

5-2015

Functional In Vivo Screen Identifies Pygo2 As A Putative Gene To Promote Prostate Cancer

Xiaolu Pan

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Medicine and Health Sciences Commons](#)

Recommended Citation

Pan, Xiaolu, "Functional In Vivo Screen Identifies Pygo2 As A Putative Gene To Promote Prostate Cancer" (2015). *Dissertations and Theses (Open Access)*. 587.

https://digitalcommons.library.tmc.edu/utgsbs_dissertations/587

This Thesis (MS) is brought to you for free and open access by the MD Anderson UTHealth Houston Graduate School at DigitalCommons@TMC. It has been accepted for inclusion in Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digcommons@library.tmc.edu.

FUNCTIONAL IN VIVO SCREEN IDENTIFIES PYGO2 AS A PUTATIVE GENE TO PROMOTE PROSTATE CANCER

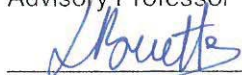
by

Xiaolu Pan, MB

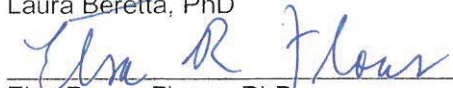
APPROVED:



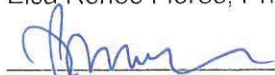
Ronald DePinho, MD
Advisory Professor



Laura Beretta, PhD



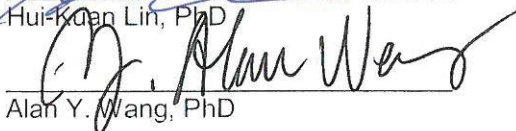
Elsa Renee Flores, PhD



Prashad Ram, PhD



Hui-Kuan Lin, PhD



Alan Y. Wang, PhD

APPROVED:

Dean, The University of Texas

Graduate School of Biomedical Sciences at Houston

FUNCTIONAL IN VIVO SCREEN IDENTIFIES PYGO2 AS A PUTATIVE GENE TO PROMOTE PROSTATE CANCER

by

Xiaolu Pan, MB

APPROVED:

Ronald DePinho, MD
Advisory Professor

Laura Beretta, PhD

Elsa Renee Flores, PhD

Prahlad Ram, PhD

Hui-Kuan Lin, PhD

Alan Wang, PhD

APPROVED:

Dean, The University of Texas

Graduate School of Biomedical Sciences at Houston

FUNCTIONAL IN VIVO SCREEN IDENTIFIES PYGO2 AS A PUTATIVE GENE TO PROMOTE PROSTATE CANCER

A

THESIS

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences
in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

by

Xiaolu Pan, MB
Houston, Texas

May, 2015

Acknowledgements

I would like to thank Dr. Ronald DePinho for his mentorship towards the completion of my thesis. His knowledge and vision have constantly inspired everyone around him including me. I am also grateful to Dr. Alan Y. Wang, who gave me the precious opportunity to study in Dr. DePinho's laboratory and he has provided significant input in my project and my career development. I would like to express my sincere gratitude to Drs. Xin Lu and Guocan Wang, who supervised me in the lab and taught me many molecular biology techniques as well as animal handling and surgery. I am grateful that they respect and trust me as a young scientist from the beginning of my training. Their passion for science, sustaining commitment for new discoveries and scientific mindset are inspiring. During the past three years, I have had the opportunities to interact with many members from Drs. DePinho, Draetta and Chin's laboratories (Wantong Yao, Haoqiang Yin, Avinish Kapoor, Baoli Hu, Flourian Muller, April, Melody, Nikunj, Ellen and Fred, Qiuyun Wang, Ram Vandahana, Pingna Deng, Eun-Jung Jin, Edward Chang, Trang Tieu) who have generously provided invaluable contribution in completion of my thesis project. This thesis could not have been completed without tremendous scientific input from Dr. Ram Prahlaad, Dr. Terrance Wu and Samir Amin; who provided necessary bioinformatic analyses. I would also like to thank Shan Jiang for her gentle caring of my mouse colonies and me. I would like to acknowledge the fruitful ongoing collaborations with Dr. Chunru Lin on elucidating

interaction networks. I am also very grateful to my advisory committee members who are so supportive all the time and who appreciate my work.

I would like to thank Dr. Xiaoping Huo, my dearest mother who dedicated her master thesis to me and now I will dedicate mine to her. Special thanks also goes to all my family and friends, especially those I met in Houston, for sharing their precious love and support.

Functional *in vivo* Screen Identifies PYGO2 as a Putative Gene to Promote Prostate Cancer

by

Xiaolu Pan, MB

Poor prognosis of prostate cancer is correlated with rampant chromosomal copy number alterations, highlighting the potential function of genes with copy number gains and losses in driving prostate cancer progression. To identify putative genes promoting prostate cancer, an *in vivo* tumorigenesis screen was performed for 286 genes that are recurrently amplified and overexpressed in human prostate cancer. The transcription co-activator protein PYGO2 was identified as a major hit for further *in vivo* functional and clinical validation. Overexpression of PYGO2 could enhance primary tumor growth as well as local invasion to lymph nodes using AR-positive prostate cancer cell line LNCaP. PYGO2 may mediate its pro-tumor function through upregulation of genes including WNT2, ADAMTS2, IGFBP3 and downregulation of KISS1. Tissue microarray analysis indicated that PYGO2 upregulation was correlated with higher Gleason score in prostate cancer. Collectively, the results suggest PYGO2 as a potential prognostic marker as well as a therapeutic target. Additional functional characterization of PYGO2 in prostate cancer pathogenesis is warranted and ongoing.

Table of Contents

Chapter 1	Introduction.....	1
1.1	Introduction	2
1.2	Anatomy of human and mouse prostate	5
1.3	The biology and progression of PCa.....	5
1.4	Subtypes of PCa	7
1.5	PCa genomics and amplified genes.	7
Chapter 2	Functional <i>in vivo</i> screen identifies putative drivers of prostate cancer	9
2.1	Introduction	10
2.2	Material and methods	12
2.2.1	Bioinformatic analysis of global gene expression to determine progression genes.	12
2.2.2	Cell culture	15
2.2.3	Establishment of the library of blasticidin-selected stable cell lines overexpressing 286 candidate genes	15
2.2.4	Xenograft studies and <i>in vivo</i> ORF screen for putative genes promoting PCa.....	17
2.2.5	RNA extraction, cDNA synthesis and quantitative RT-PCR.....	18
2.2.6	Genomic sequencing	18

2.2.7	Protein lysates and western blot assays	18
2.2.8	Proliferation analysis.....	19
2.2.9	Soft agar analysis	19
2.2.10	Migration assay.....	19
2.2.11	Matrigel invasion assay.....	20
2.2.12	Tissue specimen, histology and immunohistochemistry	20
2.2.13	Expression profiling	21
2.2.14	Bioinformatic study and statistical analysis	21
2.3	Results.....	21
2.3.1	Identification of putative genes contributing to tumor progression	21
2.3.2	Functional validation using cell proliferation assay	28
2.3.3	Genes promoting anchorage-independent growth.....	29
2.3.4	Functional validation by migration/invasion assay	31
2.3.5	Alteration of MAPK pathway activity	33
2.3.6	PYGO2 up-regulation was correlated with PCa pathological aggressiveness.....	35
2.3.7	PYGO2 promoted LNCaP local invasion to draining lymph nodes	38
2.3.8	Elevated expression of Pygo2 in Wnt-upregulated genetic engineered mouse models (GEMMs) and in PCa patients.....	42
2.3.9	PYGO2 expression in established PCa cell lines.....	51

Chapter 3	Discussion and future directions	53
3.1	Discussion	54
3.1.1	Functional <i>in vivo</i> screen identify genes with potential to promote PCa progression	54
3.1.2	Activation of MAPK in PCa	56
3.1.3	PYGO2 function in PCa	57
3.1.4	PYGO2 upregulation in PCa	59
3.1.5	PYGO2 mechanism for PCa	60

List of Figures

Figure 1.1 Currently histopathological and molecular genetic model of PCa development. Taken from (Shen and Abate-Shen, 2010) with permission from Cold Spring Harbor Laboratory (CSHL) Press	3
Figure 2.1 Bioinformatic genomic analysis to generate amplified/overexpression gene list for PCa progression.....	14
Figure 2.2 Workflow of <i>in vivo</i> ORF screen for amplified/overexpressed genes .	16
Figure 2.3 Vector structure for LentiORF library establishment (Taken From Thermo Fisher Scientific Inc., Technical Manual, Thermo Scientific Open Biosystems Precision LentiORF Collection).....	17
Figure 2.4 Representative hits growth of ORFs	25
Figure 2.5 Representative validation of the expression level of interested genes in injected LHMK lines	26
Figure 2.6 2D proliferation assays showed similar growth rates among LHMK cells overexpressing ORFs	28
Figure 2.7 ORFs facilitated anchorage-independent growth of LHMK cells.....	30
Figure 2.8 Hits showed promoting LHMK <i>in vitro</i> migration (a) and invasion (b) in Boyden chamber assay.	32
Figure 2.9 Immunoblotting analysis of various cellular signaling proteins in cell lysates of LHMK cells overexpressing ORFs	34
Figure 2.10 Expression intensity of PYGO2 in PCa patient samples from low to high (0-4).	36

Figure 2.11 Representative PDX samples showed strong positive staining of PYGO2	38
Figure 2.12 PYGO2 enhanced LNCaP subcutaneous tumor progression <i>in vivo</i>	39
Figure 2.13 PYGO2 promotes LNCaP local invasion to draining lymph nodes...	41
Figure 2.14 Wnt pathway enrichment in local (a) and metastasis (b) PCa samples with higher PYGO2 expression level at RNA level.....	43
Figure 2.15 <i>Pygo2</i> IHC on anterior and dorsolateral prostate tissue from wild type, <i>PB-Cre⁺ PTEN^{L/L}</i> and <i>PB-Cre⁺ PTEN^{L/L} APC^{L/L}</i> mice.....	45
Figure 2.16 Microarray analysis of LHMK and overexpression cells.....	47
Figure 2.17 Expression level of WNT2, IGFBP3, ADAMTS2 and KISS1 and PYGO2 in LHMK and LHMK-PYGO2	51
Figure 2.18 Investigation of PYGO2 expression level of PCa cell lines	52
Figure 3.1 PYGO2 is upregulated in metastases in several ONCOMINE® datasets	59
Figure 3.2 Current working hypothesis of PYGO2 in PCa	63

List of Tables

Table 2.1 List of ORFs facilitating LHMK xenograft growth in injected mice in 6 months	25
Table 2.2 List of ORFs that promoted multiple tumor growth.....	28
Table 2.3 PYGO2 up-regulation in PCa patient samples (n=71).....	35
Table 2.4 PYGO2 up-regulation is associated with high Gleason Score (n=71).....	37
Table 2.5 High penetrance of LNCaP local invasion to draining lymph nodes in PYGO2 overexpressed cells	40
Table 2.6 Up-regulated and down-regulated genes in PCa	50
Table 3.1 Amplification of prioritized hit in PCa genomic datasets.....	55

List of Abbreviations

µg	microgram
µM	micrometer
ADT	androgen deprivation therapy
AP	anterior prostate
AR	androgen receptor
ATCC	American Type Culture Collection
BF	bright field
BPH	benign prostatic hyperplasia
CRPC	castration resistant prostate cancer
DLP	dorsolateral prostate
DMEM	Dulbecco's Modified Eagle's Medium
FBS	fetal bovine serum
FDR	false discovery rate
FITC	fluorescein isothiocyanate
GEMM	genetically engineered mouse model
GFP	Green fluorescent protein
GISITIC	Genomic Identification of Significant Targets in Cancer
HE	hematoxylin and eosin staining
IHC	immunohistochemistry
KEGG	Kyoto Encyclopedia of Genes and Genomes
Kg	kilogram

mg	milligram
mL	milliliter
nM	nanometer
ORF	Open-Reading Frame
PBS	phosphate buffered saline
Pca	prostate cancer
PDX	patient-derived xenograft
PHD	plant homeodomain
PIN	prostatic intraepithelial neoplasia
PrEC	prostate epithelial cells
qPCR	quantitative real time polymerase chain reaction
RFP	red fluorescent protein
RPMI	Roswell Park Memorial Institute (medium)
RT-PCR	reverse transcription polymerase chain reaction
s. q.	subcutaneous(ly)
SDE	standard deviation ellipse
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCGA	The Cancer Genome Atlas

Chapter 1 Introduction

1.1 Introduction

Prostate cancer (PCa) is the most common non-cutaneous cancer in men. In 2013, around 238,590 new cases and 29,720 related deaths were reported (Siegel *et al.*, 2013). Several decades of PCa research has helped to delineate the core progression pathways for human PCa, which sequentially develops from normal epithelium, to prostatic intraepithelial neoplasia (PIN), to latent and then clinically manifested adenocarcinoma, and ultimately to distant metastases especially to bone and also to other organs (Figure 1.1) (Shen and Abate-Shen, 2010b). Accompanying the key histopathological features are some well characterized molecular changes such as TMPRSS2-ERG translocation, PTEN inactivation and EZH2 overexpression. Despite the pathological and molecular understanding, current effective therapies for PCa are largely limited to androgen deprivation therapy (ADT) with an initial success rate of close to 90%. However, most PCa patients develop refractory disease to surgical or chemical castration, resulting in the development of castration resistant prostate cancer (CRPC) with high morbidity and mortality. Therefore, there remains a critical need for better understanding the etiology of aggressive PCa, in particular, the identification of bona fide PCa genes and specific molecular mechanisms that may potentially generate new therapeutic interventions.

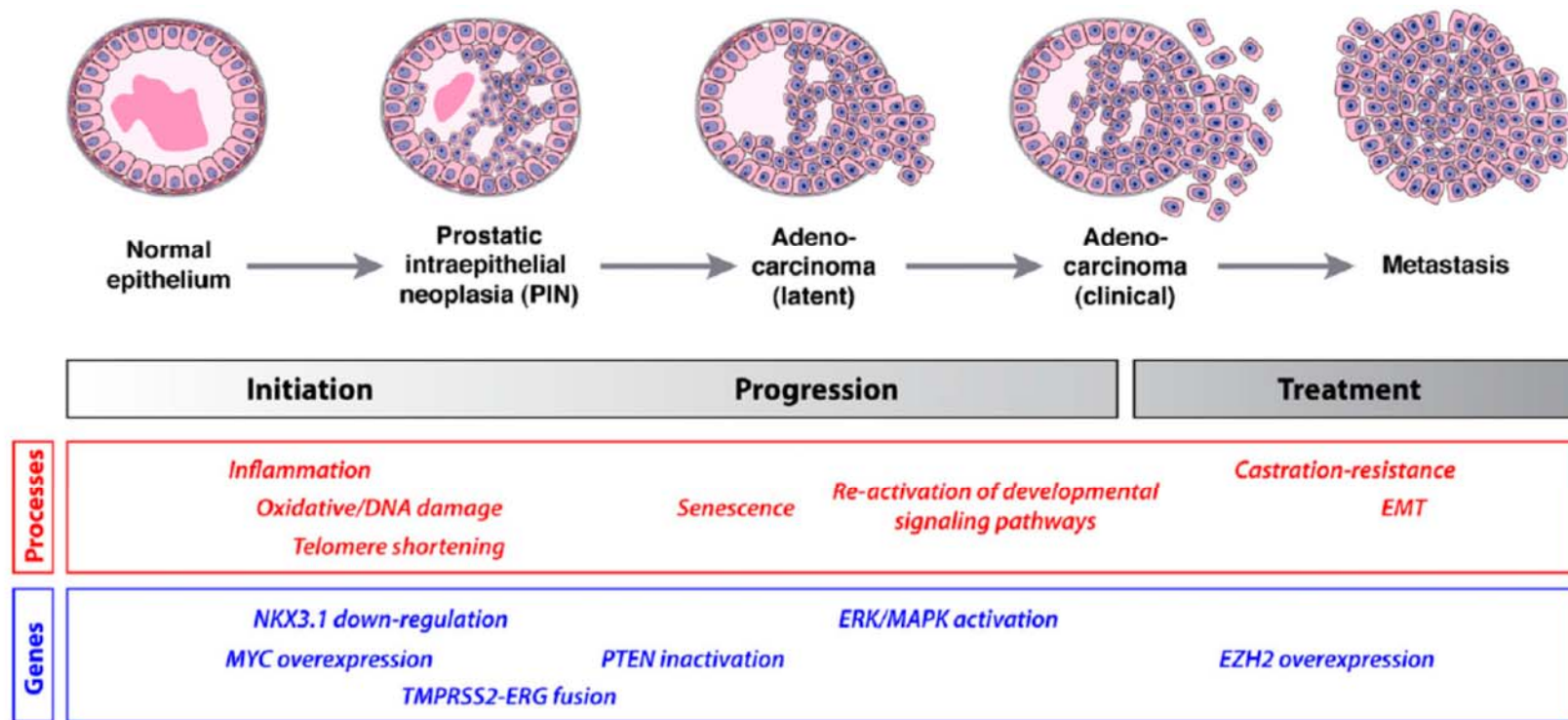


Figure 1.1 Currently histopathological and molecular genetic model of PCa development. Taken from (Shen and Abate-Shen, 2010) with permission from Cold Spring Harbor Laboratory (CSHL) Press

Recent genomic profiling of various cancer types, in particular efforts triumphed by The Cancer Genome Atlas (TCGA), has revealed that different cancer types displaying distinct mutation rates. Compared to cancers such as melanoma and lung cancer, PCa has a relatively low mutation rate. It highlights the potential implication of other types of genetic or epigenetic alternations in driving cancer progression (Lawrence *et al.*, 2013). The clinical significance of genomic instability in promoting PCa was indicated by the observed correlation of rampant genomic gains and losses with poor prognosis, while transcriptomic profiles were unable to provide significant prognostic value (Taylor *et al.*, 2010). Therefore, a critically important research direction for dissecting PCa genetics would be the identification of driver genes that are embedded in the amplified or deleted peaks of the PCa genomes. My thesis project was set to address this question by taking a genome-informed functional screen approach, in particular, an open-reading frame (ORF) screen in mice for genes that are recurrently amplified and overexpressed in PCa genomes. 286 genes were screened for their functional contribution to promoting prostate tumorigenesis in mice. Several hits are identified and in the stage of further validation and characterization. In particular, the major focus of this study is on PYGO2 which may function as a tumor-promoting gene and may serve as a potential prognostic marker based on my preliminary data.

