University of Alberta

Characterization of PAX3 target genes and its roles in malignant melanoma

by

Yisu Li

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science

Medical Sciences – Medical Genetics

Canada

ABSTRACT

PAX3 plays important roles in development, postnatally, and in multiple

diseases. PAX3 contains two DNA-binding domains: paired domain (PD) and

homeodomain (HD). To understand their role in target gene recognition, known

PAX3 targets were analyzed *in silico*. This revealed that only sites containing

both PD and HD recognition motifs met criteria of phylogenetic conservation and

efficient binding. Motif analyses also indicated that these composite sites were

over-represented in genomic fragments recovered based on PAX3 binding.

Therefore, the composite sequence was used for *de novo* discovery of PAX3

targets. A genome scan identified three target genes relevant to PAX3 biology:

BOC, PRDM12 and *LBX1*. Lastly, analyses of PAX3 in melanoma revealed that

its expression was cell cycle regulated and that attenuation of PAX3 in mouse

melanoma cells caused a significant reduction in cell number due to G0/G1

accumulation. My results improve understanding of PAX3 regulatory behavior

and its role in melanoma.

ACKNOWLEDGEMENT

First of all, I would like to thank Dr. Alan Underhill for taking me as a master student even though he knew that I was not illegible for most major scholarships as an International student. Although the master study was different from what I imagined and I was not able to do some experiments that I intended to do, I still have to thank Alan for offering me the opportunity to have hand-one experiences in molecular biology and funding me for most of my graduate researches. I would certainly acknowledge the Faculty of Medicine and Dentistry for offering me one and half years of Graduate Student Assistantship and the Department of Medical Genetics for covering half of my tuition fees throughout my graduate study.

I would also like to thank Dr. Fred Berry and Dr. Heather McDermid for contributing their time to be on my graduate committee, and Dr. Rachel Wevrick to agree to chair my defense at the last minute.

My master research would not have gone thus far without the help from other students, technicians and professors. I am addressing my biggest acknowledgement to Ms. Ning Hu for her technical assistance in the lab and her caring throughout my graduate study, as well as Leanna Tsang for being a wonderful lab mate. I would like to thank Dr. Gareth Corry for helping me with my EMSA experiments in the lab; Dr. Ing Swie Goping and her former technician Matthew Czernick for showing me FACS analysis; and Dr. Xuejun Sun at the Cross Cancer Institute microscope facility for his assistance in using microscopes.

Additional thanks also go to assistance I received from flow cytometry facilities at the Heritage Medical Research Centre and the Cross Cancer Institute.

Last but not least, I really have to thank my parents back in Shanghai, China for your love and support in every aspect of my life. I also have to thank my boyfriend Adam Haulena for staying with me for the three and half years of my graduate study and always giving me supports and encouragements. I love you all.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

APPENDIX III

ABBREVIATIONS

aa	amino acids
ARMS	Alveolar rhabdomyosarcoma
BCL-X_L	B-cell lymphoma-extra large
BOC	Brother of CDO
bp	Base pair
BrdU	Bromodeoxyuridine
BTRC	Beta-transducin repeat containing
CASTing	Cyclic amplification and selection of targets
CDO	Cell Adhesion Molecule-Related/Down Regulated by Oncogenesis
CHIP	Chromatin immunoprecipitation
CML	Chronic myeloid leukaemia
CNS	Central nervous system
DCT	Dopachrome tautomerase or trysinase-related protein-2
DM	Dermomyotome
DNA	Deoxyribonucleic acid
DRG	Dorsal root ganglia
DSHB	Developmental Studies Hybridoma Bank
e5	The PAX3 binding sequences at the promoter of *eve* gene, containing both PD and HD motifs
ECR	Evolutionary Conserved Region
EMSA	Electrophoretic mobility shift assay
ENS	Enteric nervous system

eve	Drosophila *even-skipped* gene
FACS	Fluorescence Activated Cell Sorting
FITC	Fluorescein isothiocyanate
FKHR	Forkhead transcription factor
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gsb	Drosophila gooseberry protein
HD	Homeodomain
HIR	Hstone cell cycle regulation
HIRA	Histone cell cycle regulation defective homolog A
HOX11	Homeobox proto-oncogene
L	Lumen
LBX1	Ladybird modeobox 1
MET	Tyrosine kinase receptor
MITF	Microphthalmina-associated transcription factor
MSCs	Melanocyte stem cells
Myc	Myelocytomatosis oncogene
MYF5	Myogenic factor 5
NT	Neural tube
OP	Octapeptide
p21	Cyclin-dependent kinase inhibitor 1
p21^{Cip1}	Cyclin-dependent kinase inhibitor
p53	Protein 53
PAC	Paclitaxel

PAX	Paired-box
PBS	Phosphate buffered saline
PD	Paired domain
PDHD	PD and HD
PI	Propidium Iodide
PNS	Peripheral nervous system
prd	Drosophila paired protein
PRDM12	Positive regulatory (PR)-domain containing protein 12
PTEN	Phosphatase and tensin homolog
Rb	Retinoblastoma protein
RB1	Retinoblastoma 1
Ret	Tyrosi receptor kinase
RMS	Rhabdomyosarcoma
RNA	Ribonucleic acid
RNAi	RNA interference
RNAiR-Pax3	RNAi-resistant PAX3 expresion plasmid
SHH	Sonic Hedgehog
siRNA	small interfering RNA
SMSC	Skeletal muscle stem cells
Sp	*Splotch*
sp^d	Splotch - delayed
TFBS	Transcription factor binding sites

TGFß	Transforming growth factor beta
TRANSFAC	Transcription factor binding prediction software
TYRP-1	Tyrosinase related protein 1
UTR	Untranslated Region
WGS	Whole-genome shotgun
WS	Waardenburg syndrome

CHAPTER 1 - INTRODUCTION

CHAPTER 1 - INTRODUCTION

1.1 PAX family proteins

With the isolation of many developmental control genes from *Drosophila* in the late 1980's, a new domain, paired-box domain (PD), containing 128 amino acids (aa), was found in the paired (prd) and gooseberry (gsb) proteins (Bopp *et al*, 1986). The Prd and gsb proteins contain another conserved motif – a homeobox domain (HD) containing 61 aa, which was considered to be the primary DNA binding motif for these proteins when they were first described. Following the discovery of pox-meso and pox-neuro, which contained only the PD (Bopp *et al*, 1989), Treisman and his colleagues revealed that the PD was able to bind to the promoter of the *even-skipped* gene (*eve*) (*e5*) independent of the HD (Treisman *et al*, 1991), and based on sequence similarity, the PD motif was found to be conserved in many diverse metazoan organisms (Burri *et al*, 1989; Walther *et al*, 1991). Together, this led to the description of the Paired-box (PAX) family of transcription factors (Walther *et al*, 1991).

In addition to the PD and HD, there are two other conserved features that exist in PAX proteins: an octapeptide (OP) between PD and HD, and a C-terminal domain rich in proline, serine, and threonine residues (Tremblay and Gruss 1994; Underhill 2000; Lang *et al*, 2007). To date, nine mammalian PAX proteins have been identified (PAX1-9) and they have been categorized into four subgroups (I-IV) according to their domain structure (Balczarek *et al*, 1997; Figure 1-1). A PD is present in all nine PAX proteins. Class I (PAX1/9) proteins also contain an OP but have no HD. In addition to both the PD and OP, class III (PAX3/7) proteins have a complete HD, while class II PAX proteins (PAX2/5/8) only possess the

first alpha-helix of the HD (Dahl *et al*, 1997). PAX proteins in class IV (PAX4/6) are distinguished from other PAX classes by containing a PD and complete HD but lacking the OP motif.

PAX proteins have been shown to be critical for embryogenesis and loss-of-function mutations cause a number of abnormalities in human and mouse (for review, see Underhill 2000; Wang *et al*, 2008; summarized in Figure 1-1). Expression of *Pax1* and *Pax9* has been observed in the sclerotome, where they contribute to skeletal development (Deutsch *et al*, 1988; Neubuser *et al*, 1995; Mise *et al*, 2008). Haploinsufficiency of *Pax1* causes the *Undulated (un)* phenotype, a mouse mutant with defects in vertebral bodies and intervertebral discs (Balling *et al*, 1988). Several *Pax* genes are expressed in the central nervous system (CNS) and play critical roles in tissue patterning and organogenesis. Murine *Pax2* is an essential factor in the formation of the kidney and urogenital system (Torres *et al*, 1995), and mutations in human *PAX2* have been associated with renal-coloboma syndrome (Sanyanusin *et al*, 1995a; Sanyanusin *et al*, 1995b). *Pax8* is expressed in the thyroid gland and excretory system during development (Plachov *et al*, 1990; Poleev *et al*, 1992), and mice lacking both *Pax2* and *Pax8* display a complete loss of kidney structure (Bouchard *et al*, 2002). Another class II protein, *Pax5* is expressed at the onset of B-cell lymphopoiesis and is essential for development of the B-cell lineage and midbrain patterning (Urbanek *et al*, 1994). Class IV (PAX4/6) proteins are important for pancreas development and play essential roles in the differentiation of insulin- and glucagon-producing cells (Sosa-Pineda *et al*, 1997; St-Onge *et al*, 1997). PAX6

3

also plays a crucial role in ocular development. Mutations in *Pax6/PAX6* cause

the mouse *Small eye* (*Sey*) mutant and human aniridia (Hill *et al*, 1991; Jordan *et*

al, 1992). Moreover, inactivation of mouse *Pax6* causes a total loss of the multi-

potency of retinal progenitor cells (Marquardt *et al*, 2001). Lastly, the Class III

proteins PAX3 and PAX7 are important regulators of embryonic and postnatal

myogenesis (for review, see Buckingham and Relaix 2007; see section 1.3 -

Myogenesis). Both PAX7 and PAX3 have also been implicated in the

maintenance and differentiation of satellite cells (for review, see Buckingham and

Montarras 2008; see section 1.3 – Melanocyte and skeletal muscle stem cells).

