in this study. Also, cellular stemness properties need to be ascertained functionally and thoroughly, instead of deriving only from phenotypic characterisation. The potential utility of CSC surface markers for disease prognosis in clinical settings is much more warranted than SP technique in terms of ease of application, shorter turnaround time and practicality.

REFERENCES

Azizah Abdul Manan, Nor Saleha Ibrahim Tamin, Noor Hashimah Abdullah, Asmah Zainal Abidin, & Mastulu Wahab. (2016). *Malaysian national cancer registry report 2007-2011*. Kuala Lumpur: National Cancer Institute, Ministry of Health, Malaysia.

Aigner, S., Ramos, C. L., Hafezi-Moghadam, A., Lawrence, M. B., Friederichs, J., Altevogt, P., & Ley, K. (1998). CD24 mediates rolling of breast carcinoma cells on P-selectin. *The FASEB Journal*, *12*(12), 1241–1251.

Akita, H., Nagano, H., Takeda, Y., Eguchi, H., Wada, H., Kobayashi, S., … Doki, Y. (2011). Ep-CAM is a significant prognostic factor in pancreatic cancer patients by suppressing cell activity. *Oncogene*, *30*(31), 3468–3476.

Akunuru, S., Palumbo, J., Zhai, Q. J., & Zheng, Y. (2011). Rac1 targeting suppresses human non-small cell lung adenocarcinoma cancer stem cell activity. *PLoS ONE*, *6*(2), e16951.

Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., & Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(7), 3983–3988.

Alison, M. R., Guppy, N. J., Lim, S. M., & Nicholson, L. J. (2010). Finding cancer stem cells: Are aldehyde dehydrogenases fit for purpose? *Journal of Pathology*, *222*(4), 335–344.

Alison, M. R., Lin, W. R., Lim, S. M., & Nicholson, L. J. (2012). Cancer stem cells: In the line of fire. *Cancer Treatment Reviews*, *38*(6), 589–598.

Arlt, A., Vorndamm, J., Müerköster, S., Yu, H., Schmidt, W. E., Fölsch, U. R., & Schäfer, H. (2002). Autocrine production of interleukin 1β confers constitutive nuclear factor κB activity and chemoresistance in pancreatic carcinoma cell lines. *Cancer Research*, *62*(3), 910–916.

Autengruber, A., Gereke, M., Hansen, G., Hennig, C., & Bruder, D. (2012). Impact of enzymatic tissue disintegration on the level of surface molecule expression and immune cell function. *European Journal of Microbiology & Immunology*, *2*(2), 112–120.

Baccelli, I., & Trumpp, A. (2012). The evolving concept of cancer and metastasis stem cells. *The Journal of Cell Biology*, *198*(3), 281–293.

Baumann, P., Cremers, N., Kroese, F., Orend, G., Chiquet-Ehrismann, R., Uede, T., … Sleeman, J. P. (2005). CD24 expression causes the acquisition of multiple cellular properties associated with tumor growth and metastasis. *Cancer Research*, *65*(23), 10783–10793.

Beier, D., Hau, P., Proescholdt, M., Lohmeier, A., Wischhusen, J., Oefner, P. J., … Beier, C. P. (2007). CD133+ and CD133− glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Research*, *67*(9), 4010–4015.

Bell, A. I., Groves, K., Kelly, G. L., Croom-Carter, D., Hui, E., Chan, A. T. C., & Rickinson, A. B. (2006). Analysis of Epstein-Barr virus latent gene expression in endemic Burkitt's lymphoma and nasopharyngeal carcinoma tumour cells by using quantitative real-time PCR assays. *The Journal of General Virology*, *87*(2006), 2885–2890.

Bertoncello, I., & Williams, B. (2004). Hematopoietic stem cell characterization by Hoechst 33342 and Rhodamine 123 staining. In T. S. Hawley & R. G. Hawley (Eds.), *Methods in molecular biology: Flow cytometry protocols* (2nd ed., pp. 181–200). Totowa, New Jersey: Humana Press.

Biddle, A., Gammon, L., Fazil, B., & Mackenzie, I. C. (2013). CD44 staining of cancer stem-like cells is influenced by down-regulation of CD44 variant isoforms and up-regulation of the standard CD44 isoform in the population of cells that have undergone epithelial-to- mesenchymal transition. *PLoS ONE*, *8*(2), e57314.

Blankenberg, F. G., Teplitz, R. L., Ellis, W., Salamat, M. S., Byung Hee Min, Hall, L., … Alvord, E. C. (1995). The influence of volumetric tumor doubling time, DNA ploidy, and histologic grade on the survival of patients with intracranial astrocytomas. *American Journal of Neuroradiology*, *16*(5), 1001–1017.

Broadley, K. W. R., Hunn, M. K., Farrand, K. J., Price, K. M., Grasso, C., Miller, R. J., … McConnell, M. J. (2011). Side population is not necessary or sufficient for a cancer stem cell phenotype in glioblastoma multiforme. *Stem Cells*, *29*(3), 452–461.

Buczacki, S., Davies, R. J., & Winton, D. J. (2011). Stem cells, quiescence and rectal carcinoma: An unexplored relationship and potential therapeutic target. *British Journal of Cancer*, *105*(9), 1253–1259.

Bunz, F. (2008). *Principles of cancer genetics* (1st ed.). Dordrecht, Netherlands: Springer.

Busson, P., Ganem, G., Flores, P., Mugneret, F., Clausse, B., Caillou, B., … Tursz, T. (1988). Establishment and characterization of three transplantable EBV-containing nasopharyngeal carcinomas. *International Journal of Cancer*, *42*, 599–606.

Calvet, C. Y., André, F. M., & Mir, L. M. (2014). The culture of cancer cell lines as tumorspheres does not systematically result in cancer stem cell enrichment. *PLoS ONE*, *9*(2), e89644.

Cao, L., Zhou, Y., Zhai, B., Liao, J., Xu, W., Zhang, R., … Yin, Z. (2011). Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines. *BMC Gastroenterology*, *11*(1), 71.

Cao, Y., Yang, L., Jiang, W., Wang, X., Liao, W., Tan, G., … Sun, L.-Q. (2014). Therapeutic evaluation of Epstein-Barr virus-encoded latent membrane protein-1 targeted DNAzyme for treating of nasopharyngeal carcinomas. *Molecular Therapy*, *22*(2), 371–377.

Chan, A. T. C. (2010). Nasopharyngeal carcinoma. *Annals of Oncology*, *21*(suppl. 7), 308–312.

Chan, J. K. C., Bray, F., McCarron, P., Foo, P., Lee, A. W. M., Yip, T., … Jia, W. H. (2005). Nasopharyngeal carcinoma. In L. Barnes, J. W. Eveson, P. Reichart, & D. Sidransky (Eds.), *Classification of tumours. Pathology and genetics of head and neck tumours* (pp. 85–97). Lyon: IARC Press.

Chan, K. C., Chan, L. S., Chok, J., Ip, Y., Lo, C., Tak, T., … Mak, N. K. (2015). Therapeutic targeting of CBP/b-catenin signaling reduces cancer stem-like population and synergistically suppresses growth of EBV-positive nasopharyngeal carcinoma cells with cisplatin. *Scientific Reports*, *5*, 9979.

Chan, K. S., Espinosa, I., Chao, M., Wong, D., Ailles, L., Diehn, M., … Weissman, I. L. (2009). Identification, molecular characterization, clinical prognosis, and therapeutic targeting of human bladder tumor-initiating cells. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(33), 14016–14021.

Chan, S. Y. Y., Choy, K. W., Tsao, S. W., Tao, Q., Tang, T., Chung, G. T. Y., & Lo, K. W. (2008). Authentication of nasopharyngeal carcinoma tumor lines. *International Journal of Cancer*, *122*(9), 2169–2171.

Chang, J. C. (2016). Cancer stem cells: Role in tumor growth, recurrence, metastasis, and treatment resistance. *Medicine (Baltimore)*, *95*(S1), 20–25.

Charafe-Jauffret, E., Ginestier, C., Iovino, F., Wicinski, J., Cervera, N., Finetti, P., … Wicha, M. S. (2009). Breast cancer cell lines contain functional cancer stem sells with metastatic capacity and a distinct molecular signature. *Cancer Research*, *69*(4), 1302–1313.

Chen, J., Wang, J., Chen, D., Yang, J., Yang, C., Zhang, Y., … Dou, J. (2013). Evaluation of characteristics of CD44+CD117+ ovarian cancer stem cells in three dimensional basement membrane extract scaffold versus two dimensional monocultures. *BMC Cell Biology*, *14*(1), 7.

Chen, R., Nishimura, M. C., Bumbaca, S. M., Kharbanda, S., Forrest, W. F., Kasman, I. M., … Phillips, H. S. (2010). A hierarchy of self-renewing tumor-initiating cell types in glioblastoma. *Cancer Cell*, *17*(4), 362–375.

Chen, S., Xu, Y., Chen, Y., Li, X., Mou, W., Wang, L., … Li, N. (2012). SOX2 gene regulates the transcriptional network of oncogenes and affects tumorigenesis of human lung cancer cells. *PLoS ONE*, *7*(5), e36326.

Cheng, Y., Kwok, A., Cheung, L., Mun, J., Ko, Y., Phoon, Y. P., … Lung, M. L. (2013). Physiological β -catenin signaling controls self- renewal networks and generation of stem-like cells from nasopharyngeal carcinoma. *BMC Cell Biology*, *14*, 44.

Cheung, S. T., Huang, D. P., Hui, A. B. Y., Lo, K. W., Ko, C. W., Tsang, Y. S., … Lee, J. C. K. (1999). Nasopharyngeal carcinoma cell line (C666-1) consistently harbouring Epstein-Barr virus. *International Journal of Cancer*, *83*(1), 121–126.

Chia, W. K., Teo, M., Wang, W. W., Lee, B., Ang, S. F., Tai, W. M., … Toh, H. C. (2014). Adoptive T-cell transfer and chemotherapy in the first-line treatment of metastatic and/or locally recurrent nasopharyngeal carcinoma. *Molecular Therapy*, *22*(1), 132–139.

Chua, M. L. K., Wee, J. T. S., Hui, E. P., & Chan, A. T. C. (2016). Nasopharyngeal carcinoma. *The Lancet*, *387*(10022), 1012–1024.

Clarke, M. F., Dick, J. E., Dirks, P. B., Eaves, C. J., Jamieson, C. H. M., Jones, D. L., … Wahl, G. M. (2006). Cancer stem cells - Perspectives on current status and future directions: AACR workshop on cancer stem cells. *Cancer Research*, *66*(19), 9339–9344.

Clevers, H. (2011). The cancer stem cell: Premises, promises and challenges. *Nature Medicine*, *17*(3), 313–319.

Colotta, F., Allavena, P., Sica, A., Garlanda, C., & Mantovani, A. (2009). Cancer-related inflammation, the seventh hallmark of cancer: Links to genetic instability. *Carcinogenesis*, *30*(7), 1073–1081.

Cree, I. A., Glaysher, S., & Harvey, A. L. (2010). Efficacy of anti-cancer agents in cell lines versus human primary tumour tissue. *Current Opinion in Pharmacology*, *10*(4), 375–379.

Dalerba, P., Dylla, S. J., Park, I. K., Liu, R., Wang, X., Cho, R. W., … Clarke, M. F. (2007). Phenotypic characterization of human colorectal cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(24), 10158–10163.

Dashzeveg, N. K., Taftaf, R., Ramos, E. K., Torre-Healy, L., Chumakova, A., Silver, D. J., … Liu, H. (2017). New advances and challenges of targeting cancer stem cells. *Cancer Research*, *77*(19), 5222–5227.

Dean, M. (2009). ABC transporters, drug resistance, and cancer stem cells. *Journal of Mammary Gland Biology and Neoplasia*, *14*, 3–9.

Dean, M., Hamon, Y., & Chimini, G. (2001). The human ATP-binding cassette (ABC) transporter superfamily. *Journal of Lipid Research*, *42*, 1007–1017.