1.2 Anatomy of human and mouse prostate

Prostate gland is a walnut-sized male sex hormone regulated organ. It is located at the base of the bladder, surrounding the urethra. Prostate gland secretion usually constitutes 50-75% of the volume of the semen (Leo Shedlovsky, 1942). The mouse prostate is lobular with anterior (AP), ventral, dorsal, and lateral prostate. The last two lobes commonly combine as dorsolateral prostate (DLP) (Abate-Shen and Shen, 2002). Unlike the mouse prostate, the human adult prostate lacks clear lobular structure but consists of a zonal architecture. Central, periurethral transition, and peripheral zones as well as an anterior fibromuscular stroma constitute the prostate. Specifically, peripheral zone occupies ~70% of the prostate volume and harbors the majority of PCa. Most of the benign prostatic hyperplasia (BPH) arises from the transition zone.

1.3 The biology and progression of PCa

The heterogeneity of PCa is illustrated by the multifocal events in primary tumors. The neoplasm often contains multiple genetically-independent histologic foci that commonly attribute to the multi-event neoplastic transformation that parallels aging. The earliest initiation in men can be as early as 20 years of age (Sakr *et al.*, 1994; Yatani *et al.*, 1989). While some foci progress to clinical detectable cancerous lesion, latent PCa remains inactive likely due to a lack of certain genetic events that lead to an aggressive behavior.

High-grade prostatic intraepithelial neoplasia (PIN) refers to abnormal proliferation with no stromal invasion (Bostwick, 2000). Though some controversies exist (DeMarzo *et al.*, 2003), high-grade PIN is generally accepted as a pre-invasive stage for PCa (Bostwick, 2000). The aggressive lesions may progress to PCa, marked by pathologically absent basal cell layer (Bostwick *et al.*, 2004). Remote metastasis follows invasion, mostly located in bone (90%) (Bubendorf *et al.*, 2000). Surgical or chemical ADT are applied but almost all patients develop CRPC eventually (Felici *et al.*, 2012). Notably, metastases from the same patient maintain a signature pattern of copy number, mutation status, erythroblast transformation specific rearrangement methylation, and phosphorylation (Aryee *et al.*, 2013; Drake *et al.*, 2013; Grasso *et al.*, 2012; Liu *et al.*, 2009; Mehra *et al.*, 2008).

As a cancer type with inherent heterogeneity, PCa relies on a unique scoring system to inform diagnosis and prognosis (Gleason, 1992). PCa, especially prostatic glandular carcinomas, is graded by Gleason system according to the histomorphological appearance of biopsies (Mellinger *et al.*, 1967). Five basic grade patterns from 1 to 5 represent the extent of glandular differentiation. Two scores are summed with the first score, assigned to the dominant pattern (occupied more than 50%), and the one second score, assigned to the next-most frequent pattern (less than 50%). Histological grading of Gleason score is by far the most prevailing indicator for clinical outcome of patients (Humphrey, 2004). It is

proved by multiple studies linking it to the overall survival of PCa (Gonzalzo et al., 2006; Melissari et al., 2006).

1.4 Subtypes of PCa

It is notable that PCa lacks distinguishable histopathological subtypes that could guide its prognosis or treatment response; compared to other epithelial tumors, such as breast cancer or lung cancer. Most cases of the PCa are acinar adenocarcinomas that express AR. Other categories of PCa are extremely rare; such as ductal adenocarcinoma, mucinous carcinoma, and signet ring carcinoma (Grignon, 2004). Neuroendocrine small cell carcinomas representing <2% of PCa cases are generally classified as either small cell carcinoma or carcinoid tumor.

1.5 PCa genomics and amplified genes.

Large-scale genomic analysis characterizes genetic alterations in PCa. These alternations, including point mutations, deletions, re-arrangements and amplification, contribute to multi-steps of tumorigenesis.

PCa genome harbors relatively moderate number of point mutations (Berger *et al.*, 2011; Taylor *et al.*, 2010). Common mutated genes in other tumor types, including *TP53*, *PTEN* and *KRAS*, usually are not frequently observed in PCa genome (Taylor *et al.*, 2010). *TMPRSS2:ERG* gene fusions are consistently to be reported in about 50 percent of clinically localized PCa samples (Kumar-Sinha *et*

et al., 2008). The most frequent alteration in PCa genome is loss of chromosome 8q, harboring *NKX3.1*. Some significant deletion peaks target *PTEN*, *RB1* and *TP53* (Taylor *et al.*, 2010). While the mutation rate for PCa is moderate, genomes of more aggressive PCa and CRPC are featured by rampant chromosomal instability (Taylor *et al.*, Grasso *et al.*). Genomic gains can have functional consequences, as evidenced by known cancer genes altered by amplification which leads to overexpression (Santarius *et al.*, 2010). For example, a frequently amplified region in cancer genome, 8q24.21 which encompasses *MYC* and *NCOA2*, is also among the most amplified regions in aggressive PCa (Taylor *et al.*). It is also of interest to note *AR* (Xq12) amplification in metastatic CRPC (Grasso *et al.*). An unanswered question remains, as to what other genes involved in these chromosomal gains functionally contribute to the progression of PCa.

Chapter 2 Functional *in vivo* screen
identifies putative drivers of prostate
cancer

2.1 Introduction

Increasingly high-resolution genomic studies with next-generation sequencing (NGS) technology reveal recurrent focal deletions and amplifications in cancer genome (Chin *et al.*, 2011; Yates and Campbell, 2012). For prostate, a handful of large-scale genomic profiling studies have been reported, including Taylor *et al.* (MSKCC dataset), Grasso *et al.* (Michigan dataset), Barbieri *et al.* (Harvard dataset). A large portion of genomic data of prostate cancer TCGA can be downloaded and analyzed, although the paper on prostate cancer TCGA has not been published yet. From these datasets, putative amplified gene list was compiled. It was expected some of the genes on the list should play functional roles in promoting prostate tumor progression, and through a proper selection of a mouse model system which had low tumor formation background by the tumor cell line alone, potential driver genes on the list could be identified.

Following this rationale, we performed genome-informed *in vivo* screen, and 41 hits were identified. The mechanisms for promoting PCa progression of the top 12 amplified genes were further analyzed. Moving the project forward *PYGO2* was set as focus based on the following reasons: 1) overexpression of *PYGO2* induced anchorage-independent colony formation in LHMK cell line; 2) *PYGO2* promoted *in vitro* migration and invasion of LHMK cell line; 3) Recent studies in development biology has suggested *PYGO2* as an important gene regulating stem cell function

and tissue homeostasis in skin and mammary and spermiogenesis (Gu *et al.*, 2013; Nair *et al.*, 2008; Sun *et al.*, 2014).

Pygopus was first reported as an essential transcription co-activator with Arm/ β -catenin-Tcf complex for Wnt signal transduction pathways in *Drosophila* (Belenkaya *et al.*, 2002). Two mouse and human pygopus genes *Pygo1* and *Pygo2*, have been identified. The later expressed in a broader range of adult and developing tissue consistent with the Wnt signaling activity to regulate proper development and maintenance (Li *et al.*, 2004). In contrast with *Pygo1* null mice, *Pygo2* null mice presented serious developmental defects like lens agenesis and a kidney phenotype with high penetrance, exencephaly, and cleft palate incomplete penetrance and exhibit perinatal lethality (Li *et al.*, 2007; Schwab *et al.*, 2007).

PYGO2 has highly conserved plant homeodomain (PHD) in its C-terminus that is associated with histone modifications (Miller *et al.*, 2013). It was reported to directly bind to histone H3 (Fiedler *et al.*, 2008) (Gu *et al.*, 2013) (Kessler *et al.*, 2009) and recruited histone-modifying enzymes to generate more H3K4me as active histone marks to facilitate transcription (Gu *et al.*, 2009). The recruitment of PYGO2 was associated with Wnt pathway activation (Städeli and Basler, 2005) and Rb attenuation (Tzenov *et al.*, 2013). Moreover, PYGO2 has also been implicated in histone acetylation independent of Wnt signaling (Nair *et al.*, 2008). Interestingly, a recent paper reported lncRNA PCGEM1 recruited PYGO2 to enhance AR-bound enhancers targeting gene promoters (Yang *et al.*, 2013).

Despite the reported up-regulation in skin (Sun *et al.*, 2014), hepatocellular carcinoma (Zhang *et al.*, 2014), and lung cancer (Zhou *et al.*, 2014), the function of PYGO2 has not been thoroughly investigated in PCa.

2.2 Material and methods

2.2.1 Bioinformatic analysis of global gene expression to determine progression genes.

Exome sequencing data is frequently adopted in current study to investigate the somatic copy-number amplification in cancer genome. To generate a list for putative genes that are recurrently amplified in PCa and also expressed at higher level in metastasis than primary tumor, Dr. Xin Lu and Dr. Terrance Wu conducted an integrative bioinformatics analysis on four exome sequencing datasets, Grasso and colleagues (Grasso *et al.*, 2012), Taylor and colleagues (Aravindaram and Yang, 2010), Barbieri and colleagues (Barbieri *et al.*, 2012) and the Cancer Genome Atlas (TCGA; Research Networks, <https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>, DOI # 2012-10-04). In the analysis, recurrently amplified genes across multiple datasets were further selected based on copy number – expression correlation ($p < 0.01$) and displaying higher expression level in metastasis compared with primary tumor in at least 3 out of 8 microarray-based expression datasets from Oncomine® (Figure 2.1) Lu and Wu also reasoned that some pro-metastasis genes can be upregulated through non-amplification mechanisms therefore a separate branch of analysis was applied. Genes were selected if they were highly expressed in metastatic compared to

primary tumor in 6 out of 8 Oncomine® datasets (Figure 2.1). To elaborate the process in more details, majority of the genes are collected to identify the putative driver genes at amplicons. The genes at focal amplification peaks were selected by GISTIC (Bubendorf *et al.*, 2000) analysis on the previously mentioned prostate exome sequencing studies. G-score was set as a cut-off to eliminate unimportant passenger alterations. The evaluation of copy number-expression correlation was followed to identify the amplified genes correlated with up-regulation. A cutoff was also set to spotlight genes more enriched in metastasis comparing to primary sites because they are more relevant to aggressive PCa progression. It is based on the following criteria: call if $p < 0.05$ (student t test) in 3 out of 8 expression datasets from Oncomine® (<https://www.oncomine.org>). On the other hand, in order to include some of the functionally important genes with high expression levels but without discernible amplification cases in analyzed patient samples, we added genes based on the following criteria: call if $p < 0.05$ (student t test) in 6 out of 8 expression datasets from Oncomine® (<https://www.oncomine.org>). Following these strategies, in total 741 genes were compiled. From the MDA LentiORF library, 286 ORFs were available, whose vectors were the reagents for the functional *in vivo* screen

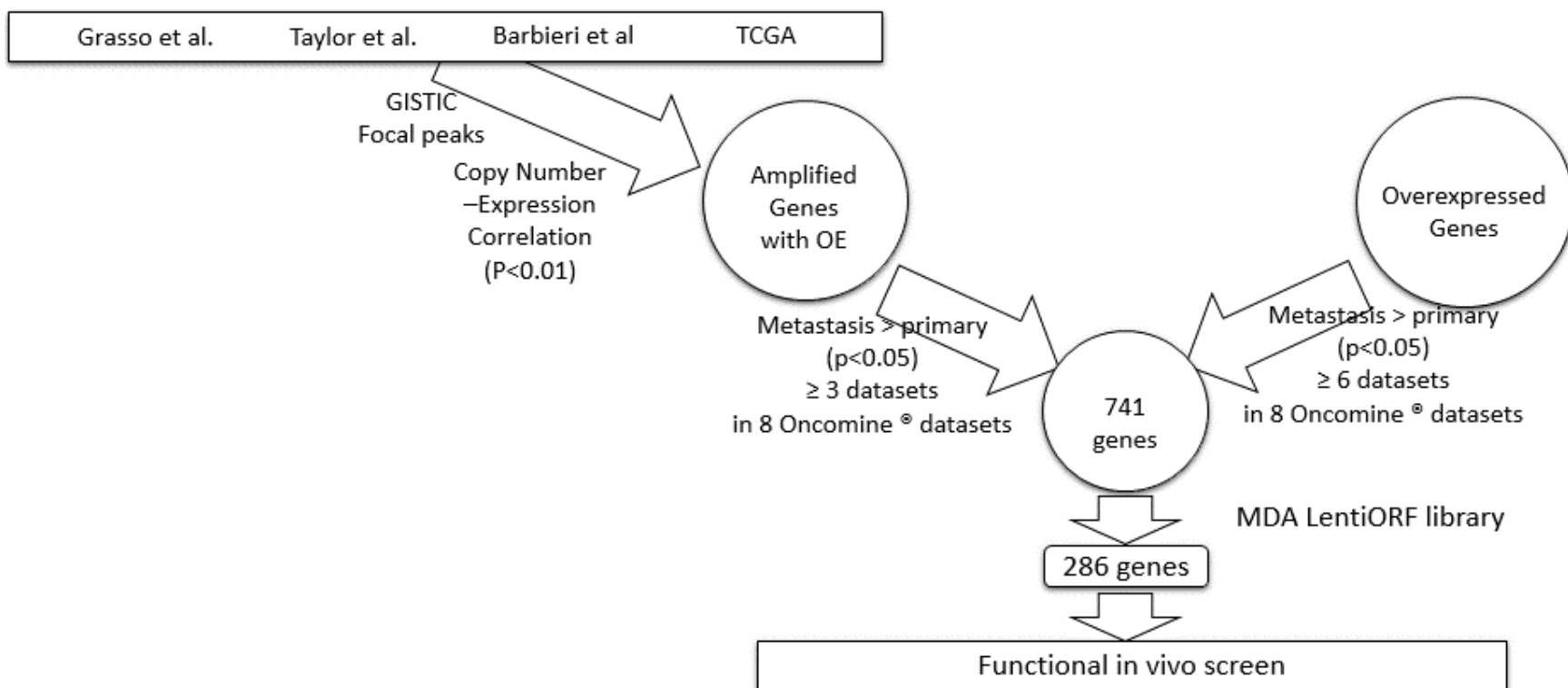


Figure 2.1 Bioinformatic genomic analysis to generate amplified/overexpression gene list for PCa progression

2.2.2 Cell culture

HEK293T cell were obtained from American Type Culture Collection (ATCC, Manassas, VA). LHMK and LHMK-AR cell lines were generous gifts from Dr. Bill Hahn's lab. Cells were cultured in DMEM (Live technologies) with 10% fetal bovine serum (FBS) (Live technologies) in a humidified incubator at 37°C in a 5% CO₂ atmosphere. BPH-1 were obtained from American Type Culture Collection (ATCC, Manassas, VA) and maintained in RPMI (Live technologies) with 10% FBS (Live technologies) in a humidified incubator at 37°C in a 5% CO₂ atmosphere.

2.2.3 Establishment of the library of blasticidin-selected stable cell lines overexpressing 286 candidate genes

The following work flow was used to establish the library of stable cell lines overexpressing putative driver genes of PCa. Glycerol stocks contained bacteria with desired open reading frames (ORFs) constructs were obtained from MD Anderson core. The ORFs were cloned by Gateway cloning to the pLOC vector. The pLOC vector contains blasticidin-selection cite and turboGFP as indicated in the Figure 2.11. Bacteria were then individually multiplied in 96-well-plate format, followed by midi-prep (Qiagen). The result pool was validated by sequencing 10 clones randomly. High throughout-put normalization to the same concentration was conducted by QiAgility (Qiagen) in 96 well plates. Virus packaging was achieved to target 286 genes individually. Virus was infected with seeded LHMK cells in 96-well plates. 24 hours post-infection, desired clones were obtained by blasticidin selection (10ug/mL) for 1 week to maintain the cells while they were expanded for 2 passages to acquire enough cell for *in vivo* injection.

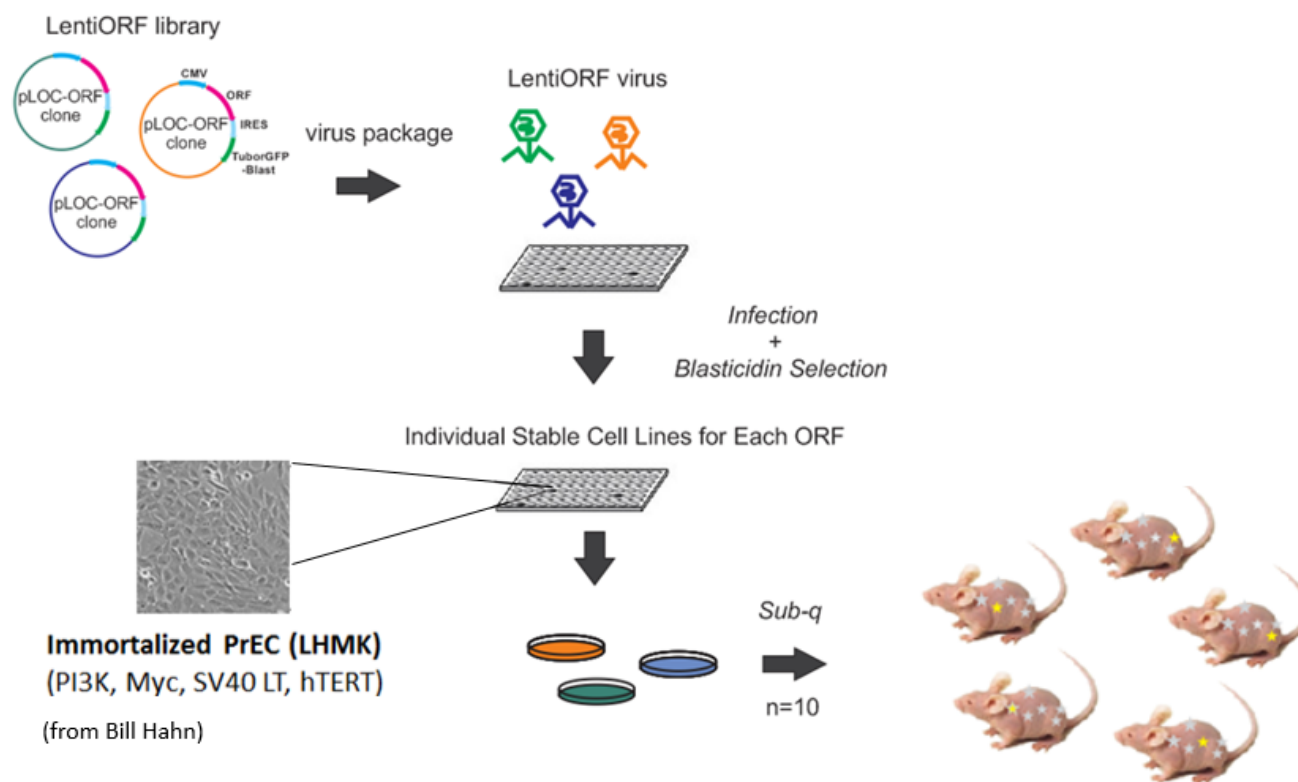


Figure 2.2 Workflow of *in vivo* ORF screen for amplified/overexpressed genes

Schematic representation of the approach. LHMK cells were infected with single ORFs and injected subcutaneously to allow tumor formation with 10 replications. The injection sites on left side of the mice were marked as stars shown in the figure.