PAX3 also plays essential roles in melanogenesis (for review, see Kubic *et al*,

2008; see section 1.3 - Melanogenesis). Haploinsufficiency at *Pax3/PAX3* leads to

defects in the development of the neural tube and neural crest in the mouse mutant

Splotch and Waardenburg syndrome in humans (see section 1.2 – Waardenburg

Syndrome). Deregulation of *PAX3* expression has been implicated in

tumorigenesis and malignancy (see section 1.2 – Association of PAX3 with

malignancy). The following sections will focus on PAX3, including its

developmental roles, associated pathogeneses, gene structure, target genes and

DNA-binding, as well as structure-function relationships.

Group	Paired domain	Octapeptide	Homeodomain	Chromosome Mouse	Chromosome Human	Mutants Mouse	Mutants Human	Tumor-associated
I Pax1				2	20p11.2	*Un*	-	N
Pax9				12	14q12-q13	-	Oligodontia	N
II Pax2				19	19q24.3-q25.1	*Krd/1Neu*	Renal-coloboma	Y
Pax5				4	9p13	-	-	Y
Pax8				2	2q12-q14	-	Congenital hypothyroidism	Y
III Pax3				1	2q35	*Sp*	Waardenburg	Y
Pax7				4	1p36.2-p36.12	-	-	Y
IV Pax4				6	7q32	-	-	N
Pax6				2	11p13	*Sey*	Aniridia	N

Figure 1-1. Summary of PAX family proteins (Underhill 2000; Wang *et al*, 2008). Nine PAX proteins are classified into 4 subgroups I, II, III and IV according to their conserved domain structure. The paired domain (PD), octapeptide (OP) and homeodomain (HD) are indicated as light blue, dark red and orange ellipses. Class II proteins contain a truncated HD, which is elucidated as a smaller orange ellipse. Chromosomal locations of each gene are included for both mouse and human. Mouse mutants and human *PAX*-related syndromes are listed (*un, undulated*; *Krd, kidney and retinal defects*; *1Neu, 1-bp insert frame-shift mutation in Pax2 gene*; *Sp, Splotch*; *Sey, Small eye*). PAX proteins have been implicated in cancer pathology are indicated as "Y" in the "Tumor-associated" column, and "N" means their relationships to cancer development have not been established.

1.2 PAX3 biology

Pax3 transcripts are first observed on embryonic day 8.5-9 (E8.5-E9) in mouse and are restricted to the dorsal part of the neural tube and the adjacent dermomyotome (Goulding *et al*, 1991). *Pax3* expression extends throughout the dorsal spinal cord from E11 to E13 and in the limb buds from E10 to E11, becoming undetectable after E14 (Goulding *et al*, 1991). After birth, the expression of *Pax3* is restricted to adult satellite cells (or skeletal muscle stem cells) and melanocyte stem cells (Relaix 2006; Osawa *et al*, 2005).

Studies of *Splotch* mice mapped this allele in *Pax3* on mouse chromosome 1. These mice revealed multiple defects in muscle and neural crest derived tissues caused by PAX3 loss-of-function and they have been used as a model for understanding neural tube defects in humans (Epstein *et al*, 1991). Homozygous *Sp* mice die prematurely, approximately 13 or 14 days of gestation. Most of them develop spina bifida (lumbosacral rachischisis), exencephaly (cranioschisis) and multiple neural crest defects that affect the peripheral and enteric nervous systems, heart, Schwann cells, melanocytes in the skin and inner ear, as well as an absence of limb musculature. Heterozygous *Sp* mice exhibit white-spotting on the belly and forehead, and overall growth retardation. These observations established that PAX3 has important roles in neural tube closure, neural crest formation, myogenesis and melanogenesis (for review, see Moase and Trasler 1992; Chi and Epstein 2002; see Figure 1-2 for a summary of PAX3 regulatory networks). The following sections cover these basic roles of PAX3 in greater detail.

6

Neurogenesis

Pax3 was first suggested to have a role in neurogenesis due to its expression in the neural tube and neural crest during embryogenesis (Goulding *et al*, 1991). Later experiments in PAX3-deficient *Splotch* mice revealed a reduction of dorsal root ganglia (DRG) and sensory neurons, as well as defects in neural patterning and development (Conway *et al*, 1997; Koblar *et al*, 1999; Tremblay *et al*, 1995; Moase and Trasler 1992; Tremblay and Gruss 1994). Wnt signaling is essential for early neural crest formation and differentiation (for review, see Wu *et al*, 2003). In Xenopus, *Pax3* expression was shown to be induced by *XWnt-8* in the lateral neural plate during early embryo development (Bang *et al*, 1999). It was also demonstrated that the presence of both *Pax3* and another important neural crest development factor, *Zic1* (zinc finger factor related 1), was required along with Wnt for the determination of neural cell fate in *Xenopus* (Sato *et al*, 2005). Moreover, PAX3 was suggested to regulate enteric nervous system (ENS) development by directly regulating the expression of *c-Ret* (Lang *et al*, 2000), which is a tyrosine receptor kinase required for the survival and proliferation of enteric ganglion precursors in the ENS (Durbec *et al*, 1996). Interestingly, it should be noted that there have been no enteric ganglion defects reported in patients with Waardenburg syndrome thus far. Although *Pax3/PAX3* has been implicated in multiple aspects of neurogenesis, the genetic network that PAX3 regulates is still to be discovered.

Melanogenesis

7

Melanocytes are pigment-producing cells derived from the neural crest

cells. In mice, melanoblasts differentiate from the neural crest at ~E8.5-9 and then

migrate to the dermis and terminally differentiate into melanocytes. By E14.5,

melanocytes populate the epidermis, hair follicles, iris, and cochlea of the inner

ear (Erickson 1993; Mackenzie *et al*, 1997; Blake and Ziman 2005). PAX3 is

thought to play important roles in the commitment of cells to the melanoblast

lineage and/or promoting melanoblast proliferation (Hornyak *et al*, 2001).

Another important melanocyte transcription factor is Microphthalmina-associated

transcription factor (MITF), which is considered to be the "master regulator" of

the melanocyte lineage and its expression is required for the survival and

migration of melanoblasts (Opdecamp *et al*, 1997; Hornyak *et al*, 2001). Loss-of-

function of *Mitf/MITF* has been shown to lead to defects in melanocytes of the

skin, hair and eyes, as well as in mast cells and osteoclasts in both human and

mouse (Tassabehji *et al*, 1994; Hodgkinson *et al*, 1993). Mutations in the human

MITF cause Waardenburg syndrome type 2, which displays phenotypic similarity

with PAX3 haploinsufficiency in WS3 (Nobukuni *et al*, 1996; Epstein *et al*, 1991;

Chalepakis *et al*, 1994). This epistasis led to the discovery of a 26bp region in the

human *MITF* promoter that was directly bound by PAX3 (Watanabe *et al*, 1998),

thus placing PAX3 on top of the melanocyte developmental hierarchy.

PAX3 is also thought to be involved in the pigmentation process in

melanocytes by regulating the expression of two genes that encode proteins

required for melanin-synthesis: DCT and TYRP1. *Dct* encodes dopachrome

tautomerase or tyrosinase-related protein-2 (TRP2) (Hornyak *et al*, 2001) and is

8

required for pigment production (Jackson *et al*, 1992) and has been implicated in

melanoblast survival (Steel *et al*, 1992). MITF has been shown to co-activate *Dct*

with SOX10, another important melanocyte transcription factor (Jiao *et al*, 2004;

Ludwig *et al*, 2004). PAX3 can compete for binding to the same region of the *Dct*

promoter as MITF and SOX10, thereby inhibiting *Dct* transcription in

melanocytes (Lang *et al*, 2005). In contrast to *Dct*, PAX3 has been shown to

activate transcription of *Tyrp1* that encodes tyrosinase related protein 1 to

promote melanin synthesis (Galibert *et al*, 1999). Although the entire mechanism

whereby PAX3 regulates melanocyte development and pigmentation has not been

mapped out, the importance of PAX3 in this lineage is clear from these studies.

Myogenesis

The myogenic lineage is known to be regulated by two major "myogenic

regulatory factors" or MRFs - MYF5 and MyoD (Tajbakhsh and Buckingham

2000). Mice carrying homozygous mutations in both *Myf5* and *MyoD* completely

lack skeletal muscle and myoblasts (Rudnicki *et al*, 1993). In *Myf5* null mice,

muscle progenitor cells were present but they were not able to migrate to the

proper sites for differentiation (Tajbakhsh *et al*, 1996), whereas *MyoD* null

mutants had no obvious skeletal abnormalities but show elevated *Myf5*

transcription (Rudnicki *et al*, 1992). Although this indicates that *Myf5* is able to

compensate for *MyoD* in skeletal muscle formation during embryogensis, *MyoD*

null myoblasts did display proliferation and growth defects *in vitro* (Montarras *et*

al, 2000) and *MyoD* null animals showed a severe deficiency in muscle

regeneration after birth (Megeney *et al*, 1996). Together, these data suggest that

9

Myf5 and MyoD can act redundantly in skeletal muscle, but also have distinct

functions.

Skeletal muscle in vertebrate embryos is derived mostly from the somitic

mesoderm (Christ and Ordahl 1995; Tajbakhsh and Buckingham 2000). During

somite maturation, myogenic progenitor cells in the dermomyotome layer of the

somite begin to differentiate and form the underlying myotome and can also

delaminate and migrate to distal locations such as the limbs or diaphragm

(Buckingham 2006). *Pax3* is expressed in myogenic progenitor cells of the

hypaxial dermomyotome (Goulding *et al*, 1991) and in migratory myoblasts, but

is down-regulated when their destination is reached (Goulding *et al*, 1994).

Moreover, *Pax3* null embryos exhibit a complete lack of skeletal muscle

formation in the limbs due to the fact that precursors fail to delaminate and enter

migratory pathways (Goulding *et al*, 1994; Bober *et al*, 1994). A similar

phenotype was observed in mouse embryos that lack expression of the tyrosine

kinase receptor MET (Bladt *et al*, 1995). The presence of MET is necessary for

the migration of muscle progenitor cells from the dermomyotome (Dietrich *et al*,

1999). In the *Splotch* mouse, *cMet* expression is greatly reduced in

dermomyotome and no *cMet* expression is observed in the limbs, raising the

possibility that the limb muscle defect in *Sp* was due to a lack of *cMet* expression

(Yang *et al*, 1996). Subsequently, a PAX3 recognition site was found within the

cMet promoter and was activated by PAX3 in reporter gene assays (Epstein *et al*,

1996), providing a mechanistic link between PAX3 and migration of myogenic

precursor cells.