De Sousa E Melo, F., Vermeulen, L., Fessler, E., & Medema, J. P. (2013). Cancer heterogeneity - A multifaceted view. *EMBO Reports*, *14*(8), 686–695.

Dembinski, J. L., & Krauss, S. (2009). Characterization and functional analysis of a slow cycling stem cell-like subpopulation in pancreas adenocarcinoma. *Clinical & Experimental Metastasis*, *26*(7), 611–623.

Deng, C. C., Liang, Y., Wu, M. S., Feng, F. T., Hu, W. R., Chen, L. Z., … Zeng, Y. X. (2013). Nigericin selectively targets cancer stem cells in nasopharyngeal carcinoma. *The International Journal of Biochemistry & Cell Biology*, *45*(9), 1997–2006.

Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., … Narod, S. A. (2007). Triple-negative breast cancer: Clinical features and patterns of recurrence. *Clinical Cancer Research*, *13*(15), 4429–4434.

Devi, B. C. R., Pisani, P., Tang, T. S., & Parkin, D. M. (2004). High incidence of nasopharyngeal carcinoma in native people of Sarawak, Borneo Island. *Cancer Epidemiology Biomarkers and Prevention*, *13*(3), 482–486.

Diestra, J. E., Scheffer, G. L., Català, I., Maliepaard, M., Schellens, J. H. M., Scheper, R. J., … Izquierdo, M. A. (2002). Frequent expression of the multi-drug resistance-associated protein BCRP/MXR/ABCP/ABCG2 in human tumours detected by the BXP-21 monoclonal antibody in paraffin-embedded material. *Journal of Pathology*, *198*(2), 213–219.

Ding, X. W., Wu, J. H., & Jiang, C. P. (2010). ABCG2: A potential marker of stem cells and novel target in stem cell and cancer therapy. *Life Sciences*, *86*(17–18), 631–637.

Dodbiba, L., Teichman, J., Fleet, A., Thai, H., Starmans, M. H. W., Navab, R., … Liu, G. (2015). Appropriateness of using patient-derived xenograft models for pharmacologic evaluation of novel therapies for esophageal/gastro-esophageal junction cancers. *PLoS ONE*, *10*(3), e0121872.

Dong, H. J., Jang, G. B., Lee, H. Y., Park, S. R., Kim, J. Y., Nam, J. S., & Hong, I. S. (2016). The Wnt/β-catenin signaling/Id2 cascade mediates the effects of hypoxia on the hierarchy of colorectal-cancer stem cells. *Scientific Reports*, *6*, 22966.

Duan, Z., Brakora, K. A., & Seiden, M. V. (2004). Inhibition of ABCB1 (MDR1) and ABCB4 (MDR3) expression by small interfering RNA and reversal of paclitaxel resistance in human ovarian cancer cells. *Molecular Cancer Therapeutics*, *3*(7), 833–838.

ElShamy, W. M., & Duhé, R. J. (2013). Overview: Cellular plasticity, cancer stem cells and metastasis. *Cancer Letters*, *341*(1), 2–8.

Engelmann, K., Shen, H., & Finn, O. J. (2008). MCF7 side population cells with characteristics of cancer stem/progenitor cells express the tumor antigen MUC1. *Cancer Research*, *68*(7), 2419–2426.

Facompre, N. D., Harmeyer, K. M., Sole, X., Kabraji, S., Belden, Z., Sahu, V., … Basu, D. (2016). JARID1B enables transit between distinct states of the stem-like cell population in oral cancers. *Cancer Research*, *76*(18), 5538–5549.

Felthaus, O., Ettl, T., Gosau, M., Driemel, O., Brockhoff, G., Reck, A., … Morsczeck, C. (2011). Cancer stem cell-like cells from a single cell of oral squamous carcinoma cell lines. *Biochemical and Biophysical Research Communications*, *407*(1), 28–33.

Feng, B. J. (2013). Descriptive, environmental and genetic epidemiology of nasopharyngeal carcinoma. In P. Busson (Ed.), *Nasopharyngeal carcinoma: Keys for translational medicine and biology* (pp. 23–41). Austin: Landes Bioscience.

Feng, X., Li, C., Liu, W., Chen, H., Zhou, W., Wang, L., … Ren, C. (2013). DLC-1, a candidate tumor suppressor gene, inhibits the proliferation, migration and tumorigenicity of human nasopharyngeal carcinoma cells. *International Journal of Oncology*, *42*, 1973–1984.

Fernando, A., Glaysher, S., Conroy, M., Pekalski, M., Smith, J., Knight, L. a, … Cree, I. a. (2006). Effect of culture conditions on the chemosensitivity of ovarian cancer cell lines. *Anti-Cancer Drugs*, *17*(8), 913–919.

Fillmore, C. M., & Kuperwasser, C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Research*, *10*(2), R25.

Forman, D., Bray, F., Brewster, D., Gombe Mbalawa, C., Kohler, B., Piñeros, M., … Ferlay, J. (Eds.). (2014). *Cancer incidence in five continents* (No. 164, Vol. X). Lyon: International Agency for Research on Cancer.

Frandsen, T. L., Holst-Hansen, C., & Brünner, N. (2001). Tumorigenicity assays. *eLS*, *10*, e0002608.

Fried, J., Doblin, J., Takamoto, S., Perez, a, Hansen, H., & Clarkson, B. (1982). Effects of Hoechst 33342 on survival and growth of two tumor cell lines and on hematopoietically normal bone marrow cells. *Cytometry*, *3*(1), 42–47.

Friedrich, R. E., Bartel-Friedrich, S., Lobeck, H., Niedobitek, G., & Arps, H. (2003). Epstein-Barr virus DNA and epithelial markers in nasopharyngeal carcinoma. *Medical Microbiology and Immunology*, *192*(3), 141–144.

Fukaya, R., Ohta, S., Yamaguchi, M., Fujii, H., Kawakami, Y., Kawase, T., & Toda, M. (2010). Isolation of cancer stem-like cells from a side population of a human glioblastoma cell line, SK-MG-1. *Cancer Letters*, *291*(2), 150–157.

Gabriel, M. T., Calleja, L. R., Chalopin, A., Ory, B., & Heymann, D. (2016). Circulating tumor cells: A review of non-EpCAM-based approaches for cell enrichment and isolation. *Clinical Chemistry*, *62*(4), 571–581.

Ghuwalewala, S., Ghatak, D., Das, P., Dey, S., Sarkar, S., Alam, N., … Roychoudhury, S. (2016). CD44highCD24low molecular signature determines the cancer stem cell

and EMT phenotype in oral squamous cell carcinoma. *Stem Cell Research*, *16*(2), 405–417.

Gires, O., Klein, C., & Baeuerle, P. (2009). On the abundance of EpCAM on cancer stem cells. *Nature Reviews Cancer*, *9*(2), 143–143.

Glaser, R., Zhang, H. Y., Yao, K. T., Zhu, H. C., Wang, F. X., Li, G. Y., … Li, Y. P. (1989). Two epithelial tumor cell lines (HNE-1 and HONE-1) latently infected with Epstein-Barr virus that were derived from nasopharyngeal carcinomas. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 9524–9528.

Golebiewska, A., Brons, N. H. C., Bjerkvig, R., & Niclou, S. P. (2011). Critical appraisal of the side population assay in stem cell and cancer stem cell research. *Cell Stem Cell*, *8*(2), 136–147.

Goodell, M. A., Brose, K., Paradis, G., Conner, A. S., & Mulligan, R. C. (1996). Isolation and functional properties of murine hematopoietic stem cells that are replicating *in vivo*. *Journal of Experimental Medicine*, *183*, 1797–1806.

Gourzones, C., Ferrand, F. R., Verillaud, B., & Busson, P. (2013). Biological tools for NPC population screening and disease monitoring. In P. Busson (Ed.), *Nasopharyngeal carcinoma: Keys for translational medicine and biology* (pp. 101–117). Austin: Landes Bioscience.

Greve, B., Kelsch, R., Spaniol, K., Eich, H. T., & Götte, M. (2012). Flow cytometry in cancer stem cell analysis and separation. *Cytometry Part A*, *81 A*(4), 284–293.

Gullo, C., Wong, K. L., & Gerrard, T. (2008). Association of Epstein-Barr virus with nasopharyngeal carcinoma and current status of development of cancer-derived cell lines. *Annals of the Academy of Medicine Singapore*, *37*(9), 769–777.

Guo, S., Qian, W., Cai, J., Zhang, L., Wery, J.-P., & Li, Q. (2016). Molecular pathology of patient tumors, patient derived xenografts and cancer cell lines. *Cancer Research*, *76*(16), 4619–4626.

Han, J. S., & Crowe, D. L. (2009). Tumor initiating cancer stem cells from human breast cancer cell lines. *International Journal of Oncology*, *34*, 1449–1453.

Han, M. E., Jeon, T. Y., Hwang, S. H., Lee, Y. S., Kim, H. J., Shim, H. H., … Oh, S. O. (2011). Cancer spheres from gastric cancer patients provide an ideal model system for cancer stem cell research. *Cellular and Molecular Life Sciences*, *68*(21), 3589–

3605.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, *144*, 646–673.

Harper, L. J., Piper, K., Common, J., Fortune, F., & Mackenzie, I. C. (2007). Stem cell patterns in cell lines derived from head and neck squamous cell carcinoma. *Journal of Oral Pathology and Medicine*, *36*(10), 594–603.

Hiraga, T., Ito, S., & Nakamura, H. (2016). EpCAM expression in breast cancer cells is associated with enhanced bone metastasis formation. *International Journal of Cancer*, *138*(7), 1698–1708.

Hoe, S. L. L., Tan, L. P., Abdul Aziz, N., Liew, K., Teow, S.-Y., Abdul Razak, F. R., … Khoo, A. S.-B. (2017). CD24, CD44 and EpCAM enrich for tumour-initiating cells in a newly established patient-derived xenograft of nasopharyngeal carcinoma. *Scientific Reports*, *7*(1), 12372.

Hoe, S. L. L., Tan, L. P., Jamal, J., Peh, S. C., Ng, C. C., Zhang, W. C., … Khoo, A. S. B. (2014). Evaluation of stem-like side population cells in a recurrent nasopharyngeal carcinoma cell line. *Cancer Cell International*, *14*(1), 101.

Hofman, V., Ilie, M. I., Long, E., Selva, E., Bonnetaud, C., Molina, T., … Hofman, P. (2011). Detection of circulating tumor cells as a prognostic factor in patients undergoing radical surgery for non-small-cell lung carcinoma: Comparison of the efficacy of the CellSearch Assay^{TM} and the isolation by size of epithelial tumor cell method. *International Journal of Cancer*, *129*(7), 1651–1660.

Hu, Y., & Smyth, G. K. (2009). ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *Journal of Immunological Methods*, *347*(1–2), 70–78.

Huang, B., Huang, Y. J., Yao, Z. J., Chen, X., Guo, S. J., Mao, X. P., … Qiu, S. P. (2013). Cancer stem cell-like side population cells in clear cell renal cell carcinoma cell line 769P. *PLoS ONE*, *8*(7), e68293.

Huang, D. P., Ho, J. H., Poon, Y. F., Chew, E. C., Saw, D., Lui, M., … Lau, W. H. (1980). Establishment of a cell line (NPC/HK1) from a differentiated squamous carcinoma of the nasopharynx. *International Journal of Cancer*, *26*(2), 127–132.

Hui, A. B., Cheung, S. T., Fong, Y., Lo, K. W., & Huang, D. P. (1998). Characterization of a new EBV-associated nasopharyngeal carcinoma cell line.

Cancer Genetics and Cytogenetics, *101*(2), 83–88.

Imrich, S., Hachmeister, M., & Gires, O. (2012). EpCAM and its potential role in tumor-initiating cells. *Cell Adhesion and Migration*, *6*(1), 30–38.