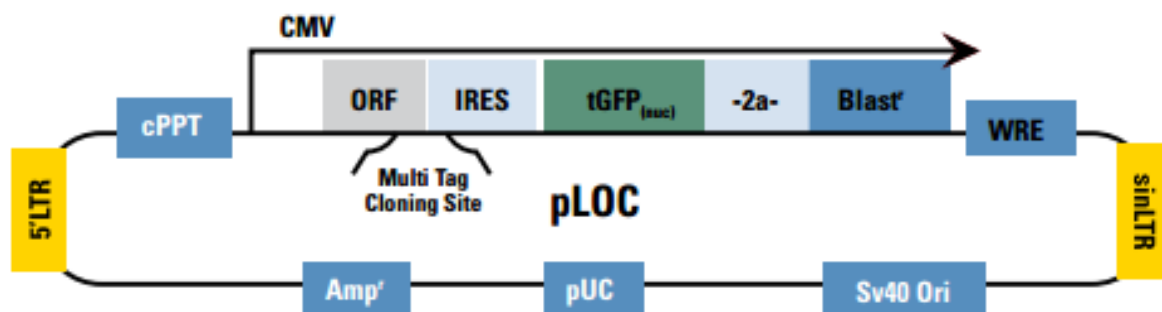


Figure 2.3 Vector structure for LentiORF library establishment (Taken From Thermo Fisher Scientific Inc., Technical Manual, Thermo Scientific Open Biosystems Precision LentiORF Collection)

2.2.4 Xenograft studies and *in vivo* ORF screen for putative genes promoting PCa

All animal experimental protocols were approved by the Institutional Animal Care and Use Committees at the University of Texas MD Anderson Cancer Center (Houston, TX) (IACUC number: 1169-RN01). Three prostate cell lines, LHMK, LHMK-AR and BPH-1, were tested for *in vivo* tumor formation by intradermal injection of 5×10^5 cells into the flanks of male nude mice (Taconic). LHMK was as parental lines because of the relatively long tumor latencies, which were seen no visible subcutaneous tumor formation in 6 months. 1×10^6 cells were injected subcutaneously into the 10 sites of each nude mouse. All the injected cells were diluted in PBS (Life technologies) and then mixed with BD Matrigel™ basement membrane matrix (BD Biosciences).

2.2.5 RNA extraction, cDNA synthesis and quantitative RT-PCR.

Total RNA was isolated from cells by RNeasy Mini kit (Qiagen). After normalization, total cDNA was synthesized by reverse transcription (RT) kit SuperScript III as manufacturer's protocol (12574-026, Invitrogen). Semi quantitative PCR was performed to detect the expression level of genes of interest, comparing level between negative control and the established cell lines by SYBR Green methods (Invitrogen). RPL30 were detected for an internal control for normalization purpose. Sequences of primers used for PCR reactions are in supplement material

2.2.6 Genomic sequencing

Plasmid DNA was prepared by QIAGEN Midiprep kit. Samples were sent to MD Anderson Sequencing Core for Sanger Sequencing. The result sequence was 'BLAST'ed by NCBI online tool compared with targeting ORF sequence.

2.2.7 Protein lysate and western blot assays

Protein lysate of PCa cell lines for the characterization of PYGO2 expression level was obtained from DePinho Lab including 22Rv1-TR,

Cells were lysed and sonicated in RIPA buffer (89901, Thermo) with protease inhibitor and phosphatase inhibitor cocktails (Roche). Bradford assay (Bio-rad) was applied for quantifying the amount of total protein. Lysates were fractionated by sodium dodecyl sulfate (SDS)-PAGE and transferred onto nitrocellulose membrane (Invitrogen). The membranes were blocked with 5% nonfat dried milk and then incubated with desired primary antibodies dilution in 5% bovine serum

albumin (overnight at 4°C. After applying second antibodies, peroxidase-conjugated anti-rabbit or anti-rabbit IgG, protein bands were detected by chemiluminescent detect system (Thermo).

Primary antibodies used in this project includes monoclonal anti-KRAS antibody (sc-30, Santa Cruz), polyclonal anti-PYGO2 (1:1000) (HPA023689, Sigma).

2.2.8 Proliferation analysis

LHMK cells with overexpression of gene of interest were plated on 24 wells plate and assessed of their confluence by IncuCyte® (Essen BioScience) for 7 days.

2.2.9 Soft agar analysis

2mL bottom agar mixture (DMEM + Glutamax with 1% FBS, 0.6% LE Agarose) (Lonza) was applied to coat a 6-well culture plate and solidified in 4°C for 1 hour. Cells were trypsinized, counted and diluted to 2×10^4 cells/mL in 2mL top agar mixture (DMEM + Glutamax with 1% FBS, 0.3% SeaPlaque® Agarose) (Lonza). Mixture was plate in the top agar and incubated with 2mL desired media at 37°C for 3 weeks. At endpoint, crystal violet was used to stain the plate. The colonies were quantified by counting 3 representative fields at 4X magnification.

2.2.10 Migration assay

Consisting of a 24-well companion plate with inserts containing 8µm pore size filters, Cell Culture Insert (BD Falcon) was used for the migration assay. Cells were first starved in DMEM with 1% FBS overnight and then seeded onto the

inserts in serum-free DMEM at a density of 5×10^5 cells/200 μ L. DMEM containing 10% FBS was placed in outer wells as a chemoattractant. After 12 hours of incubation, none-migrating cells were removed with cotton swabs while migrated cells were fixed with methanol and stained with 0.2% crystal violet containing 2% ethanol. Migrated cells on the entire chamber were observed and counted under microscope.

2.2.11 Matrigel invasion assay

Matrigel invasion assay were performed applying 8 μ L pore size BioCoat Matrigel Invasion Chamber (BD Falcon) as protocol. Starved with DMEM with 0.1% FBS overnight, 5×10^5 cells were prepared in serum-free 500 μ L DMEM and then added into the chambers. DMEM containing 10% FBS was provided in outer wells as chemoattractant. After 18 hours of incubation, none-migrating cells were removed with cotton swabs while migrating cells were fixed with methanol and stained with 0.2% crystal violet containing 2% ethanol. Invading cells on the entire membrane were observed and counted under light microscope.

2.2.12 Tissue specimen, histology and immunohistochemistry

Tissue microarray was ordered at US Biomax, Inc as catalog number PR803b, the clinical information was listed available online. Patient-derived xenograft (PDX) models were generous gifts from Dr. Nora Navone. Transgenic mice were generated from DePinho ab. After fixed in 10% formalin overnighted, tissues were embedded in paraffin. The primary antibody used in IHC was rabbit

anti-PYGO2 ordered from Sigma Aldrich (HPA-023689), which were also used in western blot.

2.2.13 Expression profiling

RNA was isolated as previous described and sent for profiling at MD Anderson Microarray Core facility by the Human Genome U133 plus 2.0 Array (Affymetrix).

2.2.14 Bioinformatic study and statistical analysis

Gene Set Enrichment Analysis (Subramanian *et al.*, 2005) was conducted as User Guide. As recommendation, FDR (False Discovery Rate) $q\text{-value} < 0.025$ was set as a cutoff for appropriate indicator of gene enrichment. Gene sets collection from Kyoto Encyclopedia of Genes and Genomes (KEGG) were included in the analysis. Statistical analyses were performed by t test and one-way ANOVA using GraphPad Prism 4. In all experiments with error bars, Standard Deviation Ellipse (SDE) was calculated to indicate the variation within all replicates.

2.3 Results

2.3.1 Identification of putative genes contributing to tumor progression

741 putative genes were identified as candidates genes associated with PCa progression by integrative comparative genomic analysis. Gain of function screening experiments were performed using the 286 available expression vectors carrying ORFs derived from the above-mentioned list. In order to establish a screening system which has low background of tumorigenesis in immunodeficient

mice thus facilitating the detection of emergence of tumors formed by ORF overexpression, we chose the cell line LHMK. LHMK was derived from a normal human prostate epithelial cell line that was transformed with SV40 large T antigen, hTERT, PI3K and c-Myc (Berger et al., 2004a). LHMK cells were selected as the screen model system because the subcutaneous tumor latency for this cell line in nude mice is more than 200 days. Thus it is expected the tumors formed by overexpression of any particular ORF would highly likely indicate a real biological consequence from a potent pro-tumor gene.

Next, 286 ORFs were screened for their *in vivo* tumor promoting activity. Individual lentivirus harboring each of the 286 ORFs was prepared and used to infect LHMK cells followed by blasticidin selection for a week. Then, 286 stable sublines were generated and maintained to enough number for subcutaneous injection (injection sites N=10) (Figure 2.2). The approach of pooling viruses into multiple pools and infecting cells followed by injection was not used in our methods. Our approach is superior in terms of avoidance of the caveat of tumor phenotype emergence due to more than one gene in the same pool and eliminating the need for further sub-screen to identify which gene(s) of a pool leads to tumor phenotype. I followed tumor growth using caliper measurement and monitored for 226 days before all mice without tumors were euthanized. In total, 41 ORFs were identified (Table 2.1).

As shown in table 2.1, the control line with RFP infection showed no tumor growth from 30 injection sites over the observation period. Among the 41 hits,

some genes with previously characterized pro-tumor functions in PCa were identified (e.g. KRAS, FGFR1, CCNE2, SDC1, EZH2 and AURKA). This result supported the validity of the screen system. The *in vivo* growth patterns of ORF-mediated tumor promotion varied and ranged from 23-226 days (Figure 2.4). Because of practicality, I only validated some but not all of the overexpression at RNA levels. Among all 10 tested ORFs, their overexpression were confirmed in the sublines (Figure 2.5).

No.	Replicates	Average tumor initiation Time (Day)	Gene name
0	0 in 30	N/A	(LHMK-RFP)
1	20 in 20	<23	KRAS
2	1 in 10	23	RBM19
3	4 in 10	35	MOS
4	3 in 10	44	FGFR1
5	1 in 10	46	SLC45A4
6	1 in 10	46	CDC20
7	1 in 10	52	BOP1
8	1 in 10	60	NCBP2
9	2 in 10	64	CCNE2
10	1 in 10	64	SDC1
11	3 in 10	74	TOMM40L
12	1 in 10	75	PPOX
13	1 in 10	76	ARL6IP1
14	2 in 10	85	EZH2
15	1 in 10	85	ST3GAL1
16	2 in 10	86	PYGO2
17	1 in 10	86	IRF5
18	1 in 10	86	CDCA4
19	1 in 10	86	WDYHV1
20	1 in 10	95	TTC35
21	1 in 10	111	OGFR
22	2 in 10	120	TAF6
23	1 in 10	131	CPSF4
24	1 in 10	152	ATP5G1
25	5 in 20	153	RPS20
26	2 in 10	154	MTBP
27	1 in 10	154	KRTCAP2
28	1 in 10	154	AURKA
29	3 in 10	164	ZKSCAN5
30	2 in 10	167	TROAP
31	1 in 10	187	POLR2H
32	1 in 10	187	SPAG1
33	1 in 10	187	WDR53
34	1 in 10	187	DERL1
35	1 in 10	188	ZNF706
36	1 in 10	189	NUSAP1
37	1 in 10	205	ESRRA
38	1 in 10	218	NASP
39	1 in 10	218	MRPL28

40	1 in 10	216	RRM1
41	1 in 10	226	MCM7

Table 2.1 List of ORFs facilitating LHMK xenograft growth in injected mice in 6 months

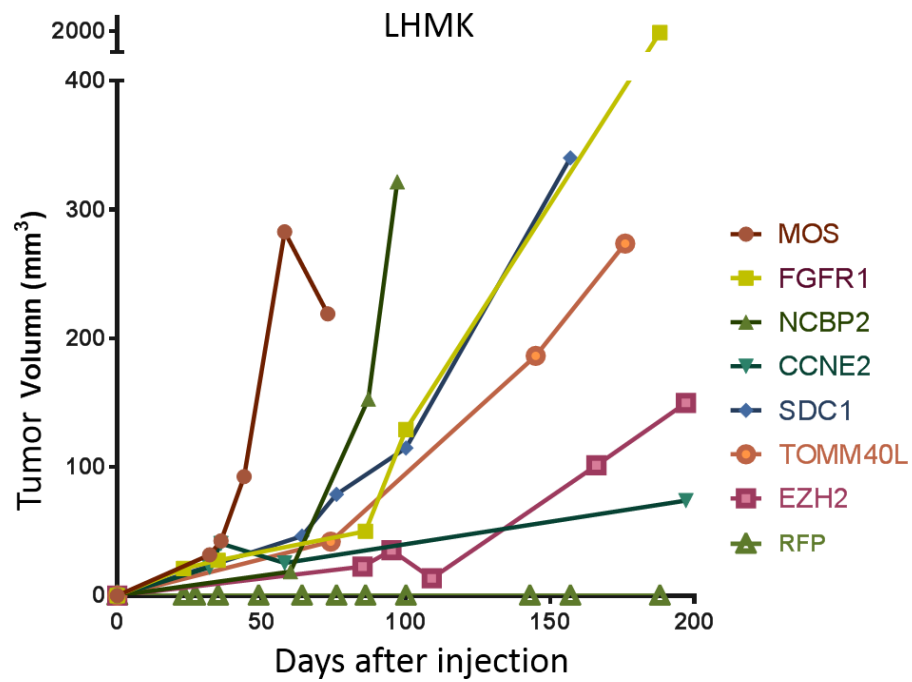


Figure 2.4 Representative hits growth of ORFs

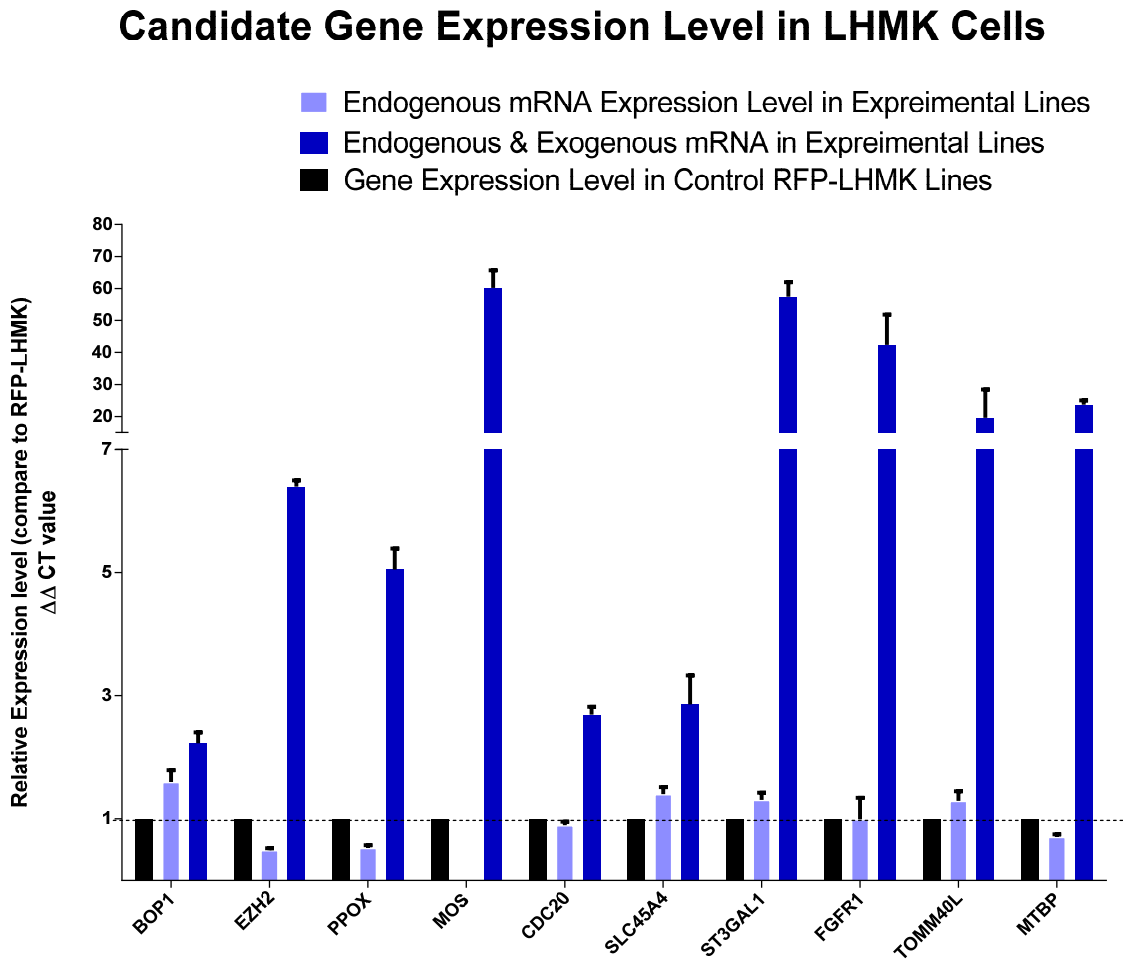


Figure 2.5 Representative validation of the expression level of interested genes in injected LHMK lines

Endogenous and exogenous mRNA expression level were validated in LHMK cell lines by different primers by qRT-PCR. All expression level were normalized to $\Delta\Delta CT$ to represent the relative expression level comparing with RFP-LHMK control. All experiments repeated 3 times. All error bars are SDM.