Although PAX3 does not directly control later events in muscle differentiation, it does function in the determination of myogenic cell fate via acting on *Myf5* and *MyoD*. For instance, although the *MyoD* gene was activated in mice null for either *Pax3* or *Myf5*, it was not in *Sp/Myf5* double homozygous embryos (Tajbakhsh *et al*, 1997). These studies indicated that *Pax3* and/or *Myf5* acted genetically upstream of *MyoD*. Further research indicated that *Pax3* is also genetically upstream of *Myf5*. Bajard *et al* characterized a PAX3 binding site within a 145bp regulatory element located -57.5kb from the *Myf5* transcription start site (Bajard *et al*, 2006). Moreover, this distal enhancer was required to confer *Myf5* expression in mature somites and limb muscles (Buchberger *et al*, 2003). This element was bound by PAX3 *in vivo* in chromatin immunoprecipitation (ChIP) experiments and was essential for *Myf5* activation and subsequent myogenesis (Bajard *et al*, 2006). As in melanocytes, PAX3 therefore functions upstream of master regulators of muscle development.

Melanocyte and skeletal muscle stem cells

Despite the fact that PAX3 expression is normally extinguished upon terminal differentiation, PAX3 expression is maintained in both melanocyte stem cells (MSCs) and skeletal muscle stem cells (or satellite cells), which reside in hair follicles and under the basal lamina of muscle fibres, respectively (Nishimura *et al*, 2002; Buckingham and Montarras 2008). These multipotent PAX3-positive cells remain quiescent until they are activated in response to external stimuli, such as tissue damage, whereupon they can self-renew within their niche and give rise

11

to progeny that will ultimately exit the cell cycle and differentiate (for review, see Lagha *et al*, 2008; Kubic *et al*, 2008).

Gene expression profiling of mouse MSCs revealed that *Pax3* and *Dct*, but not *Sox10*, *Mitf* or *Tyrp1*, were expressed in MSCs in adult tissue (Osawa *et al*, 2005). The expression of *Pax3* but not *Sox10* nor *Mitf* was suggested to facilitate maintenance of the multipotent quiescent state and prevent terminal differentiation (Nishikawa and Osawa 2007). Taken together with evidence that Wnt signaling can trigger a change in balance of PAX3, MITF and SOX10 levels (for review, see Kubic *et al*, 2008) and PAX3 has been suggested to directly regulate the expression of the Wnt signaling protein WNT1 (Fenby *et al*, 2008), this may provide a means to switch between quiescence, proliferation, and differentiation of MSCs.

Although early studies suggested PAX3 expression was terminated upon differentiation of myogenic precursors and that only PAX7 expression was retained in skeletal muscle stem cells (SMSCs or satellite cells), recent analyses find that this is not the whole picture. Specifically, PAX3 is expressed in a subgroup of SMSCs (Relaix 2006). These cells are derived from a distinct group of PAX3/PAX7-positive cells during development and continue to express PAX3 and maintain their proliferative potential even after birth (Relaix *et al*, 2005; Kassar-Duchossoy *et al*, 2005). A complicating factor in understanding the specific role PAX3 plays in these SMSCs is that both PAX3+ and PAX7+ populations are found in association with muscle fibers in adult tissue and both possess regeneration ability. Moreover, both PAX3 and PAX7 have the ability to

initiate the myogenic differentiation program by activating *MyoD* or *Myf5* alone

(Relaix 2006). Nevertheless, *Pax7* mutant mice showed a striking loss of satellite

cells in all body mass even in the presence of wide-type *Pax3* (Relaix 2006). In

addition, upon replacing *Pax3* with *Pax7*, the defect was obvious in distal muscle

masses but not the trunk (Lagha *et al,* 2008), indicating these proteins have

specialized functions in SMSCs.

1.3   PAX3-related diseases

Waardenburg Syndrome

Following the identification of *Pax3* as the gene mutated in *Splotch* mice

(Epstein *et al,* 1991), mutations in the human *PAX3* gene were reported in patients

with Waardenburg syndrome (WS) (Tassabehji *et al,* 1992; Baldwin *et al,* 1992).

WS is a rare inherited disorder that is commonly characterized by white patches

of hair, blue eyes due to lack of pigmentation in the iris, and various degrees of

hearing loss (for review, see Pingault *et al,* 2010). It was first reported in 1951 by

a Dutch ophthalmologist (Waardenburg 1951) and this form is now known as

Waardenburg syndrome type I (WS1; MIM#193500). Subsequently, WS patients

have been classified into three further subtypes noted as WS2 to WS4 (Read and

Newton 1997; Pingault *et al,* 2010), which vary in phenotype. Types 1 and 2 are

the most common with type 2 being the least severe and displaying no additional

abnormalities beyond those mentioned above. WS1 is also called WS with

dystopia canthorum, which involves lateral displacement of the ocular inner

canthi. WS3 patients are characterized by both dystopia canthorum and various

13
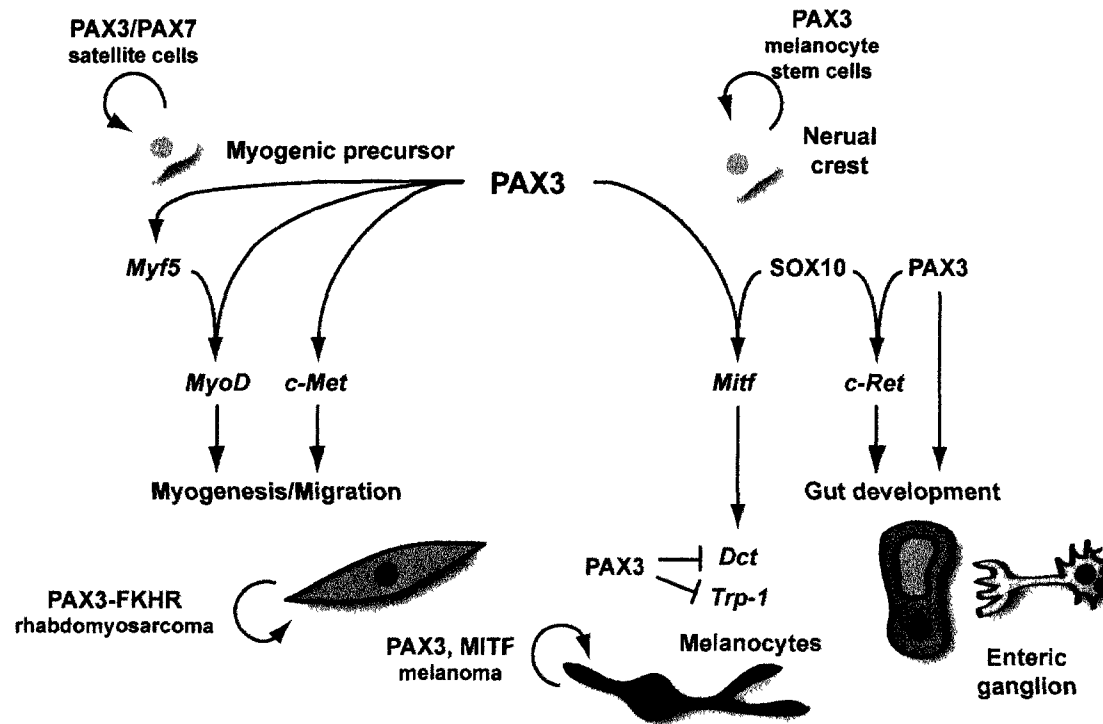
Figure 1-2. Summary of PAX3 regulatory networks. PAX3 has an essential role in gene regulation during neurogenesis, melanogenesis and myogenesis. PAX3 is involved in initiating the process of differentiation in enteric ganglion, melanocytes and skeletal muscle. PAX3 is also involved in maintaining the proliferation status of melanocyte stem cells and satellite cells. PAX3 has been suggested to regulate a list of important genes in these lineages, including *Myf5, c-Met, Mitf, Dct, Trp-1 (Tyrp1),* and *c-Ret.*

14

degrees of musculoskeletal abnormalities, especially in the upper limbs. WS4 is

the only autosomal recessive subtype described thus far and is also known as

Hirschsprung's disease or failure to innervate the enteric ganglion. Both WS1 and

WS3 are caused by disruptions in the *PAX3* gene. Since mutations in *PAX3* were

first linked to WS1 in 1992 (Tassabehji *et al*, 1992; Baldwin *et al*, 1992), about 70

distinct alterations in human *PAX3* have been described. These include missense,

nonsense and frameshift mutations, as well as larger insertions, deletions, and

aberrant splicing (for review, see Pingault *et al*, 2010). The nature of these

mutations establishes that loss-of-function in *PAX3* is responsible for the defects

seen in WS1 and WS3.

Association of PAX3 with malignancy

While insufficient PAX3 levels cause congenital defects, deregulation of

PAX3 activity has been associated with malignancy, notably in alveolar

rhabdomyosarcoma (ARMS) and melanoma, where it may promote proliferation

and survival of cancer cells (for review, see Kubic *et al*, 2008). ARMS is the most

common childhood solid tumor and is a soft tissue sarcoma derived from skeletal

muscle precursors (Lagutina *et al*, 2002). Most of these tumors are characterized

by a translocation between chromosomes 2 and 13 (t(2;13) (q35;q14)), which

forms a fusion gene (*PAX3-FKHR*) encoding the N-terminal portion of PAX3 that

contains intact PD and HD together with the C-terminal transactivation domain of

the forkhead transcription factor FKHR (Galili *et al*, 1993). This encodes a

chimeric protein that has greater transcriptional activity than either PAX3 or

FKHR alone (Fredericks *et al*, 1995), and the translocation is an important criteria

15

for ARMS diagnosis (Anderson *et al,* 2001). Although the molecular role of this

fusion gene in ARMS has not been fully uncovered, PAX3-FKHR is suggested to

promote cancer cell proliferation and metastasis of the tumor by regulating PAX3

downstream targets (Ginsberg *et al,* 1998; Keller *et al,* 2004a; Zhang and Wang

2007). In addition, down-regulation of PAX3-FKHR in ARMS cells led to

increased cell death (Bernasconi *et al,* 1996), indicating that PAX3-FKHR plays

an important role in ARMS cell survival. Nevertheless, despite the significance of

this fusion protein in ARMS pathogeneses, it is not sufficient to induce tumor

formation on its own (Lagutina *et al,* 2002), but can cooperate with additional

lesions for tumorigenesis. This was evident upon the generation of an ARMS

mouse model by conditional activation of a *Pax3-Fkhr* knock-in allele (Keller *et*

al, 2004b).