Ishiguro, T., Sato, A., Ohata, H., Ikarashi, Y., Takahashi, R., Ochiya, T., … Okamoto, K. (2016). Establishment and characterization of an in vitro model of ovarian cancer stem-like cells with an enhanced proliferative capacity. *Cancer Research*, *76*(1), 150–160.

Ishizawa, K., Rasheed, Z. A., Karisch, R., Wang, Q., Kowalski, J., Susky, E., … Matsui, W. (2010). Tumor-initiating cells are rare in many human tumors. *Cell Stem Cell*, *7*(3), 279–282.

Jakubikova, J., Adamia, S., Kost-Alimova, M., Klippel, S., Cervi, D., Daley, J. F., … Mitsiades, C. S. (2011). Lenalidomide targets clonogenic side population in multiple myeloma: Pathophysiologic and clinical implications. *Blood*, *117*(17), 4409–4419.

Jamal-Hanjani, M., Quezada, S. A., Larkin, J., & Swanton, C. (2015). Translational implications of tumor heterogeneity. *Clinical Cancer Research*, *21*(6), 1258–1266.

Janisiewicz, A. M., Shin, J. H., Murillo-Sauca, O., Kwok, S., Le, Q.-T. T., Kong, C., … Sunwoo, J. B. (2012). CD44(+) cells have cancer stem cell-like properties in nasopharyngeal carcinoma. *International Forum of Allergy & Rhinology*, *2*(6), 465–470.

Jia, W. H., Luo, X. Y., Feng, B. J., Ruan, H. L., Bei, J. X., Liu, W. S., … Zeng, Y. X. (2010). Traditional Cantonese diet and nasopharyngeal carcinoma risk: A large-scale case-control study in Guangdong, China. *BMC Cancer*, *10*, 446.

Jia, W. H., Feng, B. J., Xu, Z. L., Zhang, X. S., Huang, P., Huang, L. X., … Zeng, Y. X. (2004). Familial risk and clustering of nasopharyngeal carcinoma in Guangdong, China. *Cancer*, *101*(2), 363–369.

Jiang, R., Niu, X., Huang, Y., & Wang, X. (2016). β-catenin is important for cancer stem cell generation and tumorigenic activity in nasopharyngeal carcinoma. *Acta Biochimica et Biophysica Sinica*, *48*(3), 229–237.

Jin, J., Ouyang, Z., & Wang, Z. (2014). Association of fruit and vegetables with the risk of nasopharyngeal cancer: Evidence from a meta-analysis. *Scientific Reports*, *4*(3), 5229.

Jonker, J. W., Freeman, J., Bolscher, E., Musters, S., Alvi, A. J., Titley, I., … Dale, T. C. (2005). Contribution of the ABC transporters Bcrp1 and Mdr1a/1b to the side population phenotype in mammary gland and bone marrow of mice. *Stem Cells*, *23*(8), 1059–1065.

Joo, K. M., Kim, S. Y., Jin, X., Song, S. Y., Kong, D. S., Lee, J. I., … Nam, D. H. (2008). Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas. *Laboratory Investigation*, *88*(8), 808–815.

Julien, S., Merino-Trigo, A., Lacroix, L., Pocard, M., Goeŕé, D., Mariani, P., … Berthet, C. (2012). Characterization of a large panel of patient-derived tumor xenografts representing the clinical heterogeneity of human colorectal cancer. *Clinical Cancer Research*, *18*(19), 5314–5328.

Kathawala, R. J., Gupta, P., Ashby, C. R., & Chen, Z. S. (2015). The modulation of ABC transporter-mediated multidrug resistance in cancer: A review of the past decade. *Drug Resistance Updates*, *18*, 1–17.

Kato, K., Takao, T., Kuboyama, A., Tanaka, Y., Ohgami, T., Yamaguchi, S., … Wake, N. (2010). Endometrial cancer side-population cells show prominent migration and have a potential to differentiate into the mesenchymal cell lineage. *American Journal of Pathology*, *176*(1), 381–392.

Kelly, P. N., Dakic, A., Adams, J. M., Nutt, S. L., & Strasser, A. (2007). Tumor growth need not be driven by rare cancer stem cells. *Science*, *317*(5836), 337.

Khan, M. I., Czarnecka, A. M., Helbrecht, I., Bartnik, E., Lian, F., & Szczylik, C. (2015). Current approaches in identification and isolation of human renal cell carcinoma cancer stem cells. *Stem Cell Research & Therapy*, *6*(1), 178.

Khoo, A. S. B., & Pua, K. C. (2013). Diagnosis and clinical evaluation of nasopharyngeal carcinoma. In P. Busson (Ed.), *Nasopharyngeal carcinoma: Keys for translational medicine and biology* (pp. 1–9). Austin: Landes Bioscience.

Kondo, S., Wakisaka, N., Muramatsu, M., Zen, Y., Endo, K., Murono, S., … Carolina, N. (2011). Epstein-Barr virus latent membrane protein 1 induces cancer stem/progenitor-like cells in nasopharyngeal epithelial cell lines. *Journal of Virology*, *85*(21), 11255–11264.

Kong, Q. L., Hu, L. J., Cao, J. Y., Huang, Y. J., Xu, L. H., Liang, Y., … Zeng, M. S. (2010). Epstein-Barr virus-encoded LMP2A induces an epithelial- mesenchymal transition and increases the number of side population stem-like cancer cells in nasopharyngeal carcinoma. *PLoS Pathogens*, *6*(6), e100094.

Krampitz, G. W., George, B. M., Willingham, S. B., Volkmer, J.-P., Weiskopf, K., Jahchan, N., … Weissman, I. L. (2016). Identification of tumorigenic cells and therapeutic targets in pancreatic neuroendocrine tumors. *Proceedings of the National Academy of Sciences of USA*, *113*(16), 4464–4469.

Kroepil, F., Dulian, A., Vallböhmer, D., Geddert, H., Krieg, A., Vay, C., … Stoecklein, N. H. (2013). High EpCAM expression is linked to proliferation and lauren classification in gastric cancer. *BMC Research Notes*, *6*, 253.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., … Dick, J. E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, *367*(6464), 645–648.

Lau, W. M., Teng, E., Chong, H. S., Lopez, K. A. P., Tay, A. Y. L., Salto-Tellez, M., … Chan, S. L. (2014). CD44v8-10 is a cancer-specific marker for gastric cancer stem cells. *Cancer Research*, *74*(9), 2630–2641.

Leccia, F., Del Vecchio, L., Mariotti, E., Di Noto, R., Morel, A.-P., Puisieux, A., … Ansieau, S. (2014). ABCG2, a novel antigen to sort luminal progenitors of BRCA1- breast cancer cells. *Molecular Cancer*, *13*, 213.

Lee, A. W. M., Ma, B. B. Y., Ng, W. T., & Chan, A. T. C. (2015). Management of nasopharyngeal carcinoma: Current practice and future perspective. *Journal of Clinical Oncology*, *33*(29), 3356–3364.

Lee, A. W. M., Sze, W. M., Au, J. S. K., Leung, S. F., Leung, T. W., Chua, D. T. T., … Lau, W. H. (2005). Treatment results for nasopharyngeal carcinoma in the modern era: The Hong Kong experience. *International Journal of Radiation Oncology Biology Physics*, *61*(4), 1107–1116.

Lee, H. J., Choe, G., Jheon, S., Sung, S.-W., Lee, C.-T., & Chung, J.-H. (2010). CD24, a novel cancer biomarker, predicting disease-free survival of non-small cell lung carcinomas: A retrospective study of prognostic factor analysis from the viewpoint of forthcoming (seventh) new TNM classification. *Journal of Thoracic Oncology*, *5*(5), 649–657.

Lee, T. K. W., Castilho, A., Cheung, V. C. H., Tang, K. H., Ma, S., & Ng, I. O. L. (2011). CD24 + liver tumor-initiating cells drive self-renewal and tumor initiation through STAT3-mediated NANOG regulation. *Cell Stem Cell*, *9*(1), 50–63.

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., … Simeone, D. M. (2007). Identification of pancreatic cancer stem cells. *Cancer Research*, *67*(3), 1030–1037.

Li, H., Chen, X., Calhoun-Davis, T., Claypool, K., & Tang, D. G. (2008). PC3 human prostate carcinoma cell holoclones contain self-renewing tumor-initiating cells. *Cancer Research*, *68*(6), 1820–1825.

Liang, Y., Zhong, Z., Huang, Y., Deng, W., Cao, J., Tsao, G., … Zeng, Y. X. (2010). Stem-like cancer cells are inducible by increasing genomic instability in cancer cells. *Journal of Biological Chemistry*, *285*(7), 4931–4940.

Lim, Y. C., Oh, S. Y., Cha, Y. Y., Kim, S. H., Jin, X., & Kim, H. (2011). Cancer stem cell traits in squamospheres derived from primary head and neck squamous cell carcinomas. *Oral Oncology*, *47*(2), 83–91.

Lin, D. C., Meng, X., Hazawa, M., Nagata, Y., Varela, A. M., Xu, L., … Koeffler, H. P. (2014a). The genomic landscape of nasopharyngeal carcinoma. *Nature Genetics*, *46*(8), 866–871.

Lin, J. C., Wang, W. Y., Chen, K. Y., Wei, Y. H., Liang, W. M., Jan, J. S., & Jiang, R. S. (2004). Quantification of plasma Epstein–Barr virus DNA in patients with advanced nasopharyngeal carcinoma. *New England Journal of Medicine*, *350*(24), 2461–2470.

Lin, Z., Khong, B., Kwok, S., Cao, H., West, R. B., Le, Q. T., & Kong, C. S. (2014b). Human papillomavirus 16 detected in nasopharyngeal carcinomas in white Americans but not in endemic southern Chinese patients. *Head and Neck*, *36*(5), 709–714.

Lin, Z. S., Chu, H. C., Yen, Y. C., Lewis, B. C., & Chen, Y. W. (2012). Krüppel-like factor 4, a tumor suppressor in hepatocellular carcinoma cells reverts epithelial mesenchymal transition by suppressing slug expression. *PLoS ONE*, *7*(8), e43593.

Liu, Y. T., Dai, J. J., Xu, C. H., Lu, Y. K., Fan, Y. Y., Zhang, X. L., … Chen, Y. M. (2012). Greater intake of fruit and vegetables is associated with lower risk of nasopharyngeal carcinoma in Chinese adults: A case-control study. *Cancer Causes and Control*, *23*(4), 589–599.

Liu, Z., Yang, H., Luo, W., Jiang, Q., Mai, C., Chen, Y., … Fang, W. (2013). Loss of cytoplasmic KLF4 expression is correlated with the progression and poor prognosis of nasopharyngeal carcinoma. *Histopathology*, *63*(3), 362–370.

Lo, K. W. (2016). Somatic genetic changes in EBV-associated nasopharyngeal carcinoma. *BMC Proceedings*, *10*(S1), 1.

Lo, K. W., To, K. F., & Huang, D. P. (2004). Focus on nasopharyngeal carcinoma. *Cancer Cell*, *5*(5), 423–428.

Locke, M., Heywood, M., Fawell, S., & Mackenzie, I. C. (2005). Retention of intrinsic stem cell hierarchies in carcinoma-derived cell lines. *Cancer Research*, *65*(19), 8944–8950.

Lu, Z. X., Ye, M., Yan, G. R., Li, Q., Tang, M., Lee, L. M., … Cao, Y. (2005). Effect of EBV LMP1 targeted DNAzymes on cell proliferation and apoptosis. *Cancer Gene Therapy*, *12*(7), 647–654.

Lun, S. W. M., Cheung, S. T., & Lo, K. W. (2014). Cancer stem-like cells in Epstein-Barr virus-associated nasopharyngeal carcinoma. *Chinese Journal of Cancer*, *33*(11), 529–538.

Lun, S. W. M., Cheung, S. T., Cheung, P. F. Y., To, K. F., Woo, J. K. S., Choy, K. W., … Lo, K. W. (2012). CD44+ cancer stem-like cells in EBV-associated nasopharyngeal carcinoma. *PLoS ONE*, *7*(12), e52426.