The penetrance of the hits indicate their potency to promote the tumor progression of LHMK cells. To prioritize hits for further functional analysis, 12 hits with at least 2 tumor incidences in 10 injection sites were selected (

Table 2.2). To discover new putative genes promoting PCa, the genes which has been characterized as pro-tumorous genes in PCa (i.e. KRAS, FGFR1, CCNE2, and EZH2) will not be validated in all of the following *in vitro* assays but served as positive controls in some of the assays.

Replicates	Average tumor initiation time (Day)	Gene Name
0 in 30	N/A	(LHMK-RFP)
20 in 20	<23	KRAS
4 in 10	35	MOS
3 in 10	44	FGFR1
2 in 10	64	CCNE2
3 in 10	74	TOMM40L
2 in 10	85	EZH2
2 in 10	86	PYGO2
2 in 10	120	TAF6
5 in 20	153	RPS20
2 in 10	154	MTBP
3 in 10	164	ZKSCAN5
2 in 10	167	TROAP

Table 2.2 List of ORFs that promoted multiple tumor growth

2.3.2 Functional validation using cell proliferation assay

It is possible that some or all of the 12 hits promote tumor growth *in vivo* by enhancing cell division and proliferation. To test this hypothesis, I performed 2D growth curve assay using IncuCyte®. However, in preliminary experiments of 9 hits I observed limited consistent differentiation of the curves when ORFs sublines were compared with RFP control line (Figure 2.6). I reason that the lack of observed *in vitro* growth advantage by ORF overexpression was partly caused by the already very short doubling time of LHMK (~8h).

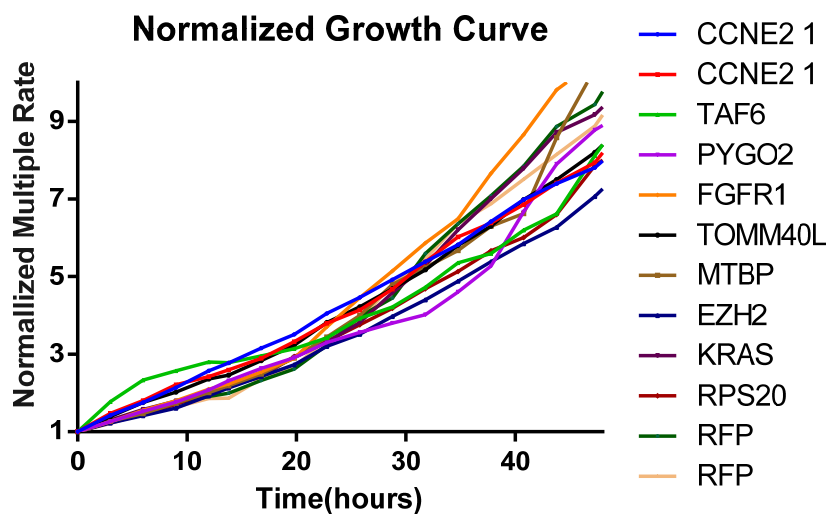


Figure 2.6 2D proliferation assays showed similar growth rates among LHMK cells overexpressing ORFs

Growth curve of top hits were obtained by IncuCyte®. Confluence of each lines was recorded in multiple time points. All the result was normalized with control lines.

2.3.3 Genes promoting anchorage-independent growth

Since I was not able to obtain proliferation differences in 2D culture, I resorted to another approach, the soft agar colony formation assay, which is routinely used to explore the anchorage independent growth and often delivers results that are more reminiscent of the *in vivo* growth properties of tumor cells. While the control LHMK line was unable to form colonies, several tested hit genes including PYGO2 showed surprisingly robust colony formation in LHMK cells in 10 days (Figure 2.7). It is interesting to note that mechanisms for promoting anchorage independent growth by different genes may be through separate pathways, as colonies formed by different ORFs displayed different sizes, for example, LHMK-BOP1 and LHMK-MOS exhibited larger colony size compared with those by LHMK-PYGO2.

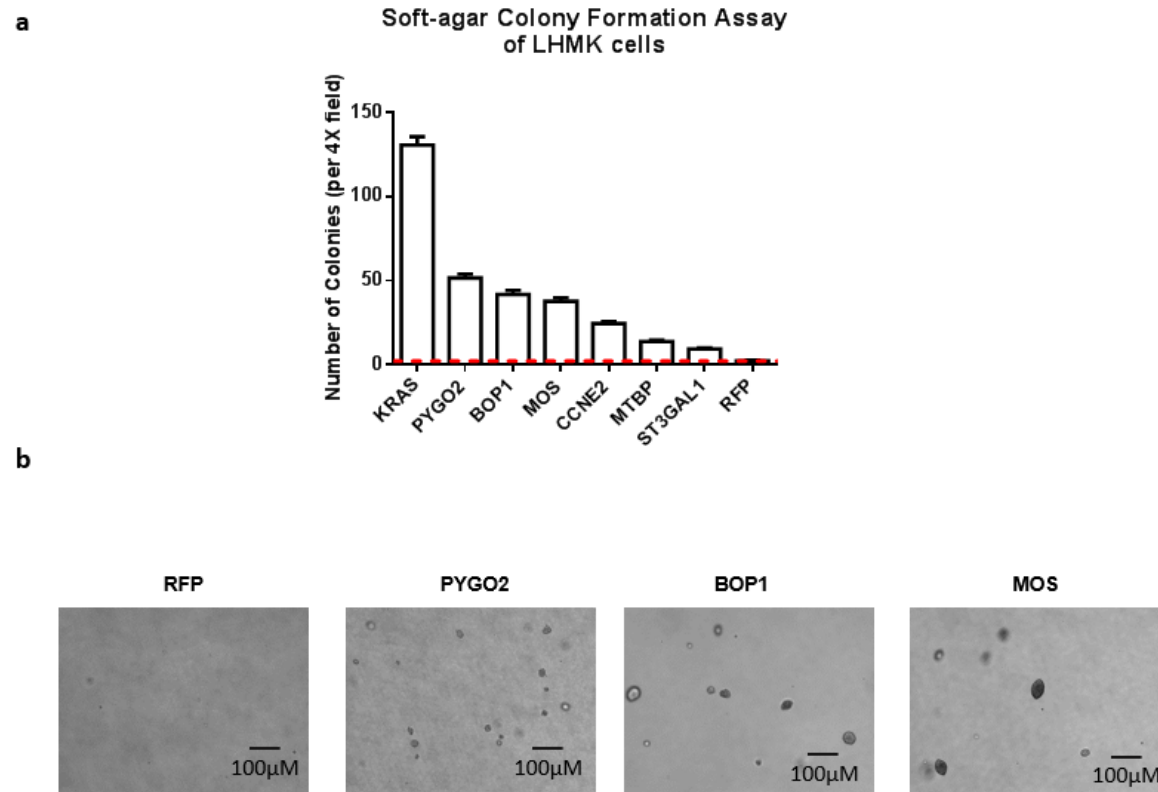


Figure 2.7 ORFs facilitated anchorage-independent growth of LHMK cells

a) Overexpression of top hit genes affects anchorage-independent growth in LHMK. Results of soft-agar colony formation assays of LHMK sublines. All experiments repeated 3 times. All error bars are SDM. b) Representative cell colonies in soft agar are shown.

2.3.4 Functional validation by migration/invasion assay

Increased mobility and invasiveness are important characteristics of malignant tumor cells. In our initial ORF list, the genes are all upregulated in metastasis at RNA level across several transcriptome datasets. Therefore, I reason that it is possible that one or more of the genes in the hit list may promote migration and invasion, properties that could be examined with in vitro Boyden chamber assays. Several genes, such as BOP1, MTBP, PYGO2, ST3GAL1 and TROAP, showed significant enhancement of migration and invasion in LHMK (Figure 2.8). Although MOS strongly promoted soft agar growth, its effect on migration and invasion was lowest in assessed genes.

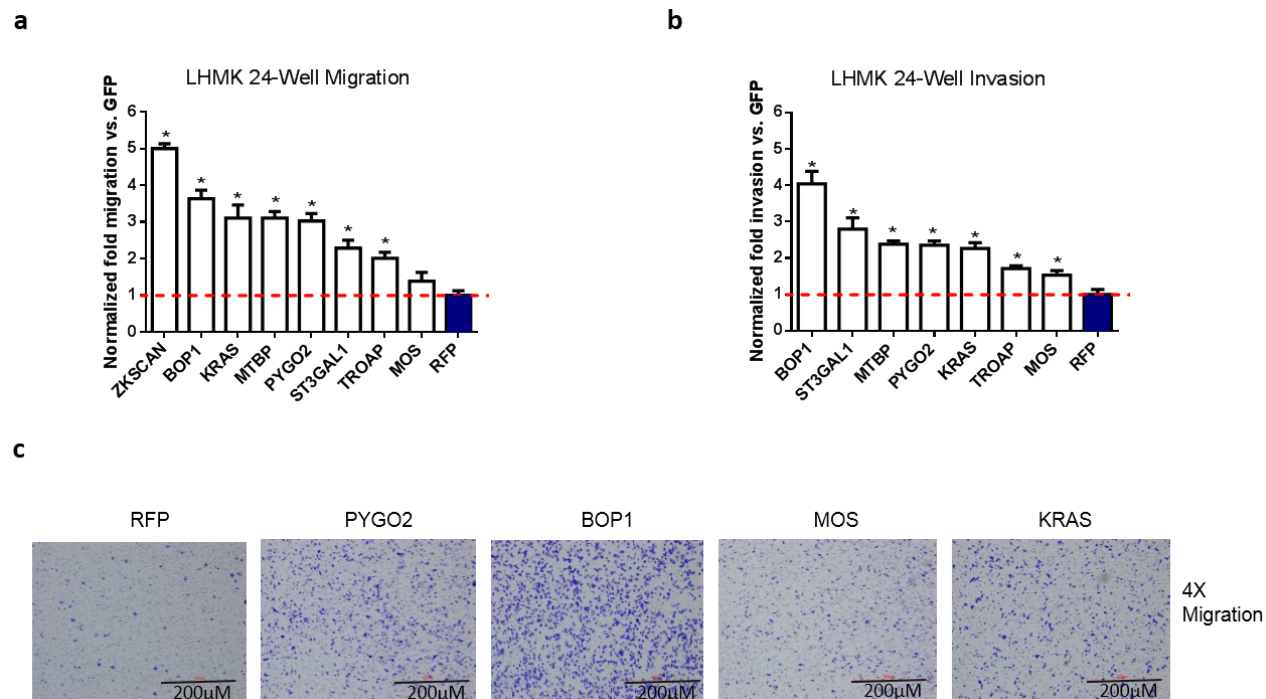


Figure 2.8 Hits showed promoting LHMK *in vitro* migration (a) and invasion (b) in Boyden chamber assay.

a & b) LHMK cells overexpressed with hit genes seeded in Boyden chamber were stimulated with FBS. RFP controls are highlighted in blue. Experiments were repeated for two independent experiments with three replicates for each. *, $P < 0.05$, student t test. All error bars are SDM. c) Migrated or invaded cells were analyzed by light microscope (magnification, 4X).

2.3.5 Alteration of MAPK pathway activity

Due to the 100% penetrance of KRAS as being the screen hit, it is of interest to assess the ability of the above-tested genes to activate MAPK pathway, the prototypical downstream signaling pathway activated by KRAS. Using western blot, both KRAS and MOS overexpression induced more pronounced phospho-MEK signal, yet both seemed to suppress phospho-AKT signals (Figure 2.9). Given MOS being identified as a serine/threonine kinase that activates the MAP kinase cascade through direct phosphorylation of the MAP kinase activator MEK (Prasad *et al.*, 2008), this result was expected.

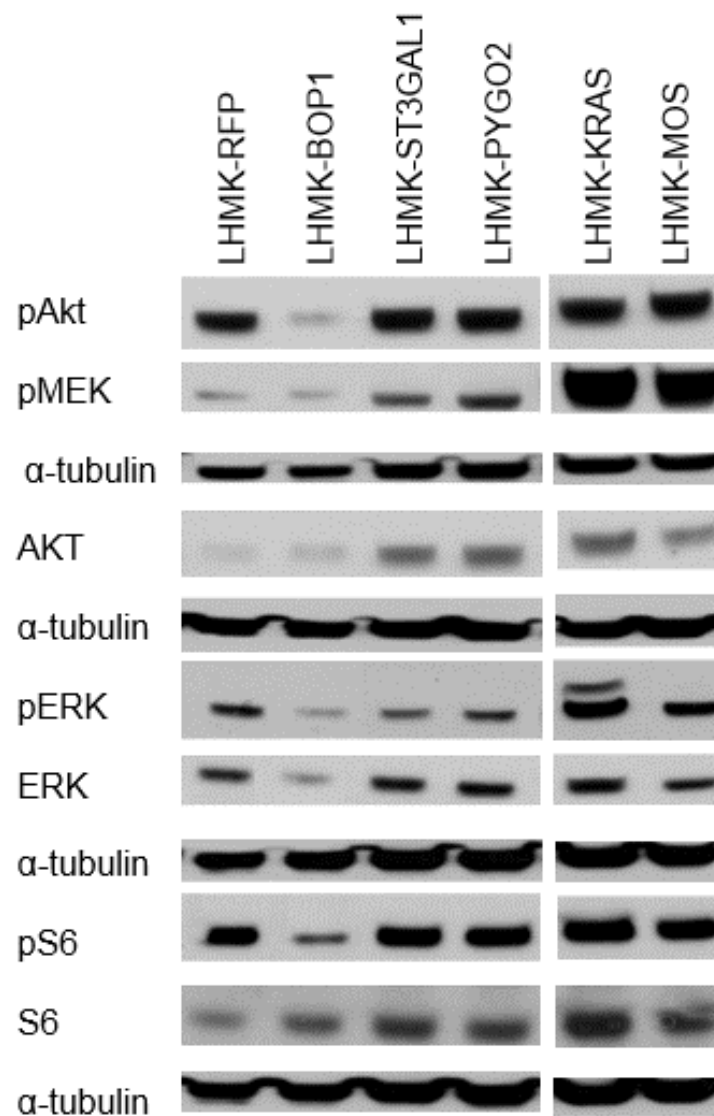


Figure 2.9 Immunoblotting analysis of various cellular signaling proteins in cell lysates of LHMK cells overexpressing ORFs

2.3.6 PYGO2 up-regulation was correlated with PCa pathological aggressiveness.

The expression of PYGO2 in normal prostate tissue and PCa was investigated by IHC using LHMK-RFP and LHMK-PYGO2 cell lines as negative and positive controls, respectively. PYGO2 staining was clearly localized in the nuclei. Notably, the expression of PYGO2 was significantly upregulated in adenocarcinoma comparing to that of normal prostate tissue (Table 2.3). Further, PYGO2 expression levels in adenocarcinoma patients and their correlation with clinicopathological factors were studied. As shown in Figure 2.10, PYGO2 staining intensity in prostate tissues was classified in four groups with scores ranging from no staining (score 0) to intense staining (score 3). Each sample was scored twice with clinical information blind.

	PYGO2 Expression	
	-	+
Normal	4	0
Adenocarcinoma	13	54

Table 2.3 PYGO2 up-regulation in PCa patient samples (n=71)

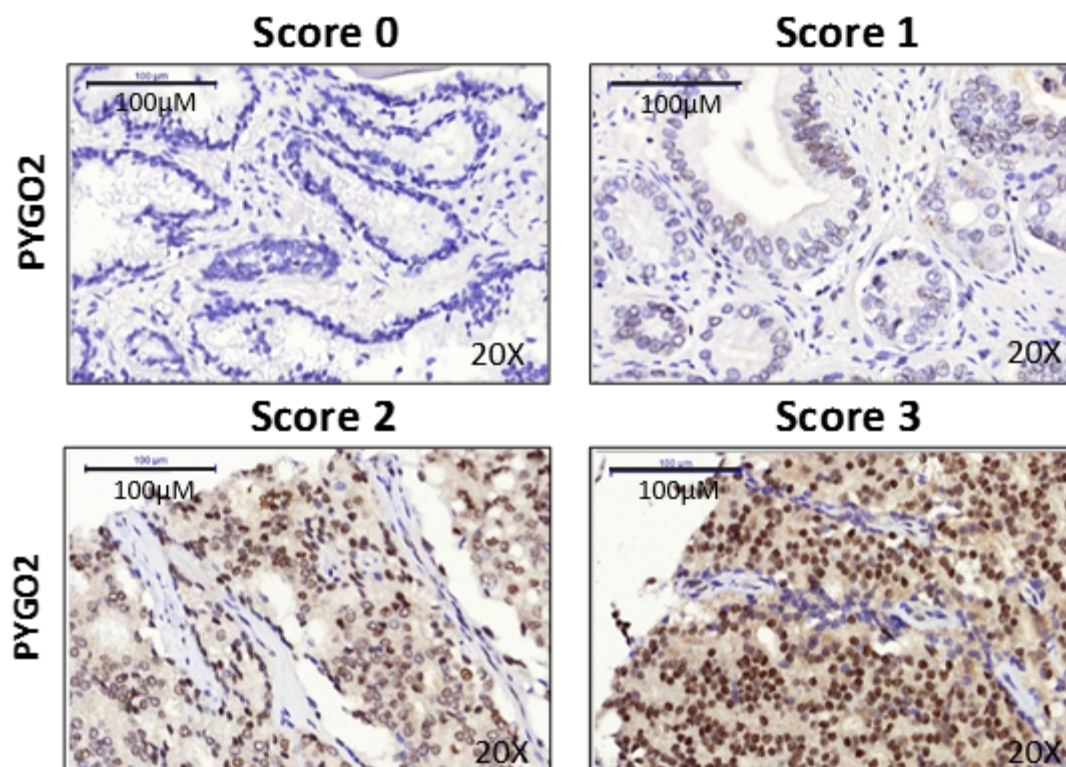


Figure 2.10 Expression intensity of PYGO2 in PCa patient samples from low to high (0-4).