Deregulation of *PAX3* has also been implicated in cutaneous malignant

melanoma (CMM), which is derived from the oncogenic transformation of

melanocytes. Clinically, melanoma is categorized into four stages: atypical nevus,

radial growth phase (RGP), vertical growth phase (VGP), and metastatic growth

phase (MGP), with both RGP and VGP being characterized as primary

melanomas (Smith *et al,* 2005). Although early stage melanoma is highly curable

by surgical procedures, metastatic melanoma has a 5-year survival rate of less

than 10% (Balch *et al,* 2001). It is also one of the few cancers that continue to

increase in incidence and mortality (Mitra and Fisher 2009). PAX3 has been

implicated in melanoma pathology since its expression has been widely observed

in melanoma-derived cell lines and human CMM tissues, but is typically absent in

16

normal differentiated melanocytes (Barr *et al*, 1999; Scholl *et al*, 2001). Although

it was suggested that higher levels of PAX3 indicate a more advanced stage of

melanoma (Ryu *et al*, 2007), a later study observed PAX3 expression through all

stages of melanoma, including benign nevi (Plummer *et al*, 2008). More recent

studies also revealed that PAX3 was not only expressed in nevi and melanoma,

but was also found in epidermal melanocytes in normal skin (Medic and Ziman

2010; He *et al*, 2010b) where it may confer an 'undifferentiated plastic state'

(Medic and Ziman 2010). It should also be noted that the samples in the Medic

study were obtained from Australia, which has more sun exposure than most

regions in the northern hemisphere. In this regard, UV irradiation has been

demonstrated to up-regulate *PAX3* in individual melanocytes in normal human

skin via inhibiting TGF-ß (Transforming growth factor beta), which is secreted by

keratinocytes to block melanocyte proliferation (Yang *et al*, 2008). PAX3

immunoreactivity is also acknowledged to be a useful diagnostic marker because

it can distinguish melanoma from other types of skin cancer (He *et al*, 2010b).

Last, *PAX3* has been suggested to be more significantly related to melanomas at

acral and non-sun-exposed body sites than those parts exposed to UV (Yang *et al*,

2008). Consistent with this, Plummer *et al* proposed that PAX3 is a useful marker

for non-chronic-sun-damaged melanoma (Plummer *et al*, 2008).

In addition to a role in melanoma cell proliferation, PAX3 has been shown

to play a role in melanoma survival because down-regulation of *PAX3* induced

apoptosis (He *et al*, 2005; Scholl *et al*, 2001). At the molecular level, this PAX3-

inducted apoptosis was related to an increase in the levels of the tumor suppressor

17

protein p53 and Caspase 3, which play important roles in programmed cell death

(He *et al*, 2005). It was also suggested that PAX3 has a direct role in apoptosis by

regulating the expression of the anti-apoptotic protein Bcl-XL (Kubic *et al*, 2008;

Margue *et al*, 2000). Likewise, PAX3 is co-expressed with the MET tyrosine

receptor kinase in both human and mouse melanoma cell lines, and PAX3 was

able to up-regulate MET expression directly and indirectly through MITF

(Mascarenhas *et al*, 2010). Because MET is a proto-oncogene that is involved in

cell survival, tumor growth, and metastasis, Mascarenhas *et al* suggested that

PAX3 promotes melanoma progression via MET (Mascarenhas *et al*, 2010). It is

therefore clear that PAX3 can influence several aspects of melanoma pathogenesis

and that more studies will be needed to help us understand their mechanistic basis.

1.4 *Pax3/PAX3* gene structure and isoforms

The murine *Pax3* gene is located on mouse chromosome 1 and is

comprised of ten exons that encode an open reading frame of 1,437 nucleotides

(Goulding *et al*, 1991). The human orthologue is identical in structure and was

mapped to chromosome 2q35 (Tsukamoto *et al*, 1992) (Figure 1-3, top panel).

Subsequently, Tsukamoto *et al* found two *PAX3* isoforms that were generated by

alternative splicing, naming them *PAX3a* and *PAX3b* (Tsukamoto *et al*, 1994).

Continuing this naming convention, Barber and colleagues named the first cloned

and sequenced transcript as *PAX3c*. They also identified three other *PAX3*

transcripts by mining EST databases made from normal human tissues, including

melanocytes, pregnant uterus, and fetal heart (Barber *et al*, 1999). This study also

18

established that the *PAX3* gene structure in fact comprised ten exons, two more

than previously described. Further analyses of human melanoma cell lines

identified two more *PAX3* isoforms, which were designated *g* and *h* (Parker *et al,*

2004). Thus, there is a minimum of seven alternatively spliced *PAX3* transcripts:

PAX3a-e, PAX3g and *PAX3h* (Figure 1-3, bottom panel).

The full length PAX3 protein possesses two DNA-binding domains PD

and HD; an eight-amino-acid conserved region OP (octapepetide), and a carboxyl-

terminal proline/serine/threonine-rich transactivation (TA) domain (Figure 1-3).

Within the *PAX3* gene structure, the two DNA-binding domains are encoded by

exons 2-4 and exons 5-6, respectively: the OP is encoded by exon 5 preceding the

start of the HD; and the c-terminal TA is encoded by exons 7 and 8 (Figure, 1-3).

Both *PAX3a* and *PAX3b* contain only exons 1-4, and they are the only two PAX3

isoforms missing the HD (Figure 1-3, bottom panel). The remaining PAX3

isoforms (PAX3c-PAX3h) possess both DNA binding domains, but differ in their

C-termini (Figure 1-3, bottom panel). PAX3c and PAX3d are encoded by the first

8 and 9 exons, respectively, and *PAX3e* comprises all 10 exons. The last two

isoforms, *PAX3g* and *PAX3h*, involve splicing of exon 7 to exon 9 or exon 10,

respectively. As a result, both transcripts encode a truncated TA domain.

PAX3c and *PAX3d* are the predominant transcripts expressed in most cell

lines and tumor tissues, and *in vitro* studies suggest they have similar DNA-

binding activity (Barber *et al,* 1999). The *PAX3e* isoform is expressed at low

levels in most cell lines and tissues. Although PAX3 has been implicated in CMM,

PAX3e transcripts were undetectable in some murine melanoma cell lines, which
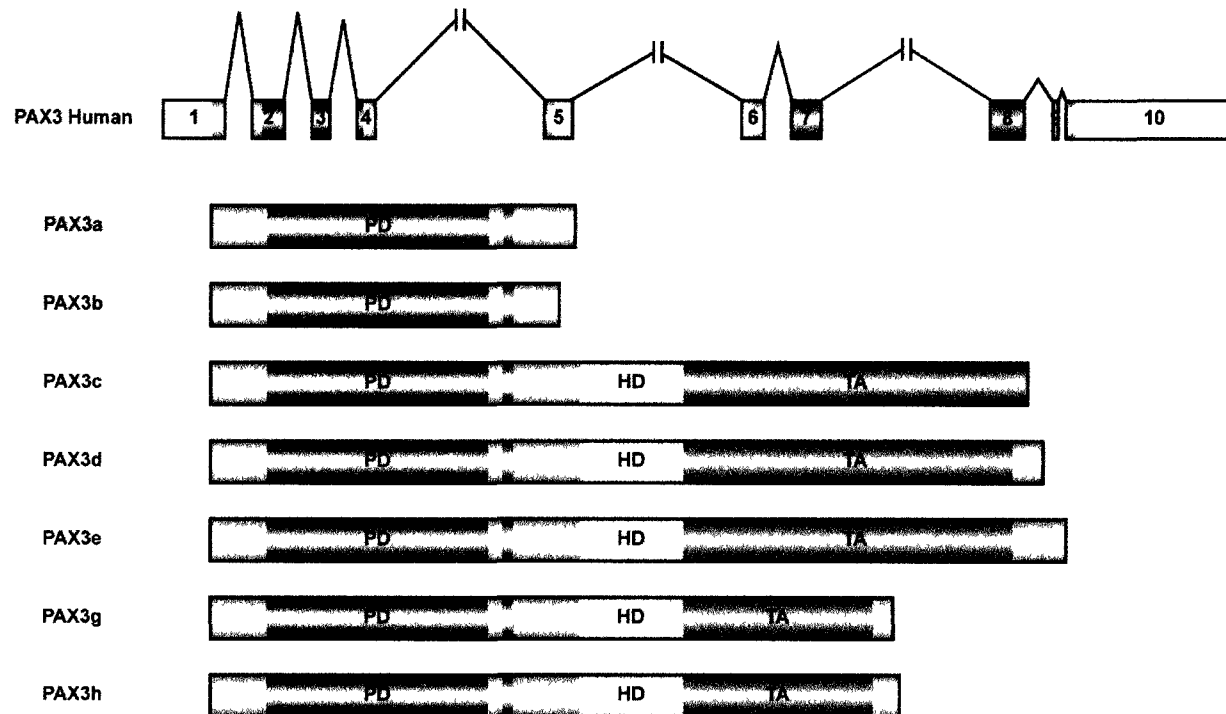
Figure 1-3. PAX3 isoforms (Wang et al, 2006). Top panel: The Human *PAX3* gene contains a total of 10 exons. The white area in the exons denotes untranslated regions. Other colours in the exons are all translated. Solid lines between two exons represent introns. Light blue, orange, yellow and dark blue indicate the regions encoding the paired domain (PD), octopeptide (OP), homeodomain (HD) and c-terminal transactivation domain (TA), respectively. Other transcribed regions are highlighted in grey. **Bottom panel:** Alternative splicing generates 7 isoforms *PAX3a-e*, *PAX3g* and *PAX3h*. The colour scheme is the same as in the top panel.