Luo, Y., Ellis, L. Z., Dallaglio, K., Takeda, M., Robinson, W. A., Robinson, S. E., … Fujita, M. (2012). Side population cells from human melanoma tumors reveal diverse mechanisms for chemoresistance. *Journal of Investigative Dermatology*, *132*(10), 2440–2450.

Ma, L., Zhang, G., Miao, X. B., Deng, X. Bin, Wu, Y., Liu, Y., … Xiao, G. H. (2013). Cancer stem-like cell properties are regulated by EGFR/AKT/b-catenin signaling and preferentially inhibited by gefitinib in nasopharyngeal carcinoma. *FEBS Journal*, *280*(9), 2027–2041.

MacLeod, R. A. F., Dirks, W. G., Matsuo, Y., Kaufmann, M., Milch, H., & Drexler, H. G. (1999). Widespread intraspecies cross-contamination of human tumor cell lines arising at source. *International Journal of Cancer*, *83*, 555–563.

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., … Weinberg, R. A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, *133*(4), 704–715.

Marcato, P., Dean, C. A., Pan, D., Araslanova, R., Gillis, M., Joshi, M., … Lee, P. W. K. (2011). Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis. *Stem Cells*, *29*(1), 32–45.

Marusyk, A., & Polyak, K. (2010). Tumor heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta*, *1805*(1), 105–117.

Massoner, P., Thomm, T., Mack, B., Untergasser, G., Martowicz, A., Bobowski, K., … Puhr, M. (2014). EpCAM is overexpressed in local and metastatic prostate cancer, suppressed by chemotherapy and modulated by MET-associated miRNA-200c/205. *British Journal of Cancer*, *111*(5), 955–964.

McConnell, B. B., & Yang, V. W. (2010). Mammalian Krueppel-like factors in health and diseases. *Physiological Reviews*, *90*(4), 1337–1381.

McEver, R. P., Beckstead, J. H., Moore, K. L., Marshall-Carlson, L., & Bainton, D. F. (1989). GMP-140, a platelet a-granule membrane protein, is also synthesized by vascular endothelial cells and is localized in Weibel-Palade bodies. *Journal of Clinical Investigation*, *84*(1), 92–99.

Meacham, C. E., & Morrison, S. J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature*, *501*(7467), 328–337.

Medema, J. P. (2013). Cancer stem cells: The challenges ahead. *Nature Cell Biology*, *15*(4), 338–44.

Medri, L. (2003). Prognostic relevance of mitotic activity in patients with node-negative breast cancer. *Modern Pathology*, *16*(11), 1067–1075.

Mitsutake, N., Iwao, A., Nagai, K., Namba, H., Ohtsuru, A., Saenko, V., & Yamashita, S. (2007). Characterization of side population in thyroid cancer cell lines: cancer stem-like cells are enriched partly but not exclusively. *Endocrinology*, *148*(4), 1797–1803.

Montanaro, F., Liadaki, K., Schienda, J., Flint, A., Gussoni, E., & Kunkel, L. M. (2004). Demystifying SP cell purification: Viability, yield, and phenotype are defined by isolation parameters. *Experimental Cell Research*, *298*(1), 144–154.

Moore, N., & Lyle, S. (2011). Quiescent, slow-cycling stem cell populations in cancer: A review of the evidence and discussion of significance. *Journal of Oncology*, *2011*, 396076.

Morreale, R. (2003). *Nasopharyngeal carcinoma: Medical illustrations*. Retrieved from https://www.cancer.net/cancer-types/nasopharyngeal-cancer/medical-illustrations

Münz, M., Kieu, C., Mack, B., Schmitt, B., Zeidler, R., & Gires, O. (2004). The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene*, *23*, 5748–5758.

Nakanishi, T., Chumsri, S., Khakpour, N., Brodie, A. H., Leyland-Jones, B., Hamburger, A. W., … Burger, A. M. (2010). Side-population cells in luminal-type breast cancer have tumour-initiating cell properties, and are regulated by HER2 expression and signalling. *British Journal of Cancer*, *102*(5), 815–826.

Nakayama, M., Ogasawara, S., Akiba, J., Ueda, K., Koura, K., Todoroki, K., … Yano, H. (2014). Side population cell fractions from hepatocellular carcinoma cell lines increased with tumor dedifferentiation, but lack characteristic features of cancer stem cells. *Journal of Gastroenterology and Hepatology*, *29*(5), 1092–1101.

Naor, D., Wallach-Dayan, S. B., Zahalka, M. A., & Sionov, R. V. (2009). Involvement of CD44, a molecule with a thousand faces, in cancer Dissemination. *Seminars in Cancer Biology*, *18*, 260–267.

Ni, J., Cozzi, P. J., Duan, W., Shigdar, S., Graham, P. H., John, K. H., & Li, Y. (2012). Role of the EpCAM (CD326) in prostate cancer metastasis and progression. *Cancer and Metastasis Reviews*, *31*(3–4), 779–791.

Norashikin Zakaria, Norazah Mohd Yusoff, Zubaidah Zakaria, Lim, M. N., Puteri J Noor Baharuddin, Kamal Shaik Fakiruddin, & Badrul Yahaya. (2015). Human non-small cell lung cancer expresses putative cancer stem cell markers and exhibits the transcriptomic profile of multipotent cells. *BMC Cancer*, *15*, 84.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, *194*(4260), 23–28.

Oh, S., Jung, H., Sook, Y., Kim, H., & Lim, Y. C. (2013). CD44-negative cells in head and neck squamous carcinoma also have stem-cell like traits. *European Journal of Cancer*, *49*(1), 272–280.

Paebst, F., Piehler, D., Brehm, W., Heller, S., Schroeck, C., Tárnok, A., & Burk, J. (2014). Comparative immunophenotyping of equine multipotent mesenchymal stromal cells: An approach toward a standardized definition. *Cytometry Part A*, *85*(8), 678–687.

Park, N. R., Cha, J. H., Jang, J. W., Bae, S. H., Jang, B. H., Kim, J.-H., … Yoon, S. K. (2016). Synergistic effects of CD44 and TGF-β1 through AKT/GSK-3β/β-catenin signaling during epithelial-mesenchymal transition in liver cancer cells. *Biochemical and Biophysical Research Communications*, *477*(4), 568–574.

Pastrana, E., Silva-Vargas, V., & Doetsch, F. (2011). Eyes wide open: A critical review of sphere-formation as an assay for stem cells. *Cell Stem Cell*, *8*(5), 486–498.

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Zhou, J., Claypool, K., & Tang, D. G. (2005). Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2+ and ABCG2- cancer cells are similarly tumorigenic. *Cancer Research*, *65*(14), 6207–6219.

Patriarca, C., Macchi, R. M., Marschner, A. K., & Mellstedt, H. (2012). Epithelial cell adhesion molecule expression (CD326) in cancer: A short review. *Cancer Treatment Reviews*, *38*(1), 68–75.

Peickert, S., Waurig, J., Dittfeld, C., Dietrich, A., Garbe, Y., Kabus, L., … Kunz-Schughart, L. A. (2012). Rapid re-expression of CD133 protein in colorectal cancer cell lines in vitro and in vivo. *Laboratory Investigation*, *92*(11), 1607–1622.

Platet, N., He, F., Inserm, U., & Michallon, C. H. U. (2007). Fluctuation of the SP/non-SP phenotype in the C6 glioma cell line. *FEBS Letters*, *581*, 1435–1440.

Port, R. J., Pinheiro-Maia, S., Hu, C., Arrand, J. R., Wei, W., Young, L. S., & Dawson, C. W. (2013). Epstein-Barr virus induction of the Hedgehog signalling pathway imposes a stem cell phenotype on human epithelial cells. *Journal of Pathology*, *231*(3), 367–377.

Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., … Ailles, L. E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(3), 973–978.

Pua, K. C., Khoo, A. S., Yap, Y. Y., Subramaniam, S. K., Ong, C. A., Gopala Krishnan, G., & Shahid, H. (2008). Nasopharyngeal carcinoma database. *Medical Journal of Malaysia*, *63*(Suppl C), 59–62.

Qi, W., Zhao, C., Zhao, L., Liu, N., Li, X., Yu, W., & Wei, L. (2014). Sorting and identification of side population cells in the human cervical cancer cell line HeLa. *Cancer Cell International*, *14*(1), 3.

Quan, Y., Yan, Y., Wang, X., Fu, Q., Wang, W., Wu, J., … Wang, Y. (2012). Impact of cell dissociation on identification of breast cancer stem cells. *Cancer Biomarkers*, *12*(3), 125–133.

Quintana, E., Shackleton, M., Sabel, M. S., Fullen, D. R., Johnson, T. M., & Morrison,

S. J. (2008). Efficient tumour formation by single human melanoma cells. *Nature*, *456*(7222), 593–598.

Robey, R. W., Shukla, S., Steadman, K., Robey, R. W., Shukla, S., Steadman, K., … Bates, S. E. (2007). Inhibition of ABCG2-mediated transport by protein kinase inhibitors with a bisindolylmaleimide or indolocarbazole structure. *Molecular Cancer Therapeutics*, *6*, 1877–1885.

Rosenbloom, D. I. S., Elliott, O., Hill, A. L., Henrich, T. J., Siliciano, J. M., & Siliciano, R. F. (2015). Designing and interpreting limiting dilution assays: General principles and applications to the latent reservoir for human immunodeficiency virus-1. *Open Forum Infectious Diseases*, *2*(4), ofv123.

Roudi, R., Madjd, Z., Ebrahimi, M., Najafi, A., Korourian, A., Shariftabrize, A., & Samadikuchaksaraei, A. (2016). Evidence for embryonic stem-like signature and epithelial-mesenchymal transition features in the spheroid cells derived from lung adenocarcinoma. *Tumor Biology*, *37*(9), 11843–11859.

Sagawa, K., Uwa, N., Daimon, T., Sakagami, M., & Tsujimura, T. (2016). Expression of CD44 variant isoforms, CD44v3 and CD44v6, are associated with prognosis in nasopharyngeal carcinoma. *The Journal of Laryngology and Otology*, *130*(9), 843–849.

Sahlberg, S. H., Spiegelberg, D., Glimelius, B., Stenerlow, B., & Nestor, M. (2014). Evaluation of cancer stem cell markers CD133, CD44, CD24: Association with AKT isoforms and radiation resistance in colon cancer cells. *PLoS ONE*, *9*(4), e94621.

Salcido, C. D., Larochelle, A., Taylor, B. J., Dunbar, C. E., & Varticovski, L. (2010). Molecular characterisation of side population cells with cancer stem cell-like characteristics in small-cell lung cancer. *British Journal of Cancer*, *102*(11), 1636–1644.

Samaeekia, R., Adorno-Cruz, V., Bockhorn, J., Chang, Y. F., Huang, S., Prat, A., … Liu, H. (2017). microRNA-206 inhibits stemness and metastasis of breast cancer by targeting MKL1/IL11 pathway. *Clinical Cancer Research*, *23*(4), 1091–1103.

Seaberg, R. M., & Van Der Kooy, D. (2003). Stem and progenitor cells: The premature desertion of rigorous definitions. *Trends in Neurosciences*, *26*(3), 125–131.

Shackleton, M., Quintana, E., Fearon, E. R., & Morrison, S. J. (2009). Heterogeneity in cancer: Cancer stem cells versus clonal evolution. *Cell*, *138*(5), 822–829.

Shen, Y., Wang, C., Chuang, H., Hwang, J. J., Chi, W., Shu, C., … Chen, Y. (2016). CD44 and CD24 coordinate the reprogramming of nasopharyngeal carcinoma cells towards a cancer stem cell phenotype through STAT3 activation. *Oncotarget*, *7*(36), 58351–58366.

Shigeishi, H., Biddle, A., Gammon, L., Emich, H., Rodini, C. O., Fazil, B., … Mackenzie, I. C. (2013). Maintenance of stem cell self-renewal in head and neck cancers requires actions of GSK3b influenced by CD44 and RHAMM . *Stem Cells*, *31*(10), 2073–2083.