		PYGO2 expression level IHC intensity score				<i>p</i> value
		0 (n=13)	1 (n=24)	2 (n=15)	3 (n=15)	
Stage	I+II	6	15	5	6	0.2937
	III+IV	7	9	10	9	
T-stage	1+2	7	17	6	9	0.2911
	3+4	6	7	9	6	
M-stage	M0	10	19	8	9	0.2834
	M1	3	5	7	6	
Gleason Score	6-7	5	8	6	0	0.0274*
	8-9	4	10	6	4	
	10	4	6	3	11	

Table 2.4 PYGO2 up-regulation is associated with high Gleason Score (n=71)

As the result shown in Table 2.4, higher PYGO2 expression was significantly associated with higher Gleason Score ($p=0.0274$). However, PYGO2 expression was not correlated with tumor staging. PYGO2 expression pattern is being further characterized with more samples recapitulating the heterogeneity and multi-stage of PCa, including samples of metastatic sites, from MD Anderson pathology bank.

Patient-derived xenograft (PDX) model has been increasingly used as a preclinical platform due to its more close histopathological and genetic resemblance to human disease. When IHC of PYGO2 was performed in four PCa PDX models that we obtained from the PDX core run by Dr. Nora Navone at MD Anderson, PYGO2 was intensely stained in the nuclei of adenocarcinoma tissues in all models (Figure 2.11).

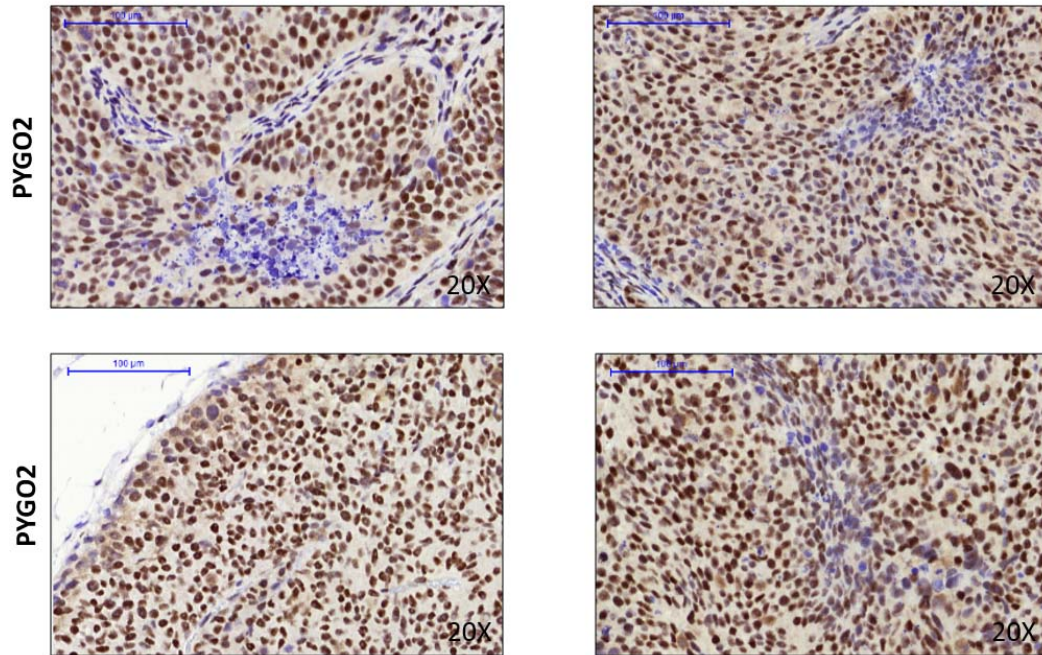


Figure 2.11 Representative PDX samples showed strong positive staining of PYGO2

2.3.7 PYGO2 promotes LNCaP local invasion to draining lymph nodes

To further characterize PYGO2 function *in vivo*, PYGO2 was overexpressed using a lentiviral construct with constitutive GFP marker in LNCaP (Figure 2.12). LNCaP, as an androgen-responsive PCa cell line, was established from a metastatic lesion of human prostatic adenocarcinoma (Horoszewicz *et al.*, 1983). PYGO2 overexpression boosted the *in vivo* tumor progression modestly as determined by tumor weight at end point (Figure 2.12). Interestingly, while metastasis was not observed in parental LNCaP cells, three of the five PYGO2 overexpressing LNCaP lines developed local invasion to draining lymph nodes (Figure 2.13) (Table 2.5).

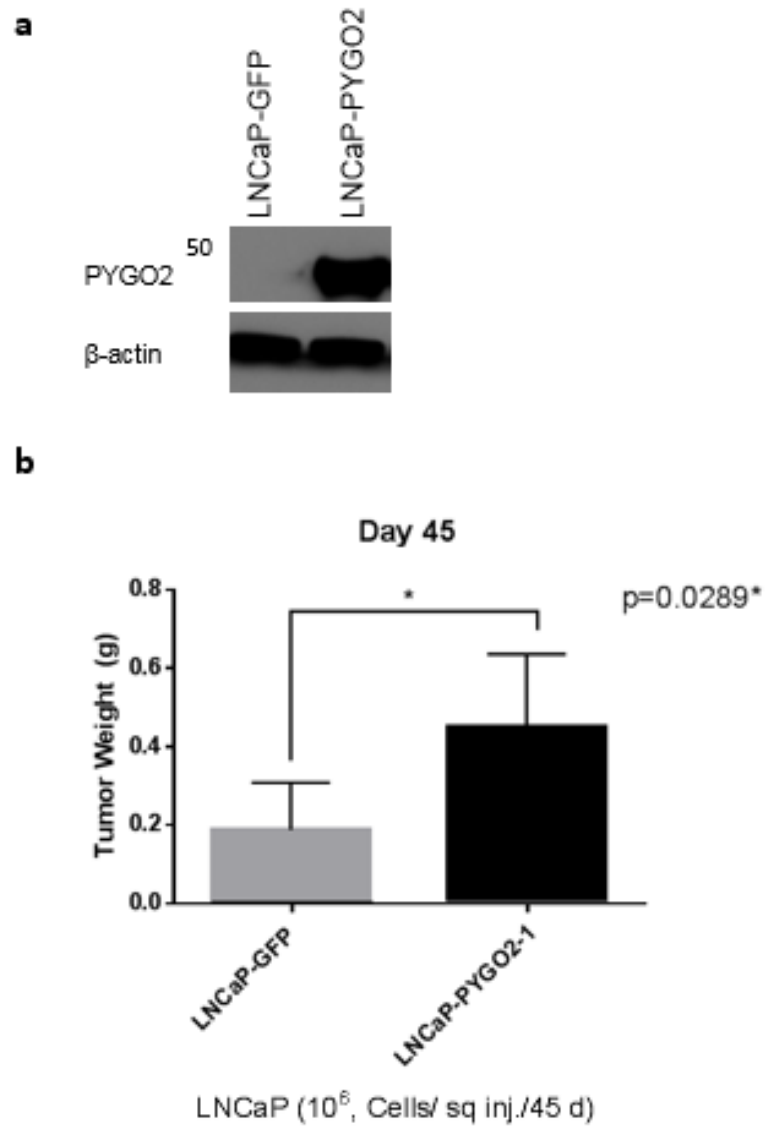


Figure 2.12 PYGO2 enhanced LNCaP subcutaneous tumor progression *in vivo*

a) Generation of PYGO2 overexpressing lines in LNCaP; b) Tumor weight increased in LNCaP overexpressing PYGO2

Line	Mouse	Mets
LNCaP-Original	1	no metastasis
	2	no primary tumor
	3	no metastasis
	4	no metastasis
	5	no metastasis
LNCaP-PYGO2	1	lymph nodes met
	2	lymph nodes met
	3	no metastasis
	4	no metastasis
	5	lymph nodes met

Table 2.5 High penetrance of LNCaP local invasion to draining lymph nodes
in PYGO2 overexpressed cells

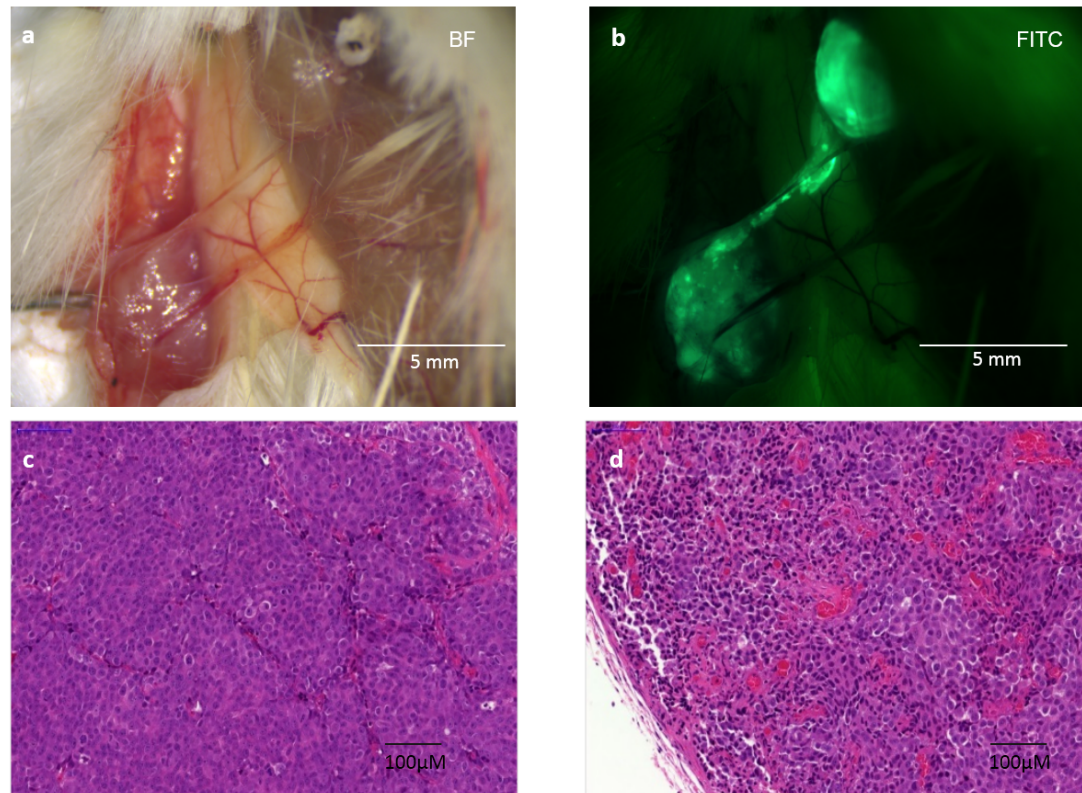


Figure 2.13 PYGO2 promotes LNCaP local invasion to draining lymph nodes

a & b) local s. q. tumor and draining lymph nodes were visualized by bright filed microscopy (BF) and FITC; c) histology of primary subcutaneous tumor by HE staining; d) draining lymph nodes were largely overtaken by invading tumor cells.

2.3.8 Elevated expression of Pygo2 in Wnt-upregulated genetic engineered mouse models (GEMMs) and in PCa patients.

To investigate which signaling pathway is enriched in clinical samples with high PYGO2 expression, I collaborate with Dr. Amin Samirkumar to perform Gene Set Enrichment Analysis (Subramanian *et al.*, 2005). From at least one publically available large-scale transcriptome datasets with both local and metastasis PCa samples (Grasso *et al.*, 2012), we observed significant statistical enrichment for Wnt signature (<http://www.genome.jp/kegg/pathway/hsa/hsa04310.html>) in both localized (FDR q-value < 0.05064655) and metastatic (FDR q-value < 0.0689247) PCa (Figure 2.14). This result suggests that the function of PYGO2 might be directly associated with Wnt pathway in PCa patients.

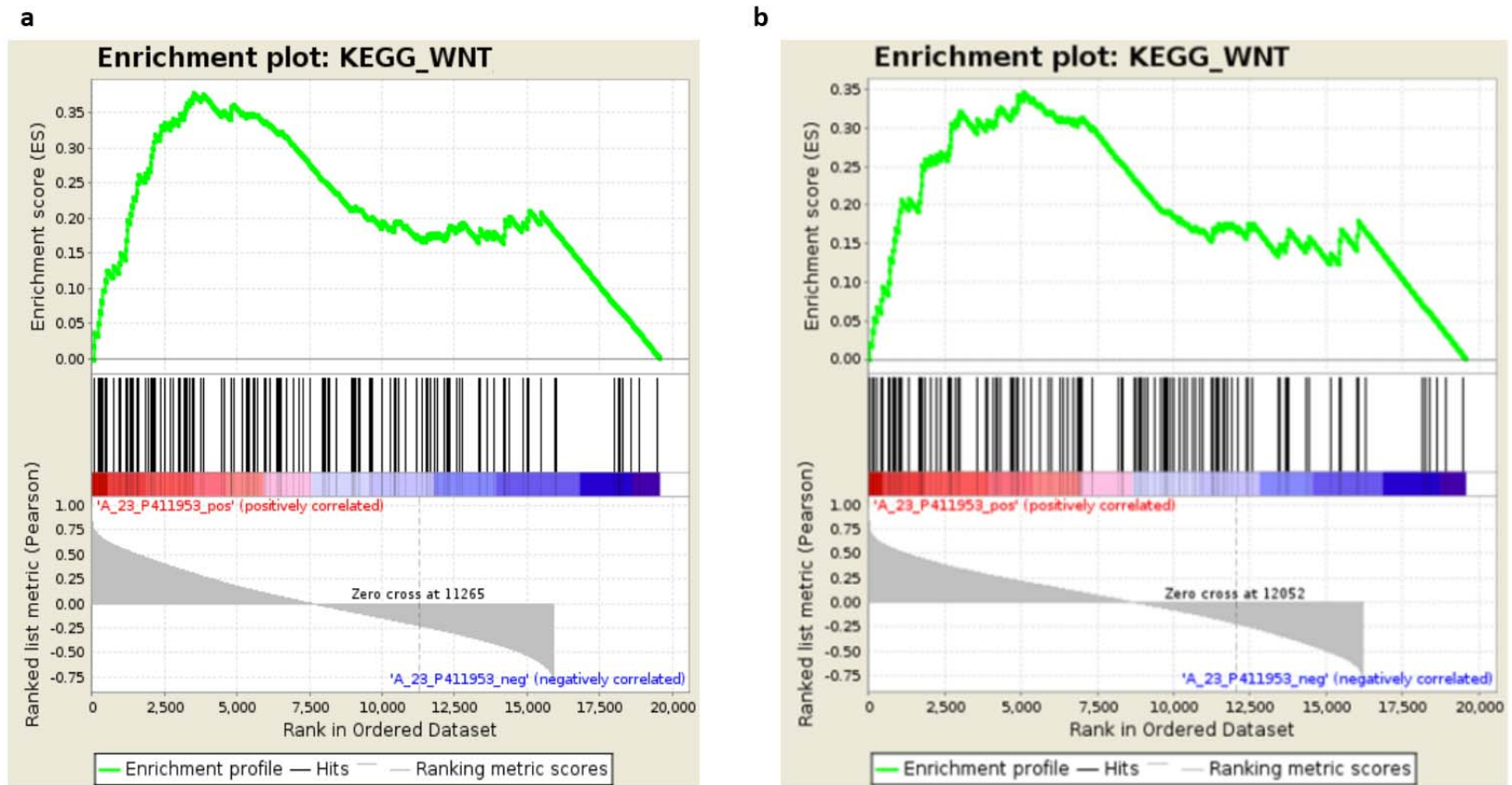


Figure 2.14 Wnt pathway enrichment in local (a) and metastasis (b) PCa samples with higher PYGO2 expression level at RNA level

Knowing that PYGO2 may be related to Wnt pathway in PCa, I took advantage of the several GEMM models, including one with prostate-specific *APC* loss, established by Dr. Xin Lu in the lab and evaluated PYGO2 expression pattern (Figure 2.15). The expression levels of Pygo2 were elevated in the prostate of *PB-Cre⁺ PTEN^{L/L}* mice comparing with wild type, and markedly enhanced in tumors from *PB-Cre⁺ PTEN^{L/L} APC^{L/L}* mice where there is an expected upregulation of Wnt signaling resulted from the deletion of *APC* gene. This result suggests that PYGO2 upregulation may be downstream of Wnt signaling, a hypothesis that is being explored in depth by Dr. Lu in the lab.