20

suggested it may have minimal significance in melanoma biology (Parker *et al,*

2004). Analyses of PAX3g did not identify a specific biological activity, but it

may inhibit transcription of PAX3 target genes by competing for binding sites

with other isoforms (Pritchard *et al,* 2003). Although PAX3h has a truncated TA

domain, it was found to have similar activity as PAX3c and PAX3d in promoting

melanocyte proliferation, migration, transformation. and survival (Wang *et al,*

2006). *PAX3b* was co-expressed in all tissues where the predominant PAX3

isoforms were detected, while *PAX3a* was only expressed in skeletal muscle and

the cerebellum (Tsukamoto *et al,* 1994). The presence of *PAX3a* and *PAX3b* was

associated with reduced cell proliferation rates (Wang *et al,* 2006). but their roles

in tumorigenesis are still unclear. It is nevertheless apparent that this collection of

PAX3 isoforms can have distinct biological activities.

1.5   PAX3 regulatory network

PAX3 regulates downstream target genes by binding to specific *cis*

regulatory elements that are recognized by its DNA-binding domains. In this

regard, PAX3 and the related *Drosophila* prd protein were first demonstrated to

interact with the *e5* sequence found upstream of the *Drosophila even-skipped*

gene (Goulding *et al,* 1991). The *e5* sequence contains both 5'-

ACCGCACGATTAG and 5'-CACCGTTCCGCTTC motifs that are recognized

by the HD and PD, respectively, and efficient binding required both domains

(Goulding *et al,* 1991). Subsequently, PCR-based selection approaches were used

to identify optimal binding motifs for the PAX3 PD. the *Drosophila* prd HD, and

for both domains of prd (Chalepakis and Gruss 1995; Wilson *et al*, 1993). The PAX3 PD sequence, called *Nβ'*, contained a core binding sequence of 5'-GTCACGCTT, which is different from that in *e5* (Chalepakis and Gruss 1995). As expected, mutations within the *Nβ'* motif can entirely disrupt the formation of a PAX3-DNA complex (Chalepakis and Gruss 1995). The prd HD motif comprised a palindrome of TAATN$_{2-3}$ATTA (called *P2* or *P3*) to which two homeodomains bound in a cooperative manner (Wilson *et al*, 1993). The use of both prd DNA-binding domains led to the identification of a composite motif (called *PH0*) that contained juxtaposed PD and HD motifs, and supported cooperative binding by the two domains of multiple PAX proteins including PAX3 (Jun and Desplan 1996). The presence of single half HD sequence (ATTA) in the composite motif revealed very limited DNA binding ability, while it produced strong binding specificity together with PD (Underhill *et al*, 1995). Mutagenesis analyses also revealed that loss of one of the two DNA binding domains leads to reduction of PAX3 binding affinity, but mutations in the HD motif led to much less disturbance in DNA binding than mutations in the PD motif (Corry and Underhill 2005).

Through the use of genetic analyses and these optimal binding sequences, a series of downstream PAX3 target genes have been identified in the myogenic and melanocyte lineages, including *Met, Trp-1, Ret, Dct, Mitf* and *Myf5* (Epstein *et al*, 1996; Galibert *et al*, 1999; Lang *et al*, 2000; Lang *et al*, 2005; Watanabe *et al*, 1998; Bajard *et al*, 2006). In each case, PAX3 has been shown to activate or

repress their transcription in reporter gene assays through binding to a specific

Nf3'-like or composite motif (for detail, see section 1.2).

PAX3 was demonstrated to regulate the expression of *Mitf*, the master

regulator of melanocytes by directly binding to a sequence located between 263bp

and 238bp upstream of the transcription start site: 5'-ATTAATACTACTGGAAC.

The ATTAAT and GGAAC (GTTCC on the complementary strand) were

proposed as binding sites for the HD and PD, and WS1 mutations in either

domain were unable to activate an *Mitf*-promoter reporter gene (Watanabe *et al*,

1998). Although this binding site is similar to the Drosophila *e5* sequence, Corry

and Underhill noted there were two suboptimal HD-binding sites near the core PD

site (Corry and Underhill 2005). Amongst these putative HD sites, a site

immediately adjacent to the PD motif had the strongest impact on PAX3 binding

(Corry and Underhill 2005). The *in-vitro* binding data also established that both

the PD and HD recognition sites (5'-TTTAGTTCC) were necessary for efficient

DNA binding (Corry and Underhill 2005). Consequently, binding of the PD and

HD to the *Mitf* promoter occurs in the same head to head fashion as described for

prd and the *PHO* motif.

The PAX3 protein has also been shown to activate the expression of *Myf5*

during myogenesis (see section 1.3). In this case, a PAX3-binding sequence (5'-

AGGCATGACTAATT) within a 145bp region located -57.5kb from the *Myf5*

transcription start site was required for enhancer activity and was bound by PAX3

in *in vitro* binding assays and in ChIP assays *in vivo* (Bajard *et al*, 2006). In

addition, Bajard and his colleagues also demonstrated the importance of this

regulatory element in transgenic mice. The element induced ectopic expression

was similar to the wild-type mouse embryo, but myogenic expression in the

somites and limb buds lost when this sequence was mutated (Bajard *et al*, 2006).

Taken together, these experiments provide strong evidence that PAX3 directly

regulates the expression of *Myf5* in the myogenic lineage through binding to the

target site in this 145bp regulatory element. Analyses thus far have therefore

revealed two types of binding modes for PAX3 recognition of target genes: PD-

only or both the PD and HD together (so-called composite motifs). The functional

significance of these distinct binding modes is not yet clear.

1.6   Objectives

Over the last two decades, the PAX family has been well studied with

regard to their biological significance during embryogenesis and after birth. In this

regard, PAX3 functions in multiple lineages during embryogenesis and postnatal

tissue regeneration, and is associated with several pathological conditions.

Moreover, both *in vitro* and *in vivo* studies have characterized many putative

PAX3 target genes. However, we still do not fully understand the roles that PAX3

plays at the genetic and molecular levels, and how these roles are affected in

PAX3-associated pathologies, such as malignant melanoma. Accordingly, the aim

of my project was to elaborate the PAX3 regulatory network through the

discovery of new PAX3 target genes. To this end, we combined both *in silico* and

in vitro methods to identify potential PAX3 target genes. At the same time, we

investigated the potential role of PAX3 in melanoma by examining the effect of

PAX3 attenuation.

CHAPTER 2 – MATERIALS AND METHODS

2.1 Phylogenetic analysis

Phylogenetic conservation of PAX3 target sites was visualized using the Evolutionary Conserved Regions (ECR) browser (http://ecrbrowser.dcode.org/). This analysis was applied to both published PAX3 target sequences and novel targets identified in this study. In order to access more species-secific sequence data than was available through ECR, these PAX3 target sequences were used to query the whole-genome shotgun (wgs) database using the NCBI nucleotide Basic Local Alignment Search Tool (blastn) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) optimized for highly similar sequences (megablast) with default parameters. The sequences obtained from these blastn searches were aligned with ClustalW (http://www.ebi.ac.uk/clustalw/) using default parameters. The sequences from the ClustalW alignment file were then depicted using WebLogo 2.8.2 (http://weblogo.berkeley.edu/logo.cgi) with default parameters to provide a graphical overview of sequence conservation across PAX3 binding sites.

2.2 Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shifts assays (EMSAs) for the characterization of PAX3 DNA-binding have been previously described (Corry and Underhill 2005). Briefly, oligonucleotides were labeled with α-[32P]-dCTP (3000 Ci/mmol, Amersham) and Klenow fragment of DNA polymerase I (Gibco-BRL, Burlington, Canada). Unincorporated nucleotides were removed using a Nick-Column (Pharmacia-Amersham). Recombinant PAX3 proteins comprising the PD or spanning the PD and HD (PDHD) were expressed in bacteria using the pET21a

27

Table 2-1. Oligonucleotides used for electrophoretic mobility shift assays

a. For Figure 3-1b

Name	Top strand sequence (5'-3')
TYRP1	GAAGGCCAATGTCACACTTGTATTTTCTGTTAG
c-MET	GACTCGGTCCCGCTTATCTCCGG
c-RET	GCCAACCACCATGTCACACTGCCCATGGGAGGG
Dct	GTGCACTTAGGGTCATGTGCTAACAAAGAGGAT
MITF	GTCATCTTTAGTTCCAGTAGTATTAATAGACAA
MYF5	GTACCATGCAATTAGTCATGCTTTTATGATTTA

b. For Figure 3-7

Name	Top strand sequence (5'-3')
MYF5	GTACCATGCAATTAGTCATGCTTTTATGATTTA
MYF5 Pdmut	GTACCATGCAATTAGAGCTGCTTTTATGATTTA
MITF	GCTATTCATCTTTAGTTCCAGTAGTATTAATAG
BOC	GAGAAGGGTAATTAGTCACTGTCACACACTG
PRDM12	GCGGCGGGCAATTTGTCACGCTGTTAAAGTGAT
LBX1/BTRC	GTTTTATTGCATTTGTCACGGTGAAGATTAAAG

Both the top and bottom strands have an overhanging 5' guanine residue that creates a template for {^{32}P}-dCTP incorporation by the Klenow fragment of DNA polymerase I.

system (Novagen) and purified by Ni-NTA-agarose chromatography. EMSAs

were performed in a 20 μL mixture containing approximately 4 pmol {^{32}P}-

labeled oligonucleotide, EMSA buffer (2 mM Tris-HCl (pH 7.5), 10 mM KCl,

1mM DTT, 0.2 mg/mL BSA, and 2% glycerol) and 1 μg p(dI•dC)-p(dI•dC)

nonspecific competitor. Protein-DNA complexes were allowed to form at room

temperature for 30 min and were resolved from free oligonucleotide by

electrophoresis (12 V/cm in 0.25X TBE) on an 8% (for PAX3-PDHD) or 10%

(for PAX3-PD) non-denaturing polyacrylamide gel (29:1

acrylamide:bisacrylamide). Polyacrylamide gels were dried under vacuum and

exposed overnight to Kodak Biomax MR film with an intensifying screen at -

80°C.