Shultz, L. D., Lyons, B. L., Burzenski, L. M., Gott, B., Chen, X., Chaleff, S., … Handgretinger, R. (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *Journal of Immunology*, *174*, 6477–6489.

Singh, R. R., Kunkalla, K., Qu, C., Schlette, E., Neelapu, S. S., Samaniego, F., & Vega, F. (2011). ABCG2 is a direct transcriptional target of Hedgehog signaling and involved in stroma-induced drug tolerance in diffuse large B-cell lymphoma. *Oncogene*, *30*(49), 4874–4886.

Singh, S. K., Clarke, I. D., Terasaki, M., Bonn, V. E., Hawkins, C., Squire, J., & Dirks, P. B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer Research*, *63*(18), 5821–5828.

Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., … Dirks, P. B. (2004). Identification of human brain tumour initiating cells. *Nature*, *432*(7015), 396–401.

Stewart, J. M., Shaw, P. A., Gedye, C., Bernardini, M. Q., Neel, B. G., & Ailles, L. E. (2011). Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(16), 6468–6473.

Strong, M. J., Baddoo, M., Nanbo, A., Xu, M., Puetter, A., & Lin, Z. (2014). Comprehensive RNA-seq analysis reveals contamination of multiple nasopharyngeal carcinoma cell lines with HeLa cell genomes. *Journal of Virology*, *88*(18), 10696–10704.

Su, J., Xu, X. H., Huang, Q., Lu, M. Q., Li, D. J., Xue, F., … Wu, Y. P. (2011). Identification of cancer stem-like CD44+ cells in human nasopharyngeal carcinoma cell line. *Archives of Medical Research*, *42*(1), 15–21.

Sung, J. M., Cho, H. J., Yi, H., Lee, C. H., Kim, H. S., Kim, D. K., … Shin, H. C. (2008). Characterization of a stem cell population in lung cancer A549 cells. *Biochemical and Biophysical Research Communications*, *371*(1), 163–167.

Takebe, N., Miele, L., Harris, P. J., Jeong, W., Bando, H., Kahn, M., … Ivy, S. P. (2015). Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: Clinical update. *Nature Reviews Clinical Oncology*, *12*(8), 445–464.

Tan, G. W., & Tan, L. P. (2017). High-throughput RT-qPCR for the analysis of circulating microRNAs. In T. Dalmay (Ed.), *MicroRNA detection and target identification: Methods and protocols* (pp. 7–20). New York, USA: Humana Press.

Tan, Y., Zhao, M., Xiang, B., Chang, C., & Lu, Q. (2016). CD24: From a hematopoietic differentiation antigen to a genetic risk factor for multiple autoimmune diseases. *Clinical Reviews in Allergy and Immunology*, *50*(1), 70–83.

Tang, D. G. (2012). Understanding cancer stem cell heterogeneity and plasticity. *Cell Research*, *22*(3), 457–472.

Tirino, V., Desiderio, V., D'Aquino, R., De Francesco, F., Pirozzi, G., Galderisi, U., … Papaccio, G. (2008). Detection and characterization of CD133+ cancer stem cells in human solid tumours. *PLoS ONE*, *3*(10), e3469.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, *65*(2), 87–108.

Trzpis, M., McLaughlin, P. M. J., de Leij, L. M. F. H., & Harmsen, M. C. (2007). Epithelial cell adhesion molecule: More than a carcinoma marker and adhesion molecule. *The American Journal of Pathology*, *171*(2), 386–395.

Tsang, C. M., Cheung, Y. C., Lui, V. W. Y., Yip, Y. L., Zhang, G., Lin, V. W., … Tsao, S. W. (2013). Berberine suppresses tumorigenicity and growth of nasopharyngeal carcinoma cells by inhibiting STAT3 activation induced by tumor associated fibroblasts. *BMC Cancer*, *13*(1), 619.

Tsao, S. W., Yip, Y. L., Tsang, C. M., Pang, P. S., Lau, V. M. Y., Zhang, G., & Lo, K. W. (2014). Etiological factors of nasopharyngeal carcinoma. *Oral Oncology*, *50*(5), 330–338.

Tysnes, B. B. (2010). Tumor-initiating and -propagating cells: cells that we would like to identify and control. *Neoplasia*, *12*(7), 506–515.

Valent, P., Bonnet, D., De Maria, R., Lapidot, T., Copland, M., Melo, J. V., … Eaves, C. (2012). Cancer stem cell definitions and terminology: The devil is in the details. *Nature Reviews Cancer*, *12*(11), 767–775.

van Staveren, W. C. G., Solis, D. Y., Hebrant, A., Detours, V., Dumont, J. E., Maenhaut, C., … Maenhaut, C. (2009). Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells? *Biochimica et Biophysica Acta*, *1795*(2), 92–103.

Vicente-Dueñas, C., Gutiérrez de Diego, J., Rodríguez, F. D., Jiménez, R., & Cobaleda, C. (2009). The role of cellular plasticity in cancer development. *Current Medicinal Chemistry*, *16*(28), 3676–3685.

Visvader, J. E., & Lindeman, G. J. (2008). Cancer stem cells in solid tumours: Accumulating evidence and unresolved questions. *Nature Reviews Cancer*, *8*(10), 755–768.

Visvader, J. E., & Lindeman, G. J. (2012). Cancer stem cells: Current status and evolving complexities. *Cell Stem Cell*, *10*(6), 717–728.

Vlashi, E., & Pajonk, F. (2015). Cancer stem cells, cancer cell plasticity and radiation therapy. *Seminars in Cancer Biology*, *31*, 28–35.

Wang, A., Chen, L., Li, C., & Zhu, Y. (2015). Heterogeneity in cancer stem cells. *Cancer Letters*, *357*, 63–68.

Wang, C., Lin, X. L., Fan, Y. Y., Liu, Y. T., Zhang, X. L., Lu, Y. K., … Chen, Y. M. (2016a). Diet quality scores and risk of nasopharyngeal carcinoma in Chinese adults: A case-control study. *Nutrients*, *8*(3), 1–11.

Wang, J., Guo, L. P., Chen, L. Z., Zeng, Y. X., & Lu, S. H. (2007). Identification of cancer stem cell-like side population cells in human nasopharyngeal carcinoma cell line. *Cancer Research*, *67*(8), 3716–3724.

Wang, J., Xin, B., Wang, H., He, X., Wei, W., Zhang, T., & Shen, X. (2016b). Gastrin regulates ABCG2 to promote the migration , invasion and side populations in pancreatic cancer cells via activation of NF- κ B signaling. *Experimental Cell Research*, *346*(1), 74–84.

Wang, L., Park, P., Zhang, H., La Marca, F., & Lin, C. Y. (2011). Prospective identification of tumorigenic osteosarcoma cancer stem cells in OS99-1 cells based on high aldehyde dehydrogenase activity. *International Journal of Cancer*, *128*(2),

294–303.

Wang, S. J., Wong, G., De Heer, A. M., Xia, W., & Bourguignon, L. Y. W. (2009). CD44 variant isoforms in head and neck squamous cell carcinoma progression. *Laryngoscope*, *119*(8), 1518–1530.

Wang, W. Y., Twu, C. W., Chen, H. H., Jiang, R. S., Wu, C. T., Liang, K. L., … Lin, J. C. (2013a). Long-term survival analysis of nasopharyngeal carcinoma by plasma Epstein-Barr virus DNA levels. *Cancer*, *119*(5), 963–970.

Wang, W. J., Wu, S. P., Liu, J. Bin, Shi, Y. S., Huang, X., Zhang, Q. B., & Yao, K. T. (2013b). MYC regulation of CHK1 and CHK2 promotes radioresistance in a stem cell-like population of nasopharyngeal carcinoma cells. *Cancer Research*, *73*(3), 1219–1231.

Wang, Y., Zhang, G. H., & Li, S. L. (2015). Isolation and phenotypic characterization of cancer stem like cells from nasopharyngeal carcinoma. *Drug Research*, *65*(6), 323–326.

Wasil, L. R., Tomaszewski, M. J., Hoji, A., & Rowe, D. T. (2013). The effect of Epstein-Barr virus latent membrane protein 2 expression on the kinetics of early B cell infection. *PLoS ONE*, *8*(1), e54010.

Weaver, V. M., Lelièvre, S., Lakins, J. N., Chrenek, M. A., Jones, J. C. R., Giancotti, F., … Bissell, M. J. (2002). b4 integrin-dependent formation of polarized three-dimensional architecture confers resistance to apoptosis in normal and malignant mammary epithelium. *Cancer Cell*, *2*(3), 205–216.

Wei, W. I., & Sham, J. S. (2005). Nasopharyngeal carcinoma. *Lancet*, *365*(9476), 2041–2054.

Weiswald, L. B., Bellet, D., & Dangles-Marie, V. (2015). Spherical cancer models in tumor biology. *Neoplasia*, *17*(1), 1–15.

Welch, D. R. (2016). Tumor heterogeneity-A "contemporary concept" founded on historical insights and predictions. *Cancer Research*, *76*(1), 4–6.

Wend, P., Holland, J. D., Ziebold, U., & Birchmeier, W. (2010). Wnt signaling in stem and cancer stem cells. *Seminars in Cell and Developmental Biology*, *21*(8), 855–863.

Williams, S. A., Anderson, W. C., Santaguida, M. T., & Dylla, S. J. (2013). Patient-derived xenografts, the cancer stem cell paradigm, and cancer pathobiology in the 21st century. *Laboratory Investigation*, *93*(9), 970–982.

Wu, A., Luo, W., Zhang, Q., Yang, Z., Zhang, G., Li, S., & Yao, K. (2013a). Aldehyde dehydrogenase 1, a functional marker for identifying cancer stem cells in human nasopharyngeal carcinoma. *Cancer Letters*, *330*(2), 181–189.

Wu, C., & Alman, B. A. (2008). Side population cells in human cancers. *Cancer Letters*, *268*(1), 1–9.

Wu, X. Q., Huang, C., He, X., Tian, Y. Y., Zhou, D. X., He, Y., … Li, J. (2013b). Feedback regulation of telomerase reverse transcriptase: new insight into the evolving field of telomerase in cancer. *Cellular Signalling*, *25*(12), 2462–2468.

Xu, C., Chen, Y. P., & Ma, J. (2016a). Clinical trials in nasopharyngeal carcinoma—Past, present and future. *Chinese Clinical Oncology*, *5*(2), 20.

Xu, S., Wen, Z., Jiang, Q., Zhu, L., Feng, S., Zhao, Y., … Zhu, Y. (2015). CD58, a novel surface marker, promotes self-renewal of tumor-initiating cells in colorectal cancer. *Oncogene*, *34*(12), 1520–1531.

Xu, X., Liu, Y., Su, J., Li, D., Hu, J., Huang, Q., … Sun, L. (2016b). Downregulation of Bmi-1 is associated with suppressed tumorigenesis and induced apoptosis in CD44+ nasopharyngeal carcinoma cancer stem-like cells. *Oncology Reports*, *35*, 923–931.

Yamamoto, Y., & Gaynor, R. B. (2001). Therapeutic potential of inhibition of the NF-κB pathway in the treatment of inflammation and cancer. *The Journal of Clinical Investigation*, *107*(2), 135–142.

Yamashita, T. T., Honda, M., Nakamoto, Y., Baba, M., Nio, K., Hara, Y., … Kaneko, S. (2013). Discrete nature of EpCAM+ and CD90+ cancer stem cells in human hepatocellular carcinoma. *Hepatology*, *57*(4), 1484–1497.

Yan, Y., Li, Z., Kong, X., Jia, Z., Zuo, X., Gagea, M., … Xie, K. (2016). KLF4-mediated suppression of CD44 signaling negatively impacts pancreatic cancer stemness and metastasis. *Cancer Research*, *76*(8), 2419–2431.