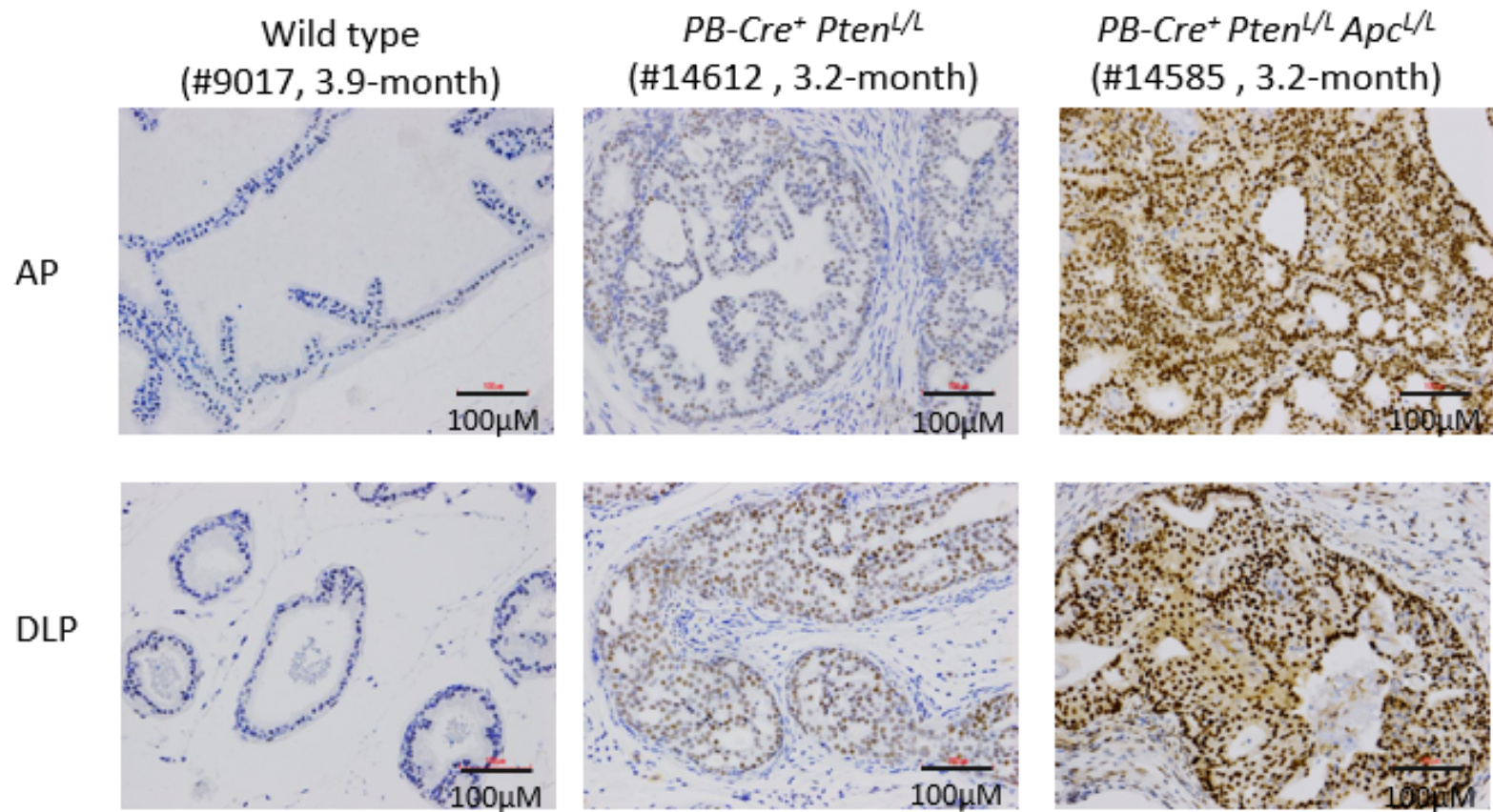


Figure 2.15 *Pygo2* IHC on anterior and dorsolateral prostate tissue from wild type, *PB-Cre⁺ PTEN^{L/L}* and *PB-Cre⁺ PTEN^{L/L} APC^{L/L}* mice

Transcriptome profiling identified differential gene expression in PYGO2 overexpressed LHMK (Figure 2.16). The cutoff was set for a fold change of greater than 2.5 times in either direction as shown in Table 2.6. Some of the genes in the list were selected for expression validation based on their potential functions in cancer after literature study, such as WNT2, KISS1, ADAMTS2, and IGFBP3. Their expression levels in LHMK-RFP and LHMK-PYGO2 cells were validated (Figure 2.17). In LHMK-PYGO2, the expression levels of WNT2, ADAMTS2 and IGFBP3 increased as expected while KISS1 was down-regulated compared to those in LHMK-RFP.

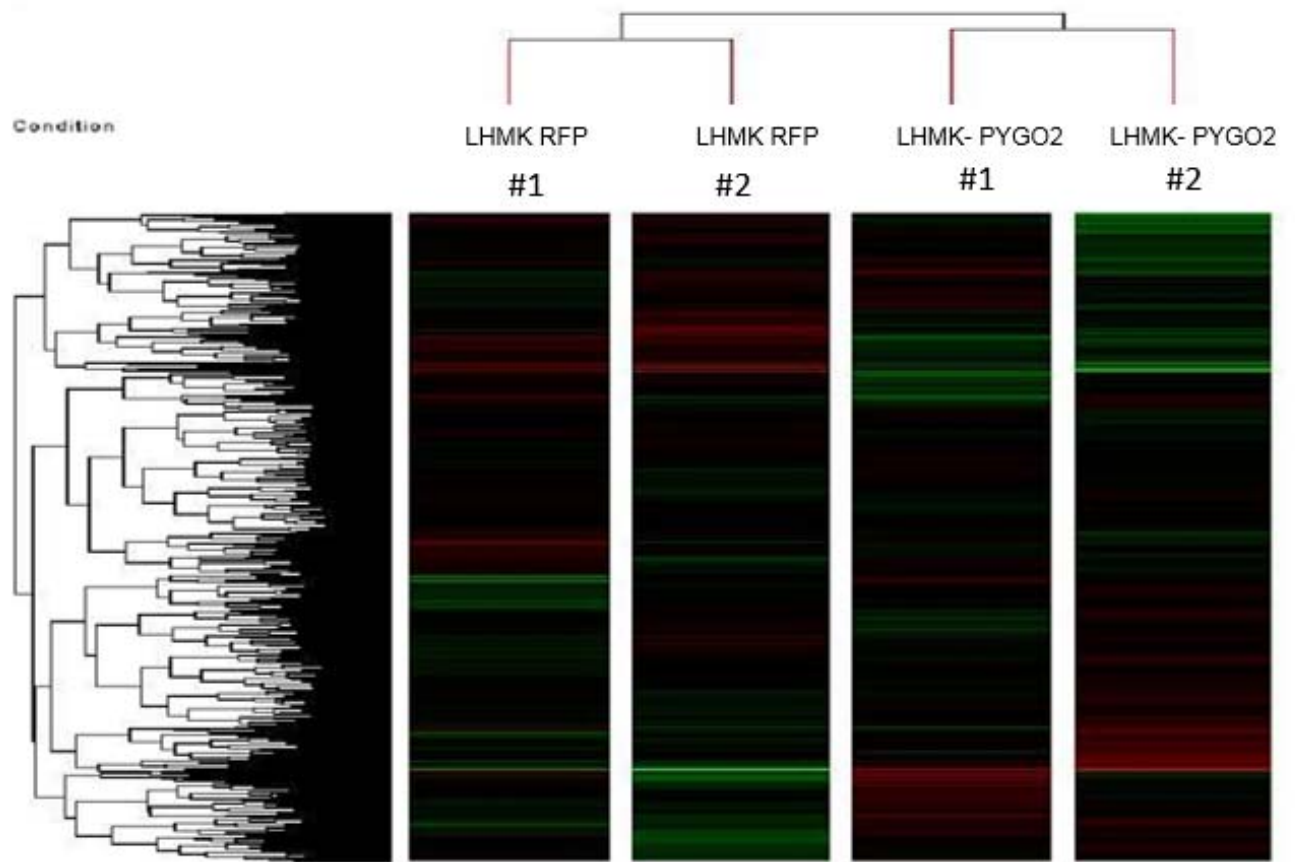


Figure 2.16 Microarray analysis of LHMK and overexpression cells

Fold Change	Regulation ([LHMK-PYGO2] vs [LHMK-RFP])	Gene Symbol
6.62884	Up	CNR1
4.9924645	Up	COL21A1
4.917648	Up	ANK3
4.568018	Up	ZNF804A
4.339982	Up	SPOCK3
4.2134004	Up	WNT2
3.9524667	Up	JPH1
3.8303087	Up	RARRES1
3.766973	Up	PCDH10
3.7488606	Up	RARRES1
3.7367597	Up	ADAMTS2
3.6632433	Up	EBF1
3.5750515	Up	HS3ST3B1
3.5061955	Up	EBF1
3.4354672	Up	LDLRAD4
3.3923466	Up	EBF1
3.3704123	Up	TNFSF10
3.3625364	Up	ISL1
3.3579595	Up	AHNAK2
3.240252	Up	FRAS1
3.2399745	Up	COL4A4
3.2212367	Up	ARHGAP24
3.1589348	Up	ANK3
3.13414	Up	LXN
3.0961728	Up	HMCN1
3.0851064	Up	PDE8B
3.048719	Up	PCDH7
3.0126672	Up	GPC6
2.9809773	Up	PLCB1
2.9315355	up	ZFHX4-AS1
2.9252326	up	MAN1C1
2.9163897	up	WISP2
2.9149344	up	GRIK2
2.9112802	up	COL3A1
2.868886	up	TNFSF10
2.8234982	up	NLGN4Y
2.8198066	up	TNFSF10
2.8013465	up	TOX
2.7908678	up	IGFBP3

2.7850227	up	FLRT2///LOC100506718
2.7831643	up	SPOCK3
2.7190666	up	C6orf141
2.71883	up	ITGA4
2.7138515	up	SYNPO
2.7098184	up	PCDH7
2.6812658	up	KAL1
2.675754	up	DCN
2.6699352	up	RCAN2
2.6692655	up	ITGA1
2.6661928	up	DNER
2.6522381	up	DCN
2.6519594	up	ZNF503
2.6400201	up	PARM1
2.634583	up	SLC44A5
2.6236303	up	NFIA
2.6227007	up	SLC4A4
2.606458	up	FLRT2///LOC100506718
2.6019416	up	PELI2
2.5822084	up	DHRS3
2.5606158	up	LMO4
2.5561075	up	CASP1
2.5494325	up	CLEC3B///EXOSC7
2.5493717	up	PBX1
2.5459576	up	TSHZ1
2.5437036	up	COL3A1
2.5408363	up	IL7
2.5340688	up	MAN1C1
2.5280597	up	FOXA1
2.5186205	up	LRRC8C
2.5179226	up	CASP1
2.5177903	up	MEGF6
2.511516	up	DKK2
2.5098848	up	GPRC5B
2.509565	up	RARB
2.508565	up	LOC100505946
2.5054262	up	CASP1
2.5045822	up	EFNB2
2.504143	up	IFITM1
2.5226448	down	DUSP4
2.5354156	down	REXO2
2.553761	down	NCAM1

2.55806	down	CLMP
2.5622652	down	USP28
2.5985959	down	MIR612///NEAT1
2.625526	down	USP28
2.6316588	down	MGST1
2.6578946	down	ZNF883
2.6836455	down	ATF7IP2///LOC100287628
2.691571	down	MGST1
2.723695	down	TMEM47
2.7501042	down	RARRES2
2.7749934	down	KISS1
2.7972617	down	POU3F2
2.8331623	down	VLDLR
2.9318223	down	DSC2
2.9504306	down	PAX6
2.9601073	down	LOC100506303///LOC100653149///LOC101060483///LOC40
2.9910436	down	OTTHUMG00000175814///RP11-13L2.4
3.0252461	down	PDZK1
3.2035272	down	ND6
3.249426	down	GAL
3.2839582	down	DCDC2
3.2858608	down	MGST1
3.382219	down	MAP7D2
3.4408128	down	NRXN3
3.7098405	down	CADM1
3.7897563	down	DSC2
3.8596961	down	NRXN3
3.9882278	down	PAX6
4.4886146	down	SYTL5
4.5460978	down	CNN1
4.622783	down	CNKSR2
4.7903833	down	NRXN3
5.27109	down	CADM1
5.482167	down	NPPB
5.779805	down	CADM1
5.8380966	down	C12orf39
6.579152	down	CXCL14

Table 2.6 Up-regulated and down-regulated genes in PCa

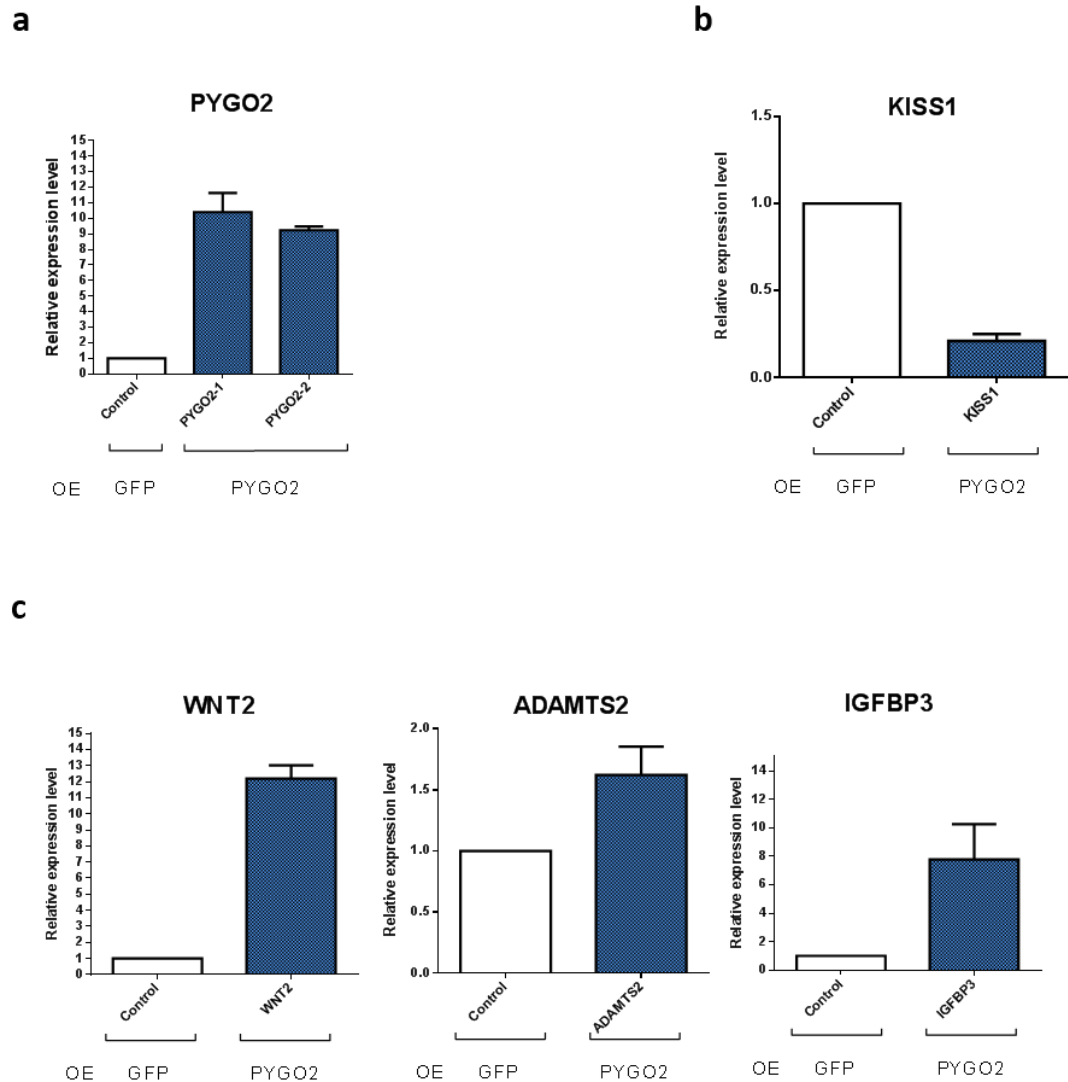


Figure 2.17 Expression level of WNT2, IGFBP3, ADAMTS2 and KISS1 and PYGO2 in LHMK and LHMK-PYGO2

PYGO2-induced gene expression in LHMK. a) PYGO2 up-regulation in LHMK overexpressed with PYGO2. b) Down-regulation of KISS1 in LHMK overexpressed with PYGO2. c) Overexpression of WNT2, IGFBP3, ADAMTS2. Bars: mean+SDE (n=3)

2.3.9 PYGO2 expression in established PCa cell lines.

For continued validation of PYGO2 function in PCa, PYGO2 expression levels in 10 PCa cell lines were characterized. As shown in Figure 2.18, PYGO2

is expressed in most PCa lines analyzed. 22Rv1 and PacMetUT cell lines showed higher expression levels comparing to the rest.

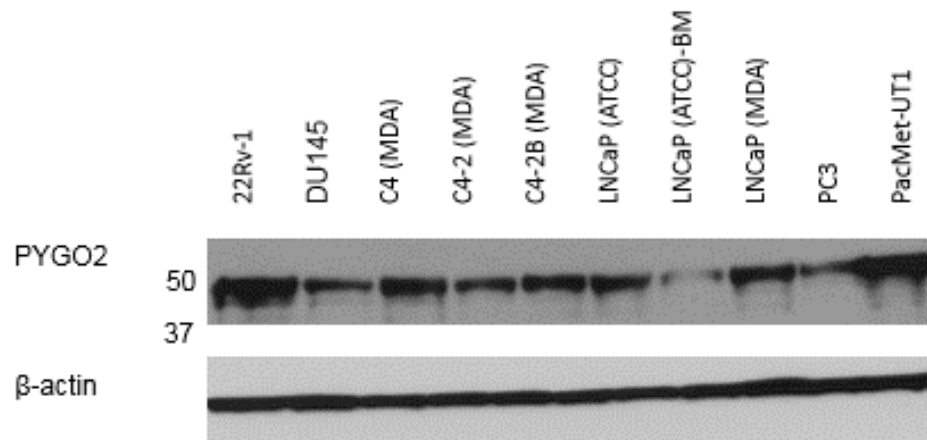


Figure 2.18 Investigation of PYGO2 expression level of PCa cell lines

Chapter 3 Discussion and future directions

3.1 Discussion

3.1.1 Functional *in vivo* screen identify genes with potential to promote PCa progression

Human cancer genome alternations consist of amplifications, indels (insertions and deletions) and mutations. Equipped with new cutting-edge tools such as exome sequencing and bioinformatics analysis, studies on copy number variations nowadays bring new perspective to discover novel bona fide oncogenes. In this study, functional *in vivo* screen advanced the current understanding based on cancer genome and transcriptome, collected the information about biological function of candidate genes in a high-throughput manner and ultimately pinpointed several putative oncogenes potentially playing various roles in multistep PCa progression.