2.3   Discovery of potential PAX3 binding sequences

A previous attempt to identify PAX3 targets used a CASTing (Cyclic

Amplification and Selection of Targets) method (Wright and Funk 1993) to

identify mouse and human genomic fragments bound by PAX3 or the PAX3-

FKHR translocation product (Barber *et al*, 2002). These PAX3 CASTing libraries

comprised 1203 human and 1260 mouse PAX3-bound genomic sequences and

were retrieved from the Genome Survey Sequence records (GSS) database at

NCBI for the discovery of over-represented motifs. Each library was analyzed

separately using Weeder (http://159.149.109.9/modtools/) and MEME

(http://meme.sdsc.edu/meme4_1_1/intro.html) following program instructions.

Briefly, repetitive elements in the PAX3 CASTing libraries were masked using

the RepeatMasker Web Server (http://www.repeatmasker.org/cgi-

bin/WEBRepeatMasker) before analysis. For Weeder analysis, large mode was

used for *Homo sapiens* (HS) or *Mus musculus* (MM) to identify over-represented

motifs of 6 to 12bp in length with 1, 2, 3 or 4 nucleotide substitutions tolerated for

motif lengths of 6, 8, 10 and 12 bp, respectively. MEME analysis was run with no

limiting criteria but with an E-value of no more than 0.1. For both MEME and

Weeder, only a single strand was considered and the top 10 scoring over-

represented motifs were recorded.

2.4   Identification of potential PAX3 target sites

2.4.1 Motif scan

A PERL program was written to identify potential PAX3 target sequences

given a user supplied genomic sequence (see Appendix I for the PERL script).

The human genome was used for motif scanning and each chromosome was

considered as a single continuous sequence so that each position had an equal

possibility of being chosen. Based on studies of phylogenetic conservation and

motif discovery in PAX3 CASTing libraries, the composite PAX3 binding site

(ATT.GTCA[C|T]G[C|G]T and its complementary sequence) were used to search

the repeat masked human genome (NCBI37 assembly 50 release, retrieved from

the Ensemble FTP site (ftp.ensembl.org)). The dot in the searched motif

represents any of four nucleotides – empirical analyses suggested this position

was neither conserved nor important in conferring efficient DNA-binding to

PAX3. The C|T and the C|G in square brackets allow for C or T and C or G at

these corresponding positions, both of which support efficient PAX3 DNA-

binding. Motifs were recovered together with 30bp of flanking sequence on both

the 5' and 3' sides. This generated a 72bp sequence that provided a unique

sequence tag so that putative target sequences could be resolved in BLAST

searches. In addition, the presence of the flanking region also provided an

opportunity to explore potential selective pressure on sequences upstream or

downstream of the PAX3 motif.

2.4.2 BLAST annotation

Mouse genomic sequences recovered from the motif scan were first

filtered using the nucleotide Basic Local Alignment Search Tool (nucleotide blast)

(http://blast.ncbi.nlm.nih.gov/Blast.cgi) against the Human genomic plus

transcript database and optimized for megablast. The regions that did not have a

corresponding human homologue were removed from subsequent analyses. In

parallel, the respective orthologues of human and mouse sequences in the PAX3

CASTing libraries were identified using a cross-species blastn analysis of the

mouse and human genomic plus transcript databases. BLAST sequence hits with

85% or greater identity were collected and their degree of conservation was

examined further as described in section 2.1.

Table 2-2. Primers for generating an RNAi-resistant PAX3 expression plasmid

Name	Original sequences (5' – 3')	Oligonucleotide sequences (5' – 3')	Restriction site
PAX3_RNAiR_F	gct gcc ccc agg **atg** acc acg ctg g	gcc GGA TCC agg **atg** acc acg ctg g	*Bam*HI
PAX3_RNAiR_R	g cga att **cta** gaa cgt cca agg ctt act ttg tcc ata ctg ccc ata ctg gta gcc	g cGA ATT **Cta** gaa cgt cca agg ctt gct ctg ccc gta ttg ccc ata ctg gta gcc	*Eco*RI

NOTE: The forward primer (PAX3_RNAiR_F) is located in exon 1 and the reverse primer (PAX3_RNAiR_R) is located in exon 8 of *Pax3*. Capitalized nucleotides in the sequences are restriction sites. Bolded nucleotides are start and stop codons. Underlined nucleotides represent changes required to generate the RNAi-resistant sequence and do not alter coding potential.

32

2.5 Plasmid constructs

Plasmids encoding hexahistidine-tagged PAX3 PD and PDHD proteins were made by inserting PCR-amplified fragments into the pET21a expression plasmid (Novagen) and have been described previously (Corry and Underhill 2005). The GFP tagged wild-type PAX3 expression plasmid was generated by cloning the full-length PAX3 open reading frame (479 aa) into pEGFP-N1 (Clontech) as previously described (Corry *et al*, 2008). This construct was used as a PCR template to make an RNAi-resistant wild-type PAX3 expression plasmid, denoted as RNAiR-PAX3. The primer set used for PCR is shown in Table 2-3. PCR reactions were carried out in a total volume of 20 μL with 50ng of DNA template, 0.5 μM of each primer, 1.0 mM dNTPs, 0.2 μL of PfuTurbo DNA polymerase (Stratagene), and 2 μL of 10X PfuTurbo polymerase buffer (Stratagene). Thirty PCR cycles were performed and comprised a 30 second denaturing step at 94°C, a 30 second primer annealing step at 63°C, and a 2 minute primer extension step at 72°C. The *Eco*RI-*Bam*HI digested PCR product was inserted into the pcDNA3.1/Hygro vector (Invitrogen) that had been sequentially digested with *Bam*HI and *Eco*RI restriction enzymes. Ligation was achieved using T4 DNA ligase (NEB) at room temperature for 2 hours.

2.6 Expression and purification of recombinant PAX3 proteins

Expression and purification of pET21a-expressed PD and PDHD PAX3 proteins has been previously described (Corry and Underhill 2005). Briefly, pET21a expression plasmids were transformed into BL21DE3 *E. coli* and a single

33

colony was used to inoculate 2 mL Luria-Bertani (LB) medium containing 5

μg/mL ampicillin. The culture was grown overnight at 37°C on an orbital shaker

at 250 rpm. A 1250 μL aliquot of this culture was used to inoculate 50 mL LB

medium (with 5 μg/mL ampicillin) and grown at 37°C with shaking (250rpm)

until an OD_{600} of 0.6 was reached. Expression of recombinant protein was induced

by adding 50 μL of 100 mM IPTG to the culture followed by 1hr incubation at

37°C with agitation (250rpm). Cells were then collected by centrifugation for 20

min at 4°C. The pellet was resuspended in a mixture of lysis buffer (50 mM

NaH_2PO_4, 300 mM NaCl, 10 mM imidazole), 100 mM

phenylmethanesulphonylfluoride (PMSF), and protease inhibitor cocktail (Sigma).

Lysozyme (1 mM) was added to a final concentration of 60 μM and the sample

was incubated on ice for 30 min. The lysate was sonicated twice for 30 s with a 10

s cooling period between each pulse, followed by centrifugation for 20 min at 4°C.

The supernatant was collected in a fresh tube and recombinant proteins were

purified using affinity chromatography on Ni-NTA agarose beads according to the

manufacturer's instructions (Qiagen). The concentration of purified protein was

determined using the Bradford method.

2.7   DNA preparation of recombinant plasmids

2.7.1 Mini-prep of DNA for sequencing

The RNAiR-PAX3 plasmid construct was transformed into DH5α and

single colonies were grown overnight in 2 mL LB medium containing 5 μg/mL

ampicillin at 37°C on a rotary shaker at 250 rpm. The plasmid DNA was isolated

and purified using QIAprep Spin Miniprep Kit according to the manufacturer's

instructions (Qiagen). Sequencing was done by the Applied Genomics Centre

(TAGC) at the University of Alberta.

2.7.2 Midi-prep of DNA for transfection

For transfection, the RNAiR-PAX3 plasmid DNA was recovered using a

QIAGEN Plasmid Midi Kit with the following modifications. A 100 mL

overnight bacterial culture was harvested by centrifugation at 6000xg for 15min at

4°C. The pellet was resuspended, lysed, and precipitated as specified in the

manufacture manual. The protein precipitate was removed by centrifugation at

8000rpm for 1hr at 4°C, followed by an additional 20min spin to remove

excessive protein precipitate. The supernatant was passed over an equilibrated

QIAGEN-tip 100 and the DNA was eluted as per the user manual. The DNA was

collected by centrifuging at 8000rpm for 1hr at 4°C and then washed with 1mL of

70% room temperature ethanol. After a 10min spin at 13,000rpm and 4°C, the

DNA pellet was resuspended in sterile ddH$_2$O and its concentration was measured

using an Eppendorf BioPhotometer.

2.8   Cell lines and culture

The murine B16F10 melanoma cell line was obtained from the American

Type Culture Collection and maintained in Dulbecco's Modified Eagle's Medium

(DMEM) supplemented with 2 mM L-glutamine and 10% fetal bovine serum at

37°C with 5% CO$_2$.

2.9 Transfection of small-interfering RNA

Three small-interfering RNA (siRNAs) targeting three different sites in the 3' region of PAX3 transcripts and an AllStars negative control were purchased from QIAGEN (sequences are listed in Table 2-2). BLOCK-iT (Invitrogen) is a fluorescein-labeled double-stranded RNA oligomer that was used to monitor transfection efficiency by cotransfection with siRNAs. Briefly, one day before transfection, 2.0×10^4 B16F10 cells in 500 μL growth medium were seeded in each well of a 24-well plate with or without a glass coverslip. Dilutions of 3 pmol siRNA in 50 μL Opti-MEM and 1 μL Lipofectamin RNAiMAX transfection reagent (Invitrogen) in 50 μL Opti-MEM were prepared separately for each well. These were incubated for 10 minutes at room temperature before being mixed together. After another 10-minute incubation at room temperature, the above transfection complex was added drop-wise to the 24-well plate. Cells were harvested 24, 48 or 72 hours after transfection for western blot, immunofluorescnece or flow cytometry analyses.