Yan, Y., Zuo, X., & Wei, D. (2015). Concise review: Emerging role of CD44 in cancer stem cells: A promising biomarker and therapeutic target. *Stem Cells Translational Medicine*, *4*(9), 1033–1043.

Yang, C. F., Peng, L. X., Huang, T. J., Yang, G. Da, Chu, Q. Q., Liang, Y. Y., … Huang, B. J. (2014a). Cancer stem-like cell characteristics induced by EB virus-encoded LMP1 contribute to radioresistance in nasopharyngeal carcinoma by suppressing the p53-mediated apoptosis pathway. *Cancer Letters*, *344*(2), 260–271.

Yang, C. H., Wang, H. L., Lin, Y. S., Kumar, K. P. S., Lin, H. C., Chang, C. J., … Lai, H. C. (2014b). Identification of CD24 as a cancer stem cell marker in human nasopharyngeal carcinoma. *PLoS ONE*, *9*(6), e99412.

Yang, C., Peng, J., Jiang, W., Zhang, Y., Chen, X., Wu, X., … Jin, K. (2013). MTOR activation in immature cells of primary nasopharyngeal carcinoma and anti-tumor effect of rapamycin in vitro and in vivo. *Cancer Letters*, *341*(2), 186–194.

Yang, C., Zhang, Y., Zhang, Y., Zhang, Z., Peng, J., Li, Z., … Liao, Z. (2015). Downregulation of cancer stem cell properties via mTOR signaling pathway inhibition by rapamycin in nasopharyngeal carcinoma. *International Journal of Oncology*, *47*(3), 909–917.

Yi, X. J., Zhao, Y. H., Qiao, L. X., Jin, C. L., Tian, J., & Li, Q. S. (2015). Aberrant Wnt/β-catenin signaling and elevated expression of stem cell proteins are associated with osteosarcoma side population cells of high tumorigenicity. *Molecular Medicine Reports*, *12*(4), 5042–5048.

Yoon, C., Park, D. J., & Schmidt, B. (2014). CD44 expression denotes a subpopulation of gastric cancer cells in which Hedgehog signaling promotes chemotherapy resistance. *Clinical Cancer Research*, *20*(15), 3974–3988.

Yori, J. L., Seachrist, D. D., Johnson, E., Lozada, K. L., Abdul-Karim, F. W., Chodosh, L. A., … Keri, R. A. (2011). Krüppel-like factor 4 inhibits tumorigenic progression and metastasis in a mouse model of breast cancer. *Neoplasia*, *13*(7), 601–610.

Yu, C., Yao, Z., Dai, J., Zhang, H., Escara-Wilke, J., Zhang, X., & Keller, E. T. (2011). ALDH activity indicates increased tumorigenic cells, but not cancer stem cells, in prostate cancer cell lines. *In Vivo*, *25*(1), 69–76.

Yu, D., Shin, H., Lee, Y. S., & Lee, Y. C. (2014). miR-106b modulates cancer stem cell characteristics through TGF- b/Smad signaling in CD44-positive gastric cancer cells. *Laboratory Investigation*, *94*(12), 1370–1381.

Yu, F., & Loh, K. S. (2014). Cancer stem cells in nasopharyngeal carcinoma : Current evidence. *Journal of Nasopharyngeal Carcinoma*, *1*(6), e6.

Zainal Ariffin Omar, & Nor Saleha Ibrahim Tamin. (2011). *National cancer registry report 2007*. Kuala Lumpur: Ministry of Health, Malaysia.

Zhang, G., Wang, Z., Luo, W., Jiao, H., Wu, J., & Jiang, C. (2013a). Expression of potential cancer stem cell marker ABCG2 is associated with malignant behaviors of hepatocellular carcinoma. *Gastroenterology Research and Practice, 2013*, 782581.

Zhang, H., Liu, W., Feng, X., Wang, L., Jiang, X., Liu, D., … Ren, C. (2012a). Identification of ABCG2(+) cells in nasopharyngeal carcinoma cells. *Oncology Reports*, *27*(4), 1177–1187.

Zhang, J., Shu, C., Song, Y., Li, Q., Huang, J., & Ma, X. (2016). Epstein-Barr virus DNA level as a novel prognostic factor in nasopharyngeal carcinoma: A meta-analysis. *Medicine*, *95*(40), e5130.

Zhang, L., Chen, Q. Y., Liu, H., Tang, L. Q., & Mai, H. Q. (2013b). Emerging treatment options for nasopharyngeal carcinoma. *Drug Design, Development and Therapy*, *7*, 37–52.

Zhang, W. C., Ng, S. C., Yang, H., Rai, A., Umashankar, S., Ma, S., … Swarup, S. (2012). Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell*, *148*(1–2), 259–272.

Zheng, D., Liao, S., Zhu, G., Luo, G., Xiao, S., He, J., … Zhou, Y. (2016). CD38 is a putative functional marker for side population cells in human nasopharyngeal carcinoma cell lines. *Molecular Carcinogenesis*, *55*(3), 300–311.

Zheng, S., & Franzmann, E. J. (2013). Comments on "CD44-negative cells in head and neck squamous carcinoma also have stem-cell like traits", Se-Yeong Oh et al., European Journal of Cancer, published online 6 July 2012. *European Journal of Cancer*, *49*(15), 3380–3381.

Zhou, Q., Facciponte, J., Jin, M., Shen, Q., & Lin, Q. (2014). Humanized NOD-SCID IL2rg –/– mice as a preclinical model for cancer research and its potential use for individualized cancer therapies. *Cancer Letters*, *344*(1), 13–19.

Zhou, S., Schuetz, J. D., Bunting, K. D., Colapietro, a M., Sampath, J., Morris, J. J., … Sorrentino, B. P. (2001). The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nature Medicine*, *7*(9), 1028–1034.

Zhuang, H. W., Mo, T. T., Hou, W. J., Xiong, G. X., Zhu, X. L., Fu, Q. L., & Wen, W. P. (2013). Biological characteristics of CD133(+) cells in nasopharyngeal carcinoma. *Oncology Reports*, *30*(1), 57–63.

Zöller, M. (2011). CD44: Can a cancer-initiating cell profit from an abundantly expressed molecule? *Nature Reviews Cancer*, *11*(4), 254–267.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Publications

i. Hoe, S. L. L., Tan, L. P., Jamal, J., Peh, S. C., Ng, C. C., Zhang, W. C., … Khoo, A. S. B. (2014). Evaluation of stem-like side population cells in a recurrent nasopharyngeal carcinoma cell line. *Cancer Cell International*, *14*(1), 101.

ii. Hoe, S. L. L., Tan, L. P., Abdul Aziz, N., Liew, K., Teow, S. Y., Abdul Razak, F. R., … Khoo, A. S. B. (2017). CD24, CD44 and EpCAM enrich for tumour-initiating cells in a newly established patient-derived xenograft of nasopharyngeal carcinoma. *Scientific Reports*, *7*(1), 12372.

Presentations

i. Hoe, S. L. L. (2015). Functional characterizations of CD24, CD44 and EpCAM cells in NPC. Paper presented at the 4th NPC Research Day, University of Malaya, Kuala Lumpur.

ii. Hoe, S. L. L. (2016). Characterization of tumour-initiating cells in nasopharyngeal carcinoma. Paper presented at the National Colloquium on Stem Cell Research, Universiti Sains Malaysia, Kota Bharu.

iii. Hoe, S. L. L. (2017). Biological properties of tumour subpopulations in nasopharyngeal carcinoma (2017). Paper presented at the 7th Regional Conference on Molecular Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur.

Hoe *et al. Cancer Cell International* 2014, **14**:101
http://www.cancerci.com/content/14/1/101

CANCER CELL INTERNATIONAL

PRIMARY RESEARCH Open Access

Evaluation of stem-like side population cells in a recurrent nasopharyngeal carcinoma cell line

Susan Ling Ling Hoe[1,2], Lu Ping Tan[1], Juliana Jamal[1,5], Suat Cheng Peh[3], Ching Ching Ng[2], Wen Cai Zhang[4], Munirah Ahmad[1] and Alan Soo Beng Khoo[1*]

Abstract

Background: Side population (SP) assay identifies cells with dye/drug extrusion ability, a characteristic of stem cells. Here, we determined if SP cells exist in a verified cell line originating from recurrent nasopharyngeal carcinoma (NPC) and a xenograft established from recurrent metastatic NPC. These cells were evaluated for stem-like properties via functional assays as well as for tumourigenicity.

Methods: We used Hoechst 33342 to identify the SP from non-SP (NSP) cells in HK1 NPC cell line and xeno-284 NPC xenograft. The cells were assayed for *in vitro* characteristics of cancer stem cells (CSC), gene expression and tumourigenicity ability. Student's t test was used to test for significance.

Results: Five to ten percent and less than 0.5% of HK1 and xeno-284 NPC cells, respectively, were SP cells. Fumitremorgin C (FTC), as opposed to verapamil, was effective in causing the cells to retain Hoechst 33342 dye. HK1 SP cells formed more holoclones, had more aldehyde dehydrogenase (ALDH) activity, divided asymmetrically and contained slow-proliferating cells. *ABCG2*, *SOX2*, *TERT*, *MYC*, Hedgehog, Notch, TGFβ and Wnt signalling pathway genes were significantly upregulated in the SP cells. However, despite these differences *in vitro*, both HK1 SP and NSP cells had an overall similar tumourigenic potential *in vivo*.

Conclusions: HK1 SP cells were ABCG2-specific as confirmed by FTC inhibition and gene expression data. Despite data from *in vitro* and gene expression experiments suggesting stem-like features, there was no significant difference in tumourigenic potential between SP and NSP cells. We conclude that SP assay alone is not sufficient to identify CSCs in HK1 cells. Our work also suggests the presence of a stem-cell like population among NPC cells which do not display increased tumourigenicity.

Keywords: Side population, Nasopharyngeal carcinoma, Slow-cycling, Stem-like, Cancer stem cells

Background

Nasopharyngeal carcinoma (NPC) is the most common malignancy arising from the nasopharynx and its causation is closely associated with the Epstein-Barr virus, environmental as well as dietary factors [1]. Majority of NPC cases present in late stages [2]. The late presentation of the disease is due to the hidden location of the tumour, which could present with either no or apparently trivial symptoms which could be dismissed by patients or even medical professionals [3]. In addition, disease recurrence, therapeutic resistance and metastasis remain major clinical problems [4].

The cancer stem cell (CSC) model hypothesizes that there is a hierarchy within the tumour cell population and only a rare subset of cancer cells has the ability to self-renew and to differentiate, leading to the recapitulation of the original tumour [5]. As such, the spread of CSCs is an important component of the process of metastasis and, reactivation of CSC proliferation is believed to be the underlying cause of disease recurrence.

CSCs are found to behave differently from the rest of tumour cells; amongst others, they have enriched tumour-forming potential, and have efficient drug extrusion systems to evade most chemotherapeutic drugs [6]. These cells undergo asymmetric divisions to give rise to daughter

* Correspondence: alankhoo@imr.gov.my
[1]Molecular Pathology Unit, Cancer Research Centre, Institute for Medical Research, Jalan Pahang, 50588 Kuala Lumpur, Malaysia
Full list of author information is available at the end of the article

SCIENTIFIC REPORTS

OPEN

CD24, CD44 and EpCAM enrich for tumour-initiating cells in a newly established patient-derived xenograft of nasopharyngeal carcinoma

Received: 11 August 2016
Accepted: 25 August 2017
Published online: 28 September 2017

Susan Ling Ling Hoe[1,2], Lu Ping Tan[1], Norazlin Abdul Aziz[1], Kitson Liew[1], Sin-Yeang Teow[3], Fazlyn Reeny Abdul Razak[1], Yoon Ming Chin[2], Nurul Ashikin Mohamed Shahrehan[1], Tai Lin Chu[1], Noor Kaslina Mohd Kornain[4], Suat-Cheng Peh[3], Cheng Eng Koay[5], Kwok-Wai Lo[6], Munirah Ahmad[1], Ching-Ching Ng[2] & Alan Soo-Beng Khoo[1]

Subpopulations of nasopharyngeal carcinoma (NPC) contain cells with differential tumourigenic properties. Our study evaluates the tumourigenic potential of CD24, CD44, EpCAM and combination of EpCAM/CD44 cells in NPC. CD44br and EpCAMbr cells enriched for higher S-phase cell content, faster-growing tumourigenic cells leading to tumours with larger volume and higher mitotic figures. Although CD44br and EpCAMbr cells significantly enriched for tumour-initiating cells (TICs), all cells could retain self-renewal property for at least four generations. Compared to CD44 marker alone, EpCAM/CD44dbr marker did not enhance for cells with faster-growing ability or higher TIC frequency. Cells expressing high CD44 or EpCAM had lower KLF4 and p21 in NPC subpopulations. KLF4-overexpressed EpCAMbr cells had slower growth while Kenpaullone inhibition of *KLF4* transcription increased *in vitro* cell proliferation. Compared to non-NPC, NPC specimens had increased expression of *EPCAM*, of which tumours from advanced stage of NPC had higher expression. Together, our study provides evidence that EpCAM is a potentially important marker in NPC.