The screen system is proved to be effective, as it successfully identified several known oncogenes in PCa, such as KRAS, CCNE2 and EZH2. However, in the *in vivo* screen, 29 out of 41 hits developed only 1 tumor out of 10 replicates. As the parental line LHMK for cell library is initially unable to form subcutaneous tumor (Berger et al., 2004b), the screen design sets a very high bar to push an extremely weakly tumorigenic cell line in nude mice to grow subcutaneous tumor. The penetrance of all the hits was general quite low while they were expected to play a potent role in PCa progression. Nevertheless, other amplified genes may also promote PCa or even be the driver since the screen was only design to identify most potent amplified genes in current screen system.

To link the hits to the context of PCa genomic and transcriptomic profile, the gene amplification and overexpression levels were re-examined in the original databases (Table 3.1). Most of the genes showed amplified in multiple databases despite that the bioinformatics study also included genes overexpressed without known amplification. Most strikingly, most of hits showed amplification in datasets from Grasso *et al* with the biggest number of malignant patients.

Name	Amplification		
	Taylor	TCGA	Grasso <i>et al</i>
MOS			+
PYGO2			+
BOP1	+		+
ST3GAL1	+		+
MTBP	+		+
ZKSCAN5		+	
TOMM40L			+
CCNE2	+		+

Table 3.1 Amplification of prioritized hit in PCa genomic datasets

Netwalker was used to explore potential interactions among hits in collaboration with Dr. Ram Prahald (data not shown). However, no clear interactions were shown, which indicated the hits per se rather than synergistically

promoted PCa progression. Moreover, it was interesting to note that 6 out of 44 genes are located on 8q24 chromosome (SLC45A4, BOP1, ST3GAL1, WDYHV1, MTBP, and DERL1) but not MYC. It indicates that other genes located at 8q24 can also promote PCa progression. I validated some but not all of their effect by *in vitro* assays (BOP1, ST3GAL1, and MTBP). They may act as driver gene working with or without MYC in PCa progression. Their potential interactions await in-depth study to identify the driver genes as well as to investigating their function in PCa.

As mentioned above, PYGO2 is selected as an interesting gene for further study because of its putative role in driving PCa progression based on its known biological functions. At the same time, BOP1 is another gene validated to promote sphere-formation, migration and invasion in LHMK cells. However, I was not able to find a good BOP1 antibody for western blot. By current commercial available antibodies, I am not able to confirm overexpression or knockdown in LHMK-BOP1 and LNCaP-BOP1 (data not shown), despite in qRT-PCR assays BOP1 expression levels are manipulated as expected. Moreover, the subcellular localization in tissue staining by current commercial available antibodies is questionable (data not shown). Thus, BOP1 awaits further study when reagents are available, as it can be a possible putative tumor promoting gene.

3.1.2 Activation of MAPK in PCa

The link of MOS and wild type KRAS being able to activate pMEK and potentiate tumorigenesis of LHMK cells was of particular interest. In PCa, RAS/RAF signaling was activated in 43% of primary tumor and 90% of metastases

(Taylor *et al.*, 2010). However, it has been shown that KRAS mutation was infrequent in PCa (Silan *et al.*, 2012). In fact, recent study that reported that MAPK activation could lead to higher tumor grade and metastasis in a transgenic mouse model, oncogenic KRAS allele G12D was used instead of overexpression of wild type KRAS (Mulholland *et al.*, 2012). Apparently, there was a level of lack of understanding how infrequent KRAS mutation can be reconciled with the hyperactivated MAPK pathway in PCa. The study shows that wild type KRAS or MOS, a MEK kinase, when overexpressed, could lead to higher pMEK signaling and higher tumorigenesis. Therefore, it is possible that this result could lead to a hypothesis that overexpression instead of oncogenic mutation serves as the mechanism for activating MAPK pathway in the context of PCa.

3.1.3 PYGO2 function in PCa

Further evaluation of PYGO2 as a hit from the screen has led to several interesting findings in functional assays both *in vitro* and *in vivo*:

1) PYGO2 increases sphere formation, invasion and migration of LHMK cells (Chapter II)

2) PYGO2 overexpression in LNCaP increases primary tumor growth and local invasion to lymph node (Chapter II)

3) Dr. Xin Lu's recent investigation on PYGO2 showed that *in vivo* silencing of PYGO2 with intratumoral injection of siRNA caused tumor shrinkage of one PDX model (data not shown).

Ongoing studies include assessing if PYGO2 knockdown in PC3 cell line can lead to reduced tumor formation in both orthotopic site and as bone metastasis. This would be of particular interest to us, as PYGO2 displayed higher expression level in metastases than in primary tumor in several ONCOMINE® datasets (Figure 3.1). The connection of PYGO2 and Wnt pathway also provides the clue that PYGO2 might be involved in bone metastasis, the most frequent metastasis type for PCa, as it has been demonstrated by previous studies that Wnt signaling plays an important role in osteoblastic bone metastasis of PCa (Ell and Kang, 2012).

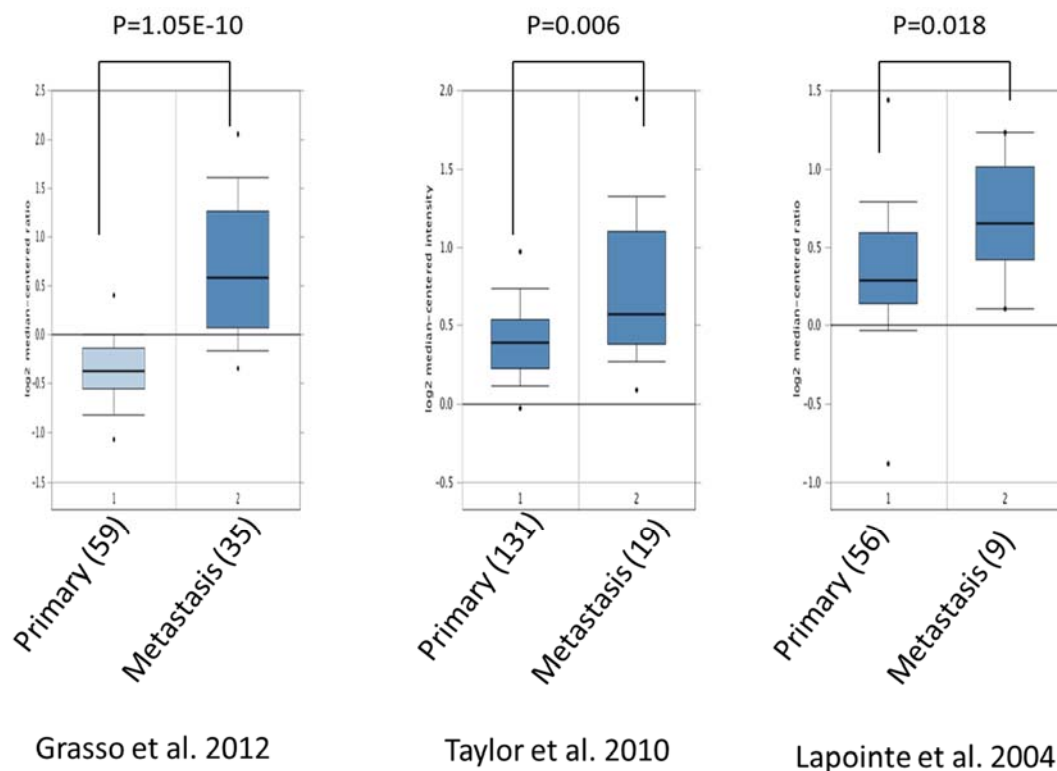


Figure 3.1 PYGO2 is upregulated in metastases in several ONCOMINE® datasets

3.1.4 PYGO2 upregulation in PCa

The preliminary assessment of PYGO2 expression in PCa TMA shows that PYGO2 is upregulated in PCa and correlated with higher Gleason score. Though Gleason score is designed to represent tumor aggressiveness of heterogenous prostate neoplasm, it is preferable to add more cases in current study. We have requested three additional TMAs from PCBN, a non-profit organization distributing prostate cancer clinical samples to the community. These TMAs include 52 Case Lymph Node Mets Array, 200 Case Grade/Stage Array, 217 Biochemical

Recurrence (CAP6) and 726 Case PSA Progression Array. While this is exciting, it raises the question of how PYGO2 is regulated in cancerous prostate. Based on what has been reported about PYGO2 and the above-indicated result, there are several possibilities:

1) PYGO2 is amplified in about 3% of PCa patients, based on an average estimate of several PCa genome databases including TCGA

2) PYGO2 might be upregulated by suppressing RB pathway in PCa. It was reported that in cervical cancer, attenuation of RB by HPV virus induces PYGO2 expression via ELF (Tzenov *et al.*, 2013). RB pathway is among the three most altered signaling pathways in PCa (Taylor *et al.*, 2010). Therefore, it is reasonable to argue that it is possible that RB pathway inactivation in PCa also induces PYGO2 upregulation. This is a possibility that Dr. Xin Lu is investigating.

3) Other possible explanations for PYGO2 upregulation may include Wnt pathway activation and certain microRNA, possibilities that Dr. Lu is investigating with Dr. Eun-Jung Jin in the lab.

3.1.5 PYGO2 mechanism for PCa

The following working model was proposed based on my current findings (Figure 3.2): PYGO2 may promote PCa through both Wnt-dependent and independent mechanisms. Both mechanisms are based on the ability of PYGO2 to bind to histone mark H3K4me and other chromatin regulators. The functional output of PYGO2-mediated chromatin regulation includes up-regulation of

important cancer related genes such as WNT2 and ADAMTS2 as well as down-regulation of KISS1. This model is under active investigation by the lab members now.

As identified and validated in various study (Li et al., 2007; Sun et al., 2014), loss of function for PYGO2 interferes with Wnt signaling pathways and leads to Wnt signaling deficiency. Meanwhile, preliminary study shows overexpression of PYGO2 activate Wnt pathways (Schlesinger et al., 2005). In our study, PYGO2 is selected for *in vivo* screen because of its focal amplification in PCa. Amplification is a possible mechanism for overexpression(Santarius et al., 2010). By amplification and/or by other unknown mechanisms(Bostwick and Qian, 2004), PYGO2 may overexpress and to further activate or promote Wnt signaling activation in PCa.

One of the major clinical devastations of PCa is bone metastasis, which is the most frequent metastatic sites(Shen and Abate-Shen, 2010a). Wnt signaling pathway has been reported to have a key role in bone metastasis (Humphrey, 2004; Pacelli and Bostwick, 1997; Regard et al., 2012). Besides other mechanism involving DKK1(Hall et al., 2005; Thudi et al., 2011), activation of β -catenin induces increased bone deposition and decreased osteoclast formation (Glass et al., 2005) Autocrine/paracrine activation Wnt signaling are reported in many studies in prostate cancer cell lines (Bostwick et al., 1993; Bostwick et al., 1995) and PCa(Chen et al., 2004). The exploration of PYGO2 in Wnt signaling especially in bone metastasis will provide new insight on this topic. It is of interest to study

PYGO2 involvement in such autocrine/paracrine Wnt signaling activation since the paracrine/autocrine positive feedback is formed when PYGO2 induces Wnt2 upregulation in LHMK cell. If it is the case in PCa, PYGO2 will actively induce WNT signaling.

After overexpression, PYGO2 may interfere downstream targets expression level by mediating chromatin regulation. The epigenetic study of PYGO2 will be conducted collaborating with Dr. Chunru Lin's group. ChIP assay would be applied to investigate PYGO2 mediated protein-DNA interaction.

To discover the downstream targets of PYGO, unbiased transcriptomic profiling is conducted. However, current validated WNT2, ADATMTS2 and KISS1 are only putative downstream targets. Shown in the results of transcriptomic profiling of PYGO2 overexpression in LHMK, other genes may also involve in PCa progression. On the list of PYGO-mediated upregulated genes, genes encoding collagens are shown to promote cancer advance, such as COL21A1(Abrahams et al., 2003) and COL3A1 (Sakr et al., 2000). We will further analyze the gene list combined with our results from other cell lines with PYGO2 overexpression and knocking-down, since such result will provide genes with epistatic change following PYGO2 dysregulation.

In summary, in my study, putative genes with amplifications and potentially promoting prostate cancer progression are identified and validated. Moreover, the study on PYGO2 will expand the current understanding of the function and mechanism of this specific gene in prostate cancer progression.

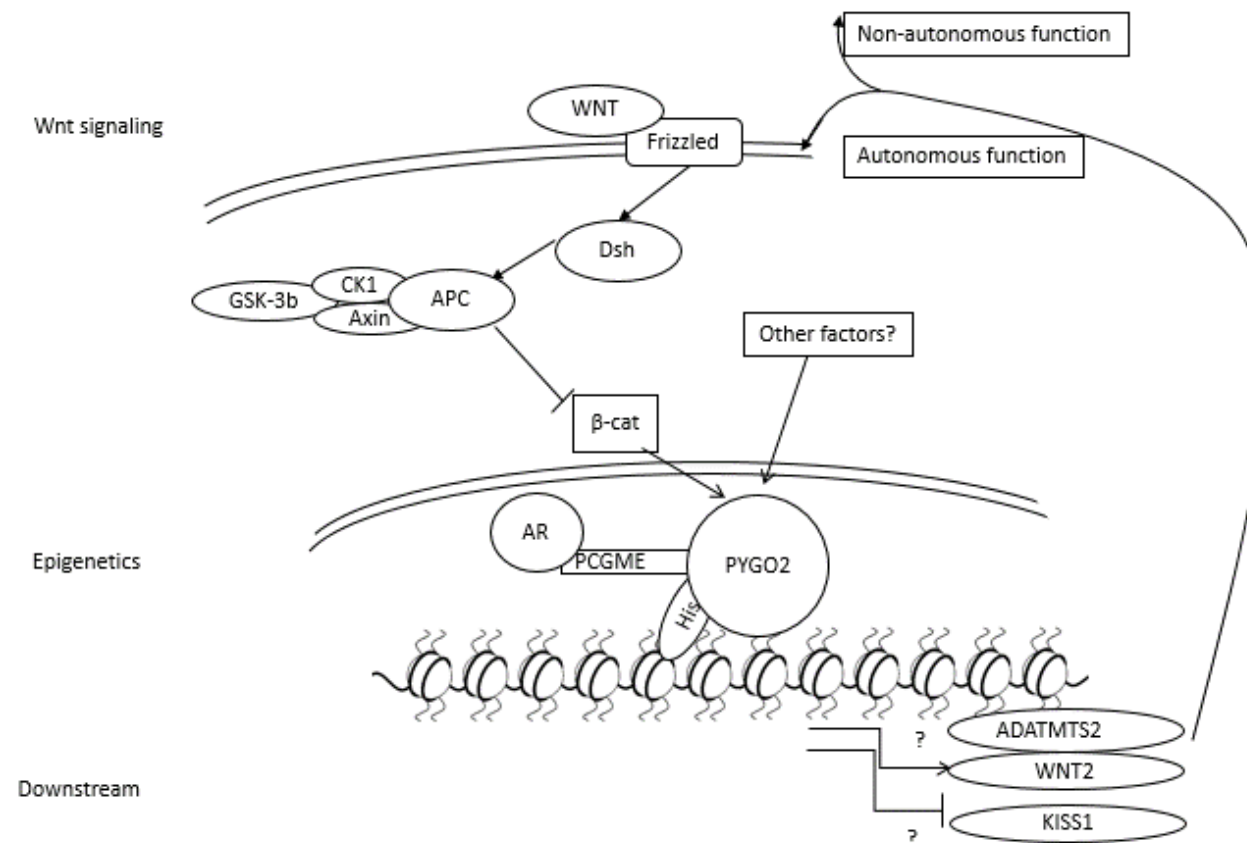


Figure 3.2 Current working hypothesis of PYGO2 in PCa

Bibliography

Abate-Shen, C., and Shen, M.M. (2002). Mouse models of prostate carcinogenesis. *Trends in genetics* : TIG 18, S1-5.

Abrahams, N.A., Bostwick, D.G., Ormsby, A.H., Qian, J., and Brainard, J.A. (2003). Distinguishing atrophy and high-grade prostatic intraepithelial neoplasia from prostatic adenocarcinoma with and without previous adjuvant hormone therapy with the aid of cytokeratin 5/6. *American journal of clinical pathology* 120, 368-376.

Aravindaram, K., and Yang, N.S. (2010). Anti-inflammatory plant natural products for cancer therapy. *Planta Medica* 76, 1103-1117.

Aryee, M.J., Liu, W., Engelmann, J.C., Nuhn, P., Gurel, M., Haffner, M.C., Esopi, D., Irizarry, R.A., Getzenberg, R.H., Nelson, W.G., Luo, J., Xu, J., Isaacs, W.B., Bova, G.S., and Yegnasubramanian, S. (2013). DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Science translational medicine* 5, 169ra110.

Barbieri, C.E., Baca, S.C., Lawrence, M.S., Demichelis, F., Blattner, M., Theurillat, J.-P., White, T.A., Stojanov, P., Van Allen, E., Stransky, N., Nickerson, E., Chae, S.-S., Boysen, G., Auclair, D., Onofrio, R.C., Park, K., Kitabayashi, N., MacDonald, T.Y., Sheikh, K., Vuong, T., Guiducci, C., Cibulskis, K., Sivachenko, A., Carter, S.L., Saksena, G., Voet, D., Hussain, W.M., Ramos, A.H., Winckler, W., Redman, M.C., Ardlie, K., Tewari, A.K., Mosquera, J.M., Rupp, N., Wild, P.J., Moch, H., Morrissey, C., Nelson, P.S., Kantoff, P.W., Gabriel, S.B., Golub, T.R., Meyerson, M., Lander, E.S., Getz, G., Rubin, M.A., and Garraway, L.A. (2012). Exome

sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 44, 685-689.

Belenkaya, T.Y., Han, C., Standley, H.J., Lin, X., Houston, D.W., Heasman, J., and Lin, X. (2002). *pygopus* Encodes a nuclear protein essential for wingless/Wnt signaling. *Development* 129, 4089-4101.