2.10 Paclitaxel treatment

Paclitaxel (PAC) (Sigma) treatment was applied to further characterize the effect of PAX3 knockdown. PAC (20 mM) was added to cells to a final concentration of 25 nM per well 24 or 48 hours following transfection of *Pax3* siRNA into B16F10 cells as described in section 2.9. Cells were harvested for FACS analysis after a 24-hour incubation.

2.11 Cotransfection of siRNA and plasmid DNA

Separate mixtures containing 12.5 pmol of AllStar Negative Control

siRNA, 12.5 pmol of Mm_PAX3_4_HP_siRNA, 12.5 pmol of

Mm_PAX3_4_HP_siRNA with 100 ng of RNAiR-PAX3 plasmid DNA, or 12.5

pmol of Mm_PAX3_4_HP_siRNA with 1000 ng of RNAiR-PAX3 plasmid DNA

were gently mixed in 500 μL of DMEM. A 5 μL aliquot of Lipofectamin

RNAiMAX transfection reagent (Invitrogen) was then added to the above mixture

and incubated at room temperature for 10 minutes to allow formation of

transfection complexes. Cells were transfected with 5 μL of Lipofectamin

RNAiMAX reagent as an additional control. The transfection complex was

transferred to the well followed by the addition of 5.0×10^5 B16F10 cells in 2 mL

of growth medium (reverse transfection). The plate was gently mixed and

incubated at 37°C with 5% CO_2 for 24 hours before harvesting for western blot

and flow cytometry analyses.

2.12 Antibodies

The PAX3-specific mouse monoclonal antibody was purchased from the

Developmental Studies Hybridoma Bank (DSHB; NICHD/University of Iowa)

and was used for immunofluorescence (1:200), western blot (1:1000) and FACS

(1:100) analyses. Monoclonal anti-GAPDH antibody (Sigma-Aldrich) was used at

a dilution of 1:10,000 as a loading control for western blotting.

Table 2-3. Sequences of Qiagen siRNAs targeting PAX3

Name	Target sequences (5' – 3')
Mm_PAX3_1_HP_siRNA	aaccggattgtcagtaggtaa
Mm_PAX3_3_HP_siRNA	ctcaagccagatattgcataa
Mm_PAX3_4_HP_siRNA	atgggcagtatggacaaagta
AllStars Negative Control	cagggtatcgacgattacaaa

2.13 Whole cell extracts

2.13.1 B16F10 cells in 24-well plates

At 24, 48 or 72 hours post-transfection, cells were washed once with cold

1X PBS; trypsinized in 200 μL trypsin; and diluted into 800 μL growth medium.

Detached cells were then transferred to a 1.5mL centrifuge tube and collected by

centrifugation at 200xg for 5 minutes. The supernatant was removed and the cell

pellet was resuspended in 40 μL of 1X sodium dodecyl sulfate (SDS) sample

buffer with 5% ß-mercaptoethanol and then stored at -80°C for later use.

2.13.2 B16F10 cells in 6-well plates

After a 24-hour transfection, cells were washed once with cold PBS. Upon

addition of 250 μL lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl and 0.5%

IGEPAL) containing 12.5 μL mammalian Protease Inhibitor Cocktail (PIC)

(Sigma-Aldrich) and 2.5 μL of 100 μM PMSF, cells were scraped off the plate

and collected to a 1.5 mL centrifuge tube. The lysate was sonicated twice for 10 s

at amplitude level 2 with a 5 minute cooling period between each cycle. The tube

was then kept on ice for 30 minutes followed by centrifugation at 13,000rpm for

10 minutes at 4°C. The supernatant was collected and stored at -80°C.

2.14 Western blot

For samples extracted from 6-well plates, approximately equal amounts

(15-20 μg) of protein were denatured at 95-100°C in 10 μL of 1X SDS sample

buffer with 5% ß -mercaptoethanol for 5 minutes before being loaded onto a 10%

SDS polyacrylamide gel with a 4% stacking gel. For whole cell protein samples

extracted from 24-well plates, protein concentrations were not measured due to the low amount of protein yield, so a 20 μL volume was used for immunoblotting. The amount of protein contained in each sample was normalized to the GAPDH loading control. Proteins were resolved by SDS-PAGE at 150V for approximately 60 minutes at room temperature and were then transferred onto a 0.45 μm nitrocellulose membrane (Bio-Rad) at 100V for 90 minutes at 4°C. The membrane was blocked in 5% skim milk in TBST (25 mM Tris-HCl pH 8.0, 0.15 M NaCl, 0.05% Tween) (milk/TBST) for 2 hours at room temperature with gentle agitation. Blots were incubated with anti-PAX3 primary antibody (1:1000 in milk/TBST) overnight at 4°C on a rocking platform, followed by four washes with TBST for 5 minutes each. Blots were then incubated with an anti-mouse secondary antibody conjugated to peroxidase (Cedarlane) (1:10,000 milk/TBST) for 1 hour at room temperature on a rotary shaker, followed by four TBST washes of 5 minutes each. Blots were rinsed twice with TBS before they were treated with Enhanced Chemiluminescence (ECL) reagent (Amersham) for 5 minutes and exposed to Fuji Super RX film for PAX3 detection. The same blot was rinsed with TBS and re-blocked in 5% milk/TBST for 1 hour at room temperature with agitation. Blots were incubated with GAPDH primary antibody (1:10,000 in milk/TBST) for 1 hour at room temperature, followed by incubation with anti-mouse secondary antibody and then developed as described above.

2.15 Immunofluorescence

Twenty-four hours following transfection, cells were rinsed 3 times with

PBS and fixed with 4% paraformaldehyde/PBS for 5 minutes at room temperature,

followed by three 5-minute PBS washes. Cells were permeablized with 0.1%

Triton X-100 in PBS, followed by 3 PBS washes. Cells were blocked in 5% BSA

in PBS (BSA/PBS) for 1 hour at 4°C and washed 3 times with 1% BSA/PBS for 5

minutes. Anti-PAX3 primary antibody was diluted in 1% BSA/PBS (1:200) and

applied to cells for 1 hour at room temperature. Cells were then washed 3 times

with 1% BSA/PBS for 5 minutes each time and underwent a 1hr incubation with a

fluorescein isothiocyanate (FITC)-conjugated anti-mouse secondary antibody

(Sigma) diluted in 1% BSA/PBS (1:400). Cells were washed 3 times with PBS for

5 minutes and then incubated with 1 μg/mL Hoechst 33258 (Sigma) (1:2500 PBS)

for 5 minutes. Coverslips were rinsed twice with PBS and then mounted onto

glass slides using 20 μL of Mowiol mounting medium.

Imaging was performed using a Leica DMRE fluorescence microscope

equipped with a monochrome 10-bit CCD camera (Qimaging) and captured using

Northern Eclipse v6.0 (Empix Imaging Inc.). Images to monitor transfection

efficiency were acquired using MetaMorph 7.5 (Molecular Devices) on a Zeiss

Axioplan 2 optical microscope equipped with a 12-bit CoolSnap HQ CCD camera

(Roper Scientific Inc.).

2.16 Fluorescence Activated Cell Sorting (FACS)

At 24, 48 or 72 hours post-transfection, cells were washed once with cold

PBS. Cells were then trypsinized using 200 μL trypsin for 3 minutes with slight

tapping and inactivated by addition of 800 μL growth medium. Detached cells

were transferred to a 1.5 mL centrifuge tube and collected by centrifugation at

200xg for 5 minutes. The supernatant was removed and cells were resuspended in

20 μL of 10X PBS and immediately fixed with 180 μL cold 70% ethanol in a

polypropylene V-bottom 96-well plate overnight at -20°C. Fixed cells were

collected by centrifugation at 200xg for 5 minutes. The cell pellet was

resuspended in 200 μL of 1X PBS and a 5 minute centrifugation at 200xg was

again applied to collect cells. The cells were then stained with 200 μL PI/PBS

solution (10% Triton X-100, 2 μg/μL RNase and 0.02 μg/μL Propidium Iodide

(PI) (Invitrogen)). Cells were transferred into 5 mL round-bottom polystyrene

FACS tubes and DNA content was analyzed on a FACSCalibur Flow Cytometer

(BD) equipped with a Doublet Discrimination Module (DDM). Data files were

analyzed using CellQuestPro software.

CHAPTER 3 – PHYLOGENETIC AND GENOMIC ANALYSES OF PAX3 BINDING SITES

Figure 3-8 was prepared by N. Hu and A. Underhill.

3.1 INTRODUCTION

PAX3 is a member of the paired box family of sequence-specific DNA-binding proteins and functions in both neural crest and myogenic lineages (Goulding *et al*, 1991; Wu *et al*, 2008). Insufficient PAX3 levels in these lineages cause congenital defects, while deregulated PAX3 activity has been associated with malignancy, notably in melanoma and alveolar rhabdomyosarcoma (Barr *et al*, 1995; Biegel *et al*, 1995; van den Broeke *et al*, 2006; Scholl *et al*, 2001). An essential step in understanding the pathogenic role of PAX3 in these cases requires identifying its transcriptional targets. To date, a number of putative PAX3 regulated genes have been proposed in the neural crest and myogenic lineages, including *MITF* (Watanabe *et al*, 1998), *Tyrp1* (Galibert *et al*, 1999), *Dct* (Lang *et al*, 2005), *MET* (Epstein *et al*, 1996), *RET* (Lang *et al*, 2000) and *Myf5* (Bajard *et al*, 2006). In these cases, PAX3 target recognition occurs through two distinct modes, exemplified by interaction of both DNA-binding domains with composite motifs in the *MITF* and *MYF5* genes, but only the paired domain for all other target sites.