Nasopharyngeal carcinoma (NPC) is a type of head and neck cancer predominantly found in the southern Chinese population, several indigenous groups of the Southeast Asia, Amazigh- and Arabic-speaking populations of North Africa, and the Inuits of North America and Greenland[1]. It is infamously associated with high dietary intake of salted and/or preserved food, exposure to tobacco and formaldehyde as well as Epstein-Barr virus (EBV) infection[1]. Although it is a highly radiosensitive malignancy, its initial nonspecific clinical presentations such as nasal blockage, blood-tainted sputum, ringing in the ears and mild hearing loss are the main obstacles to early diagnosis[2,3] and the 5-year overall survival rate for NPC patients can reduce remarkably in late-stage of the disease[4].

Studies on NPC tumour biology such as growth, stemness, invasion, metastasis, therapy resistance and EBV presence have been widely investigated mainly using NPC cell lines. Stem-like cells from NPC cell lines were commonly enriched by using functional assays such as the side population (SP) assay[5–7], spheroid assay[8] and ALDH assay[9]. The SP cells in frequently used CNE-2 and HK1 cell lines were found to be more tumour-inducing in nude mice than bulk cells[6,7]. Cells from another NPC cell line, C666-1 which grew as free-floating spheres, displayed stemness characteristics such as a propensity for tumour formation, contained higher expression levels

[1]Molecular Pathology Unit, Cancer Research Centre, Institute for Medical Research, 50588, Kuala, Lumpur, Malaysia. [2]Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603, Kuala, Lumpur, Malaysia. [3]Sunway Institute for Healthcare Development, Sunway University, 47500, Bandar Sunway, Selangor, Malaysia. [4]Department of Pathology, Faculty of Medicine, Universiti Teknologi MARA (UiTM), 47000, Sungai Buloh, Selangor, Malaysia. [5]Gleneagles Hospital (Kuala Lumpur) Sdn. Bhd., Jalan Ampang, 50450, Kuala, Lumpur, Malaysia. [6]Li Ka Shing Institute of Health Science, The Chinese University of Hong Kong, Hong Kong, Hong Kong SAR. Correspondence and requests for materials should be addressed to C.-C.N. (email: ccng@um.edu.my) or A.S.-B.K. (email: alankhoo@imr.gov.my)

APPENDICES

Appendix A: List of genes in RT² Profiler™ PCR Array Human Stem Cell

Gene symbol	Description
ABCG2	ATP-binding cassette, sub-family G (WHITE), member 2
ACAN	Aggrecan
ACTC1	Actin, alpha, cardiac muscle 1
ADAR	Adenosine deaminase, RNA-specific
ALDH1A1	Aldehyde dehydrogenase 1 family, member A1
ALDH2	Aldehyde dehydrogenase 2 family (mitochondrial)
ALPI	Alkaline phosphatase, intestinal
APC	Adenomatous polyposis coli
ASCL2	Achaete-scute complex homolog 2 (Drosophila)
AXIN1	Axin 1
BGLAP	Bone gamma-carboxyglutamate (gla) protein
BMP1	Bone morphogenetic protein 1
BMP2	Bone morphogenetic protein 2
BMP3	Bone morphogenetic protein 3
BTRC	Beta-transducin repeat containing
CCNA2	Cyclin A2
CCND1	Cyclin D1
CCND2	Cyclin D2
CCNE1	Cyclin E1
CD3D	CD3d molecule, delta (CD3-TCR complex)
CD4	CD4 molecule

Gene symbol	Description
CD44	CD44 molecule (Indian blood group)
CD8A	CD8a molecule
CD8B	CD8b molecule
CDC42	Cell division cycle 42 (GTP binding protein, 25kDa)
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)
CDH2	Cadherin 2, type 1, N-cadherin (neuronal)
CDK1	Cyclin-dependent kinase 1
COL1A1	Collagen, type I, alpha 1
COL2A1	Collagen, type II, alpha 1
COL9A1	Collagen, type IX, alpha 1
CTNNA1	Catenin (cadherin-associated protein), alpha 1, 102kDa
CXCL12	Chemokine (C-X-C motif) ligand 12
DHH	Desert hedgehog
DLL1	Delta-like 1 (Drosophila)
DLL3	Delta-like 3 (Drosophila)
DTX1	Deltex homolog 1 (Drosophila)
DTX2	Deltex homolog 2 (Drosophila)
DVL1	Dishevelled, dsh homolog 1 (Drosophila)
EP300	E1A binding protein p300
FGF1	Fibroblast growth factor 1 (acidic)
FGF2	Fibroblast growth factor 2 (basic)
FGF3	Fibroblast growth factor 3
FGF4	Fibroblast growth factor 4

Gene symbol	Description
FGFR1	Fibroblast growth factor receptor 1
FOXA2	Forkhead box A2
FRAT1	Frequently rearranged in advanced T-cell lymphomas
FZD1	Frizzled family receptor 1
GDF2	Growth differentiation factor 2
GDF3	Growth differentiation factor 3
GJA1	Gap junction protein, alpha 1, 43kDa
GJB1	Gap junction protein, beta 1, 32kDa
GJB2	Gap junction protein, beta 2, 26kDa
HDAC2	Histone deacetylase 2
HSPA9	Heat shock 70kDa protein 9 (mortalin)
IGF1	Insulin-like growth factor 1 (somatomedin C)
ISL1	ISL LIM homeobox 1
JAG1	Jagged 1
KAT2A	K(lysine) acetyltransferase 2A
KAT7	K(lysine) acetyltransferase 7
KAT8	K(lysine) acetyltransferase 8
KRT15	Keratin 15
MME	Membrane metallo-endopeptidase
MSX1	Msh homeobox 1
MYC	V-myc myelocytomatosis viral oncogene homolog (avian)
MYOD1	Myogenic differentiation 1
NCAM1	Neural cell adhesion molecule 1

Gene symbol	Description
NEUROG2	Neurogenin 2
NOTCH1	Notch 1
NOTCH2	Notch 2
NUMB	Numb homolog (Drosophila)
PARD6A	Par-6 partitioning defective 6 homolog alpha (C. elegans)
PDX1	Pancreatic and duodenal homeobox 1
PPARD	Peroxisome proliferator-activated receptor delta
PPARG	Peroxisome proliferator-activated receptor gamma
RB1	Retinoblastoma 1
S100B	S100 calcium binding protein B
SIGMAR1	Sigma non-opioid intracellular receptor 1
SOX1	SRY (sex determining region Y)-box 1
SOX2	SRY (sex determining region Y)-box 2
T	T, brachyury homolog (mouse)
TERT	Telomerase reverse transcriptase
TUBB3	Tubulin, beta 3
WNT1	Wingless-type MMTV integration site family, member 1
ACTB	Actin, beta
B2M	Beta-2-microglobulin
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HPRT1	Hypoxanthine phosphoribosyltransferase
RPLP0	Ribosomal protein, large, P0

Appendix B: List of genes in RT² Profiler™ PCR Array Human Stem Cell Signaling

Gene symbol	Description
ACVR1	Activin A receptor, type I
ACVR1B	Activin A receptor, type IB
ACVR1C	Activin A receptor, type IC
ACVR2A	Activin A receptor, type IIA
ACVR2B	Activin A receptor, type IIB
ACVRL1	Activin A receptor type II-like 1
AMHR2	Anti-Mullerian hormone receptor, type II
BCL9	B-cell CLL/lymphoma 9
BCL9L	B-cell CLL/lymphoma 9-like
BMPR1A	Bone morphogenetic protein receptor, type IA
BMPR1B	Bone morphogenetic protein receptor, type IB
BMPR2	Bone morphogenetic protein receptor, type II (serine/threonine kinase)
CDX2	Caudal type homeobox 2
CREBBP	CREB binding protein
CTNNB1	Catenin (cadherin-associated protein), beta 1, 88kDa
E2F5	E2F transcription factor 5, p130-binding
ENG	Endoglin
EP300	E1A binding protein p300
FGFR1	Fibroblast growth factor receptor 1
FGFR2	Fibroblast growth factor receptor 2
FGFR3	Fibroblast growth factor receptor 3

Gene symbol	Description
FGFR4	Fibroblast growth factor receptor 4
FZD1	Frizzled family receptor 1
FZD2	Frizzled family receptor 2
FZD3	Frizzled family receptor 3
FZD4	Frizzled family receptor 4
FZD5	Frizzled family receptor 5
FZD6	Frizzled family receptor 6
FZD7	Frizzled family receptor 7
FZD8	Frizzled family receptor 8
FZD9	Frizzled family receptor 9
GLI1	GLI family zinc finger 1
GLI2	GLI family zinc finger 2
GLI3	GLI family zinc finger 3
IL6ST	Interleukin 6 signal transducer (gp130, oncostatin M receptor)
LEF1	Lymphoid enhancer-binding factor 1
LIFR	Leukemia inhibitory factor receptor alpha
LRP5	Low density lipoprotein receptor-related protein 5
LRP6	Low density lipoprotein receptor-related protein 6
LTBP1	Latent transforming growth factor beta binding protein 1
LTBP2	Latent transforming growth factor beta binding protein 2
LTBP3	Latent transforming growth factor beta binding protein 3
LTBP4	Latent transforming growth factor beta binding protein 4
NCSTN	Nicastrin
NFAT5	Nuclear factor of activated T-cells 5, tonicity-responsive

Gene symbol	Description
NFATC1	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2
NFATC3	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3
NFATC4	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4
NOTCH1	Notch 1
NOTCH2	Notch 2
NOTCH3	Notch 3
NOTCH4	Notch 4
PSEN1	Presenilin 1
PSEN2	Presenilin 2 (Alzheimer disease 4)
PSENEN	Presenilin enhancer 2 homolog (C. elegans)
PTCH1	Patched 1
PTCHD2	Patched domain containing 2
PYGO2	Pygopus homolog 2 (Drosophila)
RBL1	Retinoblastoma-like 1 (p107)
RBL2	Retinoblastoma-like 2 (p130)
RBPJL	Recombination signal binding protein for immunoglobulin kappa J region-like
RGMA	RGM domain family, member A
SMAD1	SMAD family member 1
SMAD2	SMAD family member 2
SMAD3	SMAD family member 3
SMAD4	SMAD family member 4
SMAD5	SMAD family member 5
SMAD6	SMAD family member 6