Berger, M.F., Lawrence, M.S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A.Y., Sboner, A., Esgueva, R., Pflueger, D., Sougnez, C., Onofrio, R., Carter, S.L., Park, K., Habegger, L., Ambrogio, L., Fennell, T., Parkin, M., Saksena, G., Voet, D., Ramos, A.H., Pugh, T.J., Wilkinson, J., Fisher, S., Winckler, W., Mahan, S., Ardlie, K., Baldwin, J., Simons, J.W., Kitabayashi, N., MacDonald, T.Y., Kantoff, P.W., Chin, L., Gabriel, S.B., Gerstein, M.B., Golub, T.R., Meyerson, M., Tewari, A., Lander, E.S., Getz, G., Rubin, M.A., and Garraway, L.A. (2011). The genomic complexity of primary human prostate cancer. *Nature* 470, 214-220.

Berger, R., Febbo, P.G., Majumder, P.K., Zhao, J.J., Mukherjee, S., Signoretti, S., Campbell, K.T., Sellers, W.R., Roberts, T.M., Loda, M., Golub, T.R., and Hahn, W.C. (2004a). Androgen-Induced Differentiation and Tumorigenicity of Human Prostate Epithelial Cells. *Cancer Research* 64, 8867-8875.

Berger, R., Febbo, P.G., Majumder, P.K., Zhao, J.J., Mukherjee, S., Signoretti, S., Campbell, K.T., Sellers, W.R., Roberts, T.M., Loda, M., Golub, T.R., and Hahn, W.C. (2004b). Androgen-induced differentiation and tumorigenicity of human prostate epithelial cells. *Cancer research* 64, 8867-8875.

Bostwick, D.G. (2000). Prostatic intraepithelial neoplasia. *Current urology reports* 1, 65-70.

Bostwick, D.G., Amin, M.B., Dundore, P., Marsh, W., and Schultz, D.S. (1993). Architectural patterns of high-grade prostatic intraepithelial neoplasia. *Human pathology* 24, 298-310.

Bostwick, D.G., Liu, L., Brawer, M.K., and Qian, J. (2004). High-grade prostatic intraepithelial neoplasia. *Reviews in urology* 6, 171-179.

Bostwick, D.G., and Qian, J. (2004). High-grade prostatic intraepithelial neoplasia. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 17, 360-379.

Bostwick, D.G., Qian, J., and Frankel, K. (1995). The incidence of high grade prostatic intraepithelial neoplasia in needle biopsies. *The Journal of urology* 154, 1791-1794.

Bubendorf, L., Schopfer, A., Wagner, U., Sauter, G., Moch, H., Willi, N., Gasser, T.C., and Mihatsch, M.J. (2000). Metastatic patterns of prostate cancer: an autopsy study of 1,589 patients. *Human pathology* 31, 578-583.

Chen, G., Shukeir, N., Potti, A., Sircar, K., Aprikian, A., Goltzman, D., and Rabbani, S.A. (2004). Up-regulation of Wnt-1 and beta-catenin production in patients with advanced metastatic prostate carcinoma: potential pathogenetic and prognostic implications. *Cancer* 101, 1345-1356.

Chin, L., Hahn, W.C., Getz, G., and Meyerson, M. (2011). Making sense of cancer genomic data. *Genes & development* 25, 534-555.

DeMarzo, A.M., Nelson, W.G., Isaacs, W.B., and Epstein, J.I. (2003). Pathological and molecular aspects of prostate cancer. *Lancet* 361, 955-964.

Drake, J.M., Graham, N.A., Lee, J.K., Stoyanova, T., Faltermeier, C.M., Sud, S., Titz, B., Huang, J., Pienta, K.J., Graeber, T.G., and Witte, O.N. (2013). Metastatic castration-resistant prostate cancer reveals inpatient similarity and interpatient heterogeneity of therapeutic kinase targets. *Proceedings of the National Academy of Sciences of the United States of America* 110, E4762-4769.

Ell, B., and Kang, Y. (2012). SnapShot: Bone Metastasis. *Cell* 151, 690-690.e691.

Felici, A., Pino, M.S., and Carlini, P. (2012). A changing landscape in castration-resistant prostate cancer treatment. *Frontiers in endocrinology* 3, 85.

Fiedler, M., Sánchez-Barrena, M.J., Nekrasov, M., Mieszczanek, J., Rybin, V., Müller, J., Evans, P., and Bienz, M. (2008). Decoding of methylated histone H3 tail by the Pygo-BCL9 Wnt signaling complex. *Molecular cell* 30, 507-518.

Glass, D.A., 2nd, Bialek, P., Ahn, J.D., Starbuck, M., Patel, M.S., Clevers, H., Taketo, M.M., Long, F., McMahon, A.P., Lang, R.A., and Karsenty, G. (2005). Canonical Wnt signaling in differentiated osteoblasts controls osteoclast differentiation. *Developmental cell* 8, 751-764.

Gleason, D.F. (1992). Histologic grading of prostate cancer: a perspective. *Human pathology* 23, 273-279.

Gonzalgo, M.L., Bastian, P.J., Mangold, L.A., Trock, B.J., Epstein, J.I., Walsh, P.C., and Partin, A.W. (2006). Relationship between primary Gleason pattern on needle biopsy and clinicopathologic outcomes among men with Gleason score 7 adenocarcinoma of the prostate. *Urology* 67, 115-119.

Grasso, C.S., Wu, Y.M., Robinson, D.R., Cao, X., Dhanasekaran, S.M., Khan, A.P., Quist, M.J., Jing, X., Lonigro, R.J., Brenner, J.C., Asangani, I.A., Ateeq, B., Chun,

S.Y., Siddiqui, J., Sam, L., Anstett, M., Mehra, R., Prensner, J.R., Palanisamy, N., Ryslik, G.A., Vandin, F., Raphael, B.J., Kunju, L.P., Rhodes, D.R., Pienta, K.J., Chinnaiyan, A.M., and Tomlins, S.A. (2012). The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 487, 239-243.

Grignon, D.J. (2004). Unusual subtypes of prostate cancer. *Modern Pathol* 17, 316-327.

Gu, B., Sun, P., Yuan, Y., Moraes, R.C., Li, A., Teng, A., Agrawal, A., Rheaume, C., Bilanchone, V., Veltmaat, J.M., Takemaru, K., Millar, S., Lee, E.Y., Lewis, M.T., Li, B., and Dai, X. (2009). Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation. *The Journal of cell biology* 185, 811-826.

Gu, B., Watanabe, K., Sun, P., Fallahi, M., and Dai, X. (2013). Chromatin effector Pygo2 mediates Wnt-notch crosstalk to suppress luminal/alveolar potential of mammary stem and basal cells. *Cell stem cell* 13, 48-61.

Hall, C.L., Bafico, A., Dai, J., Aaronson, S.A., and Keller, E.T. (2005). Prostate cancer cells promote osteoblastic bone metastases through Wnts. *Cancer research* 65, 7554-7560.

Horoszewicz, J.S., Leong, S.S., Kawinski, E., Karr, J.P., Rosenthal, H., Chu, T.M., Mirand, E.A., and Murphy, G.P. (1983). LNCaP model of human prostatic carcinoma. *Cancer research* 43, 1809-1818.

Humphrey, P.A. (2004). Gleason grading and prognostic factors in carcinoma of the prostate. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 17, 292-306.

Kessler, R., Hausmann, G., and Basler, K. (2009). The PHD domain is required to link *Drosophila* Pygopus to Legless/beta-catenin and not to histone H3. *Mech Dev* 126, 752-759.

Kumar-Sinha, C., Tomlins, S.A., and Chinnaiyan, A.M. (2008). Recurrent gene fusions in prostate cancer. *Nat Rev Cancer* 8, 497-511.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., Kiezun, A., Hammerman, P.S., McKenna, A., Drier, Y., Zou, L., Ramos, A.H., Pugh, T.J., Stransky, N., Helman, E., Kim, J., Sougnez, C., Ambrogio, L., Nickerson, E., Shefler, E., Cortes, M.L., Auclair, D., Saksena, G., Voet, D., Noble, M., DiCara, D., Lin, P., Lichtenstein, L., Heiman, D.I., Fennell, T., Imielinski, M., Hernandez, B., Hodis, E., Baca, S., Dulak, A.M., Lohr, J., Landau, D.A., Wu, C.J., Melendez-Zajgla, J., Hidalgo-Miranda, A., Koren, A., McCarroll, S.A., Mora, J., Lee, R.S., Crompton, B., Onofrio, R., Parkin, M., Winckler, W., Ardlie, K., Gabriel, S.B., Roberts, C.W., Biegel, J.A., Stegmaier, K., Bass, A.J., Garraway, L.A., Meyerson, M., Golub, T.R., Gordenin, D.A., Sunyaev, S., Lander, E.S., and Getz, G. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214-218.

Leo Shedlovsky, D.B., Irving Levenstein (1942). Titrations of human seminal fluid with acids and alkalis and their effects on the survival of sperm motility. *American Journal of Physiology*.

Li, B., Mackay, D.R., Ma, J., and Dai, X. (2004). Cloning and developmental expression of mouse pygopus 2, a putative Wnt signaling component. *Genomics* 84, 398-405.

Li, B., Rheaume, C., Teng, A., Bilanchone, V., Munguia, J.E., Hu, M., Jessen, S., Piccolo, S., Waterman, M.L., and Dai, X. (2007). Developmental phenotypes and reduced Wnt signaling in mice deficient for pygopus 2. *Genesis* 45, 318-325.

Liu, W., Laitinen, S., Khan, S., Vihinen, M., Kowalski, J., Yu, G., Chen, L., Ewing, C.M., Eisenberger, M.A., Carducci, M.A., Nelson, W.G., Yegnasubramanian, S., Luo, J., Wang, Y., Xu, J., Isaacs, W.B., Visakorpi, T., and Bova, G.S. (2009). Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nature medicine* 15, 559-565.

Mehra, R., Tomlins, S.A., Yu, J., Cao, X., Wang, L., Menon, A., Rubin, M.A., Pienta, K.J., Shah, R.B., and Chinnaiyan, A.M. (2008). Characterization of TMPRSS2-ETS gene aberrations in androgen-independent metastatic prostate cancer{Mehra, 2008 #3}. *Cancer research* 68, 3584-3590.

Melissari, M., Lopez Beltran, A., Mazzucchelli, R., Froio, E., Bostwick, D.G., and Montironi, R. (2006). High grade prostatic intraepithelial neoplasia with squamous differentiation. *Journal of clinical pathology* 59, 437-439.

Mellinger, G.T., Gleason, D., and Bailar, J., 3rd (1967). The histology and prognosis of prostatic cancer. *The Journal of urology* 97, 331-337.

Miller, T.C.R., Mieszczanek, J., Sánchez-Barrena, M.J., Rutherford, T.J., Fiedler, M., and Bienz, M. (2013). Evolutionary adaptation of the fly Pygo PHD finger toward recognizing histone H3 tail methylated at arginine 2. *Structure* 21, 2208-2220.

Mulholland, D.J., Kobayashi, N., Ruscetti, M., Zhi, A., Tran, L.M., Huang, J., Gleave, M., and Wu, H. (2012). Pten Loss and RAS/MAPK Activation Cooperate

to Promote EMT and Metastasis Initiated from Prostate Cancer Stem/Progenitor Cells. *Cancer Research* 72, 1878-1889.

Nair, M., Nagamori, I., Sun, P., Mishra, D.P., Rheaume, C., Li, B., Sassone-Corsi, P., and Dai, X. (2008). Nuclear regulator Pygo2 controls spermiogenesis and histone H3 acetylation. *Developmental biology* 320, 446-455.

Pacelli, A., and Bostwick, D.G. (1997). Clinical significance of high-grade prostatic intraepithelial neoplasia in transurethral resection specimens. *Urology* 50, 355-359.

Prasad, C.K., Mahadevan, M., MacNicol, M.C., and MacNicol, A.M. (2008). Mos 3' UTR regulatory differences underlie species-specific temporal patterns of Mos mRNA cytoplasmic polyadenylation and translational recruitment during oocyte maturation. *Molecular reproduction and development* 75, 1258-1268.

Regard, J.B., Zhong, Z., Williams, B.O., and Yang, Y. (2012). Wnt signaling in bone development and disease: making stronger bone with Wnts. *Cold Spring Harbor perspectives in biology* 4.

Sakr, W.A., Billis, A., Ekman, P., Wilt, T., and Bostwick, D.G. (2000). Epidemiology of high-grade prostatic intraepithelial neoplasia. *Scandinavian journal of urology and nephrology Supplementum*, 11-18.

Sakr, W.A., Grignon, D.J., Crissman, J.D., Heilbrun, L.K., Cassin, B.J., Pontes, J.J.E., and Haas, G.P. (1994). High-Grade Prostatic Intraepithelial Neoplasia (Hgpin) and Prostatic Adenocarcinoma between the Ages of 20-69 - an Autopsy Study of 249 Cases. *In Vivo* 8, 439-443.

Santarius, T., Shipley, J., Brewer, D., Stratton, M.R., and Cooper, C.S. (2010). A census of amplified and overexpressed human cancer genes. *Nat Rev Cancer* 10, 59-64.

Schlesinger, C., Bostwick, D.G., and Iczkowski, K.A. (2005). High-grade prostatic intraepithelial neoplasia and atypical small acinar proliferation: predictive value for cancer in current practice. *The American journal of surgical pathology* 29, 1201-1207.

Schwab, K.R., Patterson, L.T., Hartman, H.A., Song, N., Lang, R.A., Lin, X., and Potter, S.S. (2007). Pygo1 and Pygo2 roles in Wnt signaling in mammalian kidney development. *BMC biology* 5, 15.

Shen, M.M., and Abate-Shen, C. (2010a). Molecular genetics of prostate cancer: new prospects for old challenges. *Genes & development* 24, 1967-2000.

Shen, M.M., and Abate-Shen, C. (2010b). Molecular genetics of prostate cancer: new prospects for old challenges. *Genes & Development* 24, 1967-2000.

Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. *Ca-Cancer J Clin* 63, 11-30.

Silan, F., Gultekin, Y., Atik, S., Kilinc, D., Alan, C., Yildiz, F., Uludag, A., and Ozdemir, O. (2012). Combined point mutations in codon 12 and 13 of KRAS oncogene in prostate carcinomas. *Molecular biology reports* 39, 1595-1599.

Städeli, R., and Basler, K. (2005). Dissecting nuclear Wingless signalling: recruitment of the transcriptional co-activator Pygopus by a chain of adaptor proteins. *Mech Dev* 122, 1171-1182.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* *102*, 15545-15550.

Sun, P., Watanabe, K., Fallahi, M., Lee, B., Afetian, M.E., Rheaume, C., Wu, D., Horsley, V., and Dai, X. (2014). Pygo2 regulates beta-catenin-induced activation of hair follicle stem/progenitor cells and skin hyperplasia. *Proceedings of the National Academy of Sciences of the United States of America* *111*, 10215-10220.

Taylor, B.S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B.S., Arora, V.K., Kaushik, P., Cerami, E., Reva, B., Antipin, Y., Mitsiades, N., Landers, T., Dolgalev, I., Major, J.E., Wilson, M., Socci, N.D., Lash, A.E., Heguy, A., Eastham, J.A., Scher, H.I., Reuter, V.E., Scardino, P.T., Sander, C., Sawyers, C.L., and Gerald, W.L. (2010). Integrative Genomic Profiling of Human Prostate Cancer. *Cancer Cell* *18*, 11-22.

Thudi, N.K., Martin, C.K., Murahari, S., Shu, S.T., Lanigan, L.G., Werbeck, J.L., Keller, E.T., McCauley, L.K., Pinzone, J.J., and Rosol, T.J. (2011). Dickkopf-1 (DKK-1) stimulated prostate cancer growth and metastasis and inhibited bone formation in osteoblastic bone metastases. *The Prostate* *71*, 615-625.

Tzenov, Y.R., Andrews, P.G., Voisey, K., Popadiuk, P., Xiong, J., Popadiuk, C., and Kao, K.R. (2013). Human papilloma virus (HPV) E7-mediated attenuation of retinoblastoma (Rb) induces hPygopus2 expression via Elf-1 in cervical cancer. *Molecular cancer research : MCR* *11*, 19-30.

Yang, L., Lin, C., Jin, C., Yang, J.C., Tanasa, B., Li, W., Merkurjev, D., Ohgi, K.A., Meng, D., Zhang, J., Evans, C.P., and Rosenfeld, M.G. (2013). lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* 500, 598-602.

Yatani, R., Kusano, I., Shiraishi, T., Hayashi, T., and Stemmermann, G.N. (1989). Latent Prostatic-Carcinoma - Pathological and Epidemiological Aspects. *Jpn J Clin Oncol* 19, 319-326.

Yates, L.R., and Campbell, P.J. (2012). Evolution of the cancer genome. *Nature reviews Genetics* 13, 795-806.

Zhang, S., Li, J., He, F., and Wang, X.M. (2014). Abnormal nuclear expression of Pygopus-2 in human primary hepatocellular carcinoma correlates with a poor prognosis. *Histopathology*.

Zhou, S.-Y., Xu, M.-L., Wang, S.-Q., Zhang, F., Wang, L., and Wang, H.-Q. (2014). Overexpression of Pygopus-2 is required for canonical Wnt activation in human lung cancer. *Oncol Lett* 7, 233-238.

Vita

Xiaolu Pan was born in Nanjing, Jiangsu, China on September 8th, 1989, the daughter of Yan Pan and Xiaoping Huo. After completing her study in Nanjing No.1 High School, Nanjing, Jiangsu in 2007, she entered Guiyang Medical University (GMU) in Guiyang, Guizhou, China. She received the degree of Bachelor of Medicine in Clinical Medicine (Radiology) from GMU in June, 2012. In September of 2012, she entered the University of Texas, Graduate School of Biomedical Sciences at Houston.

Permanent address:

84 Houzaimendongcun

18-506

Nanjing, Jiangsu, China 210016