Gene symbol	Description
SMAD7	SMAD family member 7
SMAD8	SMAD family member 8
SMAD9	SMAD family member 9
SMO	Smoothened, frizzled family receptor
SP1	Sp1 transcription factor
STAT3	Signal transducer and activator of transcription 3 (acute-phase response factor)
SUFU	Suppressor of fused homolog (Drosophila)
TCF7	Transcription factor 7 (T-cell specific, HMG-box)
TCF7L1	Transcription factor 7-like 1 (T-cell specific, HMG-box)
TCF7L2	Transcription factor 7-like 2 (T-cell specific, HMG-box)
TGFBR1	Transforming growth factor, beta receptor 1
TGFBR2	Transforming growth factor, beta receptor II (70/80kDa)
TGFBR3	Transforming growth factor, beta receptor III
TGFBRAP1	Transforming growth factor, beta receptor associated protein 1
VANGL2	Vang-like 2 (van gogh, Drosophila)
ZEB2	Zinc finger E-box binding homeobox 2
ACTB	Actin, beta
B2M	Beta-2-microglobulin
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
HPRT1	Hypoxanthine phosphoribosyltransferase
RPLP0	Ribosomal protein, large, P0

Appendix C: Short tandem repeat (STR) profiling data of C666-1, HK1, xeno-284 and xeno-B110*

Cell line	D5S818	D13S317	D7S820	D16S539	vWA	THO1	TPOX	CSF1PO	D3S1358	FGA	D8S1179	D21S11	D18S51	D19S433	D2S1338	Amelogenin	% Match
C666-1[1]	11, 12	8, 11	11, 12	10	17, 18	6, 8	8, 11	11, 15, 16	16, 17	23, 24	11, 13, 14, 15	28, 29, 30.2, 31.2	16	13, 15.2	16, 23	X,Y	
C666-1[2]	11	8, 11	11, 12	10	17, 18	6, 8	8, 11	11	16, 17	24	11, 15	28, 31.2	16	13, 15.2	15, 16, 23	X, Y	85.7
HK1[1]	11, 13	11	8, 11	9, 11	18, 19	7, 9	8, 11	11, 12	15	26	11, 13	28, 29	17	14	20, 23	X	
HK1[2]	11, 13	11	8, 11	9, 11	18, 19	7, 9	8, 11	11, 12	15	26	11, 13	28, 29	17	14	20, 23	X	100
Patient's sample	11	12	10, 11	11, 12	17, 19	9, 10	8	10	16, 18	19	14	29, 30	14, 17	13	19, 20	X, Y	
xeno-284	11	12	10, 11	11, 12	17, 19	9, 10	8	10	16	19	14, 15	29	14, 17	13	19, 20	X	91.7
Patient's sample	12	8, 9	10, 11	10	18, 20	6.3, 8.3	11	11	15, 16	23, 24	15	29, 33.2	14, 20	13	17, 23	X, Y	
xeno-B110	12	8, 9	10, 11	10	18, 20	9	11	11	16	19, 23, 24	15	29, 33.2	14, 20	13	17, 23	X, Y	90.2

[1] published data from Chan *et al.* (2008)

[2] data from cell line used in this study

* sample preparation and data analysis were performed by Pauline Balraj (HK1 cells), Norazlin Abdul Aziz (patients' samples) and Dr. Kitson Liew (C666-1, xeno-284 and xeno-B110) (Institute for Medical Research, Malaysia)

Appendix D: List of 50 genes significantly deregulated by at least 2 folds in SP cells as compared with NSP cells ($p < 0.05$)

Gene	Gene description	Fold change	p-value
Pluripotency maintenance			
SOX2	SRY (sex determining region Y)-box 2	2.87	0.002
STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	2.02	0.01
FGF signalling			
FGFR3	fibroblast growth factor receptor 3	20.45	0.03
Hedgehog signalling			
GLI2	GLI family zinc finger 2	24.02	0.000
GLI1	GLI family zinc finger 1	14.12	0.002
PTCHD2	Patched domain containing 2	9.76	0.02
SUFU	suppressor of fused homolog (Drosophila)	6.26	0.01
PTCH1	Patched 1	3.52	0.02
GLI3FL	GLI family zinc finger 3	2.40	0.03
Notch signalling			
NOTCH1	Notch 1	6.04	0.04
PSENEN	presenilin enhancer 2 homolog (C. elegans)	5.04	0.001
JAG1	Jagged 1	2.50	0.01
DTX2	Deltex homolog 2 (Drosophila)	2.27	0.03
NCSTN	nicastrin	2.26	0.004
DLL1	Delta-like 1 (Drosophila)	2.21	0.001
TGFβ signalling			
RGMA	RGM domain family, member A	74.27	0.01
ENG	endoglin	19.35	0.01
LTBP3	latent transforming growth factor beta binding protein 3	14.78	0.002
E2F5	E2F transcription factor 5, p130-binding	7.85	0.04
LTBP2	latent transforming growth factor beta binding protein 2		
ACVR1B	activin A receptor, type IB	4.90	0.003
SMAD1	SMAD family member 1	3.49	0.01
EP300	E1A binding protein p300	3.42	0.003
TGFBRAP1	transforming growth factor, beta receptor associated protein 1	2.82	0.02
BMPR2	bone morphogenetic protein receptor, type II (serine/threonine kinase)	2.76	0.03
SMAD7	SMAD family member 7	2.05	0.01

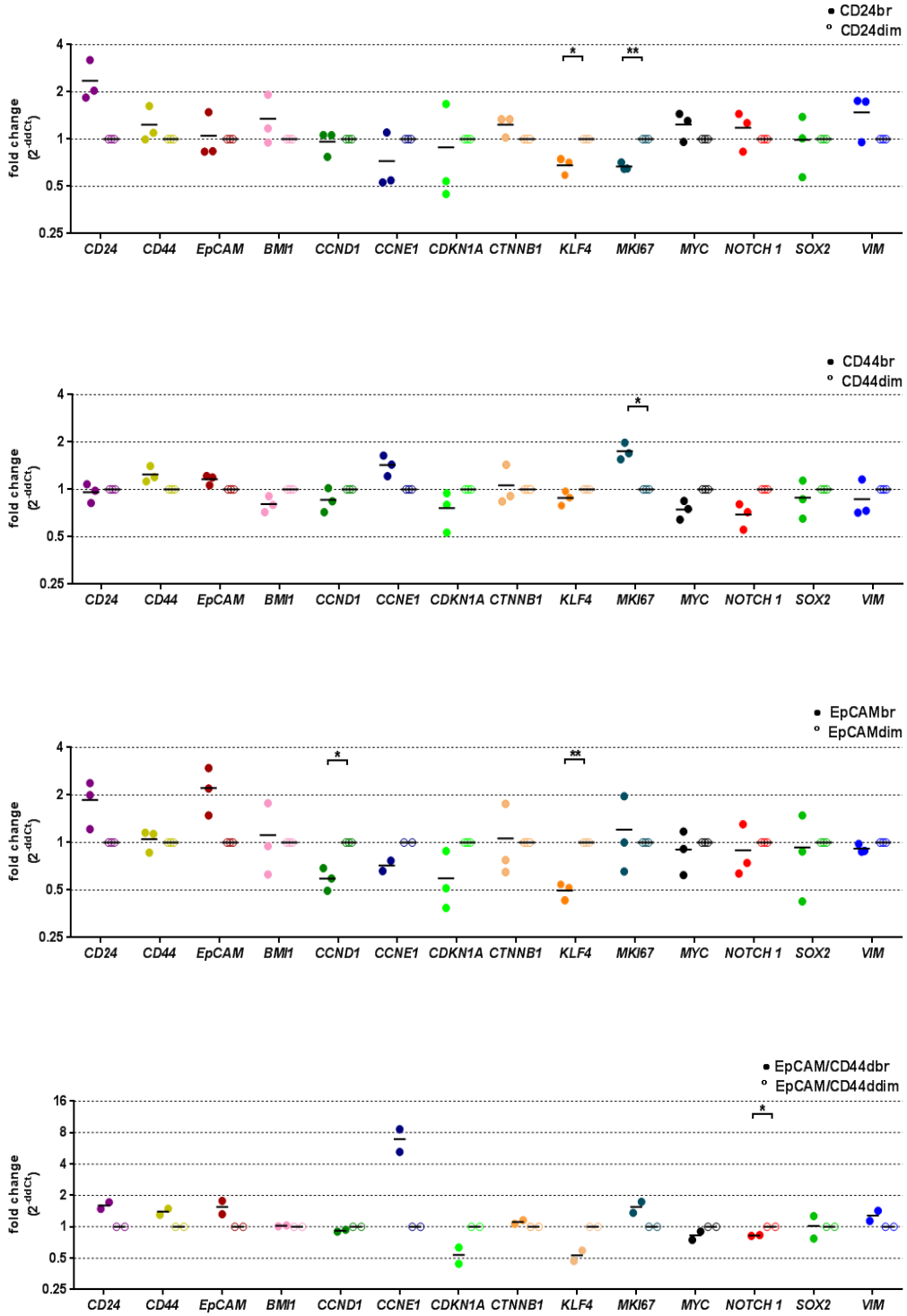
Gene	Gene description	Fold change	p-value
Wnt signalling			
LRP5	low density lipoprotein receptor-related protein 5	5.22	0.02
PYGO2	Pygopus homolog 2 (Drosophila)	5.22	0.04
BCL9	B-cell CLL/lymphoma 9	4.80	0.01
FZD7	Frizzled family receptor 7	4.48	0.01
FZD2	Frizzled family receptor 2	3.88	0.04
AXIN1	Axin 1	3.39	0.002
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	3.01	0.02
NFAT5	nuclear factor of activated T-cells 5, tonicity-responsive	2.61	0.01
FZD6	Frizzled family receptor 6	2.41	0.004
PPARD	peroxisome proliferator-activated receptor delta	2.09	0.01
BCL9L	B-cell CLL/lymphoma 9-like	2.02	0.01
Miscellaneous			
CD4	CD4 molecule	10.54	0.03
ALPI	alkaline phosphatase, intestinal	9.48	0.03
TERT	telomerase reverse transcriptase	5.44	0.003
KRT15	keratin 15	4.69	0.01
S100B	S100 calcium binding protein B	3.98	0.04
BGLAP	bone gamma-carboxyglutamate (gla) protein	3.81	0.01
T	T, brachyury homolog (mouse)	2.88	0.02
ABCG2	ATP-binding cassette, sub-family G (WHITE), member 2	2.63	0.003
KAT2A	K(lysine) acetyltransferase 2A	2.62	0.002
JAG1	Jagged 1	2.50	0.006
BMP1	bone morphogenetic protein 1	2.44	0.03
COL1A1	collagen, type I, alpha 1	2.09	0.03
TUBB3	tubulin, beta 3	2.04	0.03

Appendix E: List of TaqMan® assays used in RT-qPCR with preamplification step

Assay	TaqMan ID	Remark
HPRT1	Hs01003267_m1	Ok
RPL13A	Hs01926559_g1	Ok
UBC	Hs008 24723_m1	Ok
ABCG2	Hs01053790_m1	Failed
ALDH1A1	Hs00946916_m1	Failed
BMI1	Hs00180411_m1	Ok
CCND1	Hs00765553_m1	Ok
CCNE1	Hs01026536_m1	Ok
CD24	Hs02379687_s1	Ok
CD44 (all 8 isoforms)	Hs01075862_m1	Ok
CDKN1A	Hs00355782_m1	Ok
CTNNB1	Hs00355049_m1	Ok
EPCAM	Hs00158980_m1	Ok
GLI1	Hs01110766_m1	Undetected
KIT	Hs00174029_m1	Failed
KLF4	Hs00358836_m1	Ok
LMP1	custom	Undetected
LMP2A	custom	Undetected
MKI67	Hs01032443_m1	Ok
MYC	Hs00905030_m1	Ok
NANOG	Hs02387400_g1	Failed
NES	Hs04187831_g1	Failed
NOTCH1	Hs01062014_m1	Ok
SOX2	Hs01053049_s1	Ok
VIM	Hs00185584_m1	Ok

Failed, failed QC due to non-linear amplification or poor PCR efficiency
Undetected, more than 35% of samples below detection limit (Ct = 999)

Appendix F: Expression levels of 14 genes which passed QC, showed amplification and were analysed in CD24, CD44, EpCAM and EpCAM/CD44 groups of cells from xeno-B110



Results, mean of 2 or 3 sorting replicates per transcript. * $p < 0.05$, ** $p < 0.01$