We employed a combination of phylogenetic and genomic analyses to query the conservation and architecture of binding sites in existing PAX3 target genes, together with analyses of DNA-binding *in vitro*. Surprisingly, only the *MITF* and *MYF5* elements met both criteria of phylogenetic conservation and efficient DNA-binding. Moreover, mining of pre-existing databases that are comprised of PAX3-bound mouse and human genomic sequences revealed a bias towards composite binding sites. The conserved features of the composite binding

44

motifs provided a sufficient level of complexity to undertake an *in silico* search

for putative PAX3 targets. This identified the gene for *LBX1*, a homeodomain

protein that has been demonstrated to function in an autoregulatory loop with

PAX3 in the myogenic lineage (Mennerich and Braun 2001); *Brother of CDO*

(BOC), which encodes an Immunoglobulin/fibronectin domain family member

that, like PAX3, functions in myogenesis and neural patterning; and *PRDM12*, a

candidate tumor suppressor. Together, these data support the functionality of

composite binding motifs and reveal their presence in several genes relevant to

understanding PAX3 biology and its role in cancer.

3.3 RESULTS

3.2.1 Assessing evolutionary conservation and relative PAX3 binding affinity of representative PAX3 target sequences

It is widely believed that functionally important regions in the genome evolve slower than neutral ones (Prabhakar *et al*, 2006) so that they are more conserved across multiple species. Nevertheless, sequence conservation of PAX3 binding sites has not been systematically analyzed and was therefore undertaken in this study. Phylogenetic conservation of genomic regions comprising published PAX3 binding sites was broadly analyzed using the Evolutionary Conserved Region (ECR) Browser, which is an online graphical interface designed to visualize genomic conservation in up to 13 vertebrate genomes from human to fish (Ovcharenko *et al*, 2004). The ECR result illustrated that reported PAX3 binding sites in the *RET* and *Tyrp1* genes were not conserved between human and murine genomes (Figure 3-1a). Although the PAX3 PDHD protein was able to bind to the *Tyrp1* site *in vitro*, it interacted only weakly with the *RET* element (Figure 3-1b). Similarly, the *MET* sequence was poorly conserved and we could not detect its interaction with PAX3 in EMSAs (Figure 3-1). In addition, the EMSA results also indicated that the PAX3 PDHD protein failed to bind to the target site in the *Dct* promoter, even though it displayed a higher degree of conservation (Figure 3-1). As a result, only the regulatory elements in the *MITF* and *Myf5* genes were both efficiently recognized by the recombinant PAX3 PDHD protein in EMSAs and highly conserved from primates to frog (Figure 3-1).

a.

| MITF | MYF5 | DCT | MET | RET | TYRP1 |

b.
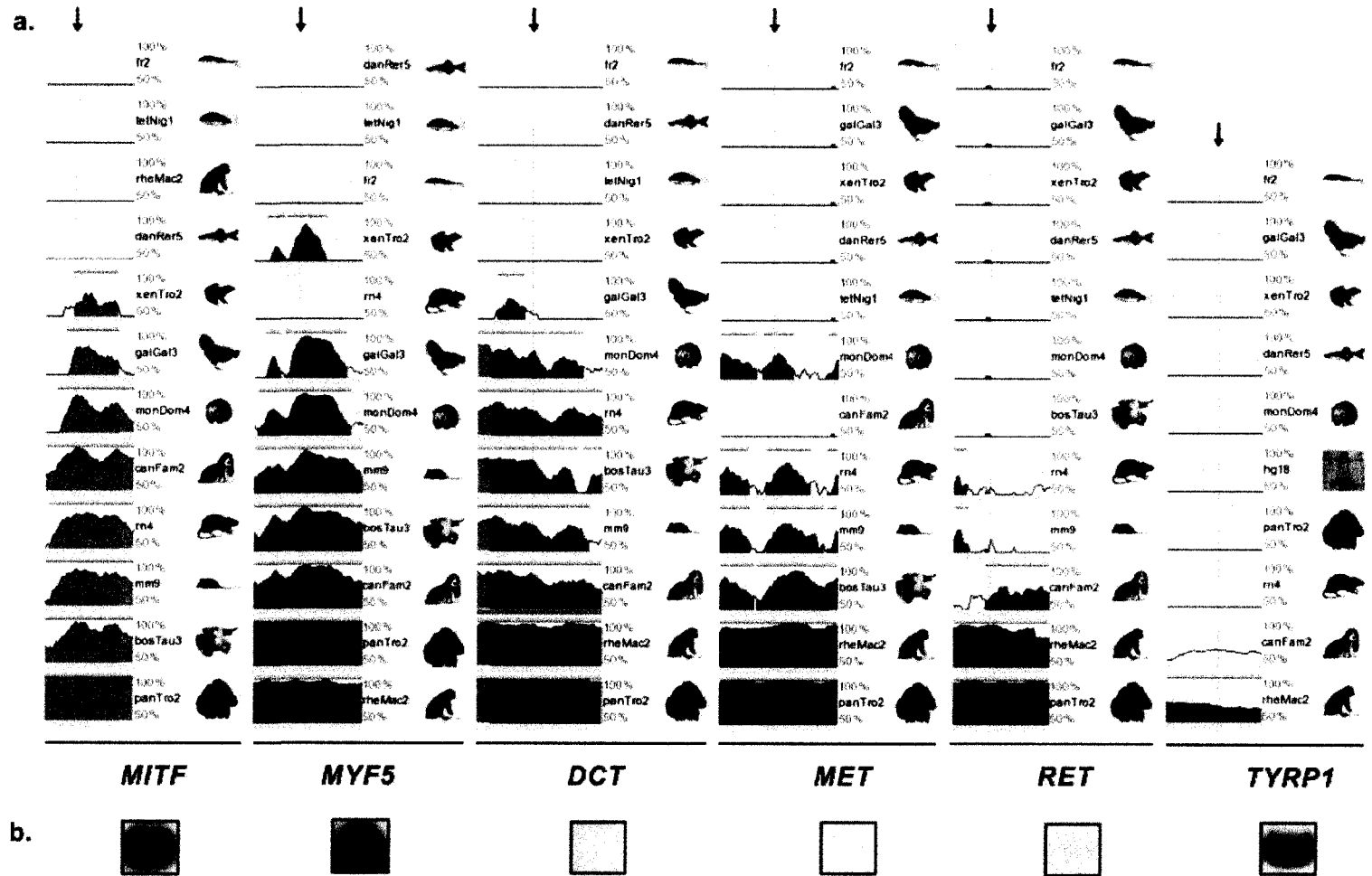
Figure 3-1. Overview of evolutionary conservation and binding affinity of representative PAX3 target sites. (a) ECR view of each putative PAX3 binding site. Black arrows on the top and red vertical lines indicate the center of each PAX3 binding site. The level of conservation is represented graphically according to percentage conservation on the right edge of each ECR output. Intronic regions are highlighted in pink; intergenic regions are highlighted in red; yellow and green regions are UTRs and repetitive elements, respectively. The species are annotated on the right side of each profile: bosTau3, *Bos taurus*; canFam2, *Canis familiaris*; danRer5, *Danio rerio*; fr2, *Takifugu rubripes*; galGal3, *Gallus gallus*; mm9, *Mus musculus*; monDom4, *Monodelphis domestica* ; rheMac2, *Macaca mulatta*; rn4, *Rattus norvegicus*; tetNig1, *Tetraodon nigroviridis*; xenTro2, *Xenopus tropicalis*; and panTro2, *Pan troglodytes*. The corresponding gene names are listed on the bottom. **(b)** EMSA results of each PAX3 target sequences using the PAX3 PDHD protein.

a. Composite type of motifs

MIITF

Homo sapiens
Pan troglodytes
Macaca mulatta
Otolemur garnettii
Canis familiaris
Loxodonta africana
Felis catus
Cavia porcellus
Tupaia belangeri
Dasypus novemcinctus
Bos taurus
Equus caballus
Myotis lucifugus
Echinops telfairi
Ornithorhynchus anatinus
Erinaceus europaeus
Sorex araneus
Monodelphis domestica
Mus musculus
Rattus norvegicus
Spermophilus tridecemlineatus
Gallus gallus

MYF5

Homo sapiens
Pan troglodytes
Macaca mulatta
Ornithorhynchus anatinus
Canis familiaris
Bos taurus
Equus caballus
Myotis lucifugus
Sorex araneus
Spermophilus tridecemlineatus
Tupaia belangeri
Monodelphis domestica
Rattus norvegicus
Mus musculus
Gallus gallus
Erinaceus europaeus

b. PD only motifs

TYRP1

Mus musculus
Rattus norvegicus

MET

Homo sapiens
Pan troglodytes
Macaca mulatta
Microcebus murinus
Spermophilus tridecemlineatus
Dasypus novemcinctus
Bos taurus
Otolemur garnettii
Ornithorhynchus anatinus
Erinaceus europaeus
Myotis lucifugus
Rattus norvegicus
Mus musculus
Monodelphis domestica
Sorex araneus

DCT

Homo sapiens
Pan troglodytes
Macaca mulatta
Tupaia belangeri
Canis familiaris
Myotis lucifugus
Oryctolagus cuniculus
Equus caballus
Dasypus novemcinctus
Bos taurus
Otolemur garnettii
Spermophilus tridecemlineatus
Echinops telfairi
Loxodonta africana
Erinaceus europaeus
Mus musculus

RET

Homo sapiens
Pan troglodytes
Macaca mulatta
Spermophilus tridecemlineatus
Otolemur garnettii
Rattus norvegicus
Oryctolagus cuniculus

49

Figure 3-2. Conservation analysis of PAX3 target sequences in higher eukaryotes. Alignment of blast hits of putative PAX3-binding regions of (a) composite type motifs in *MITF* and *MYF5* and (b) PD only motifs in *Tyrp1*, *MET*, *DCT* and *RET* promoters among higher eukaryotes. Sequences were aligned using ClustalW and conserved nucleotides are highlighted in black with BoxShade 3.21. Consensus sequences were derived via WebLogo 2.8.2. The HD and PD recognition motifs are highlighted in blue and red boxes, respectively.

50

It was noted that the ECR browser lapsed in annotating the *Myf5* promoter in *Rattus norvegicus* even though manual recovery of this region supports its conservation (Figure 3-2a). Because of this potential pitfall of the ECR algorithm and the ceiling on the number of genomes analyzed, we further examined the phylogenetic conservation of PAX3 target sequence via BLASTN, followed by sequence alignment using ClustalW (Figure 3-2). Similar to the ECR output, the *MITF* and *MYF5* promoter sequences had the best conservation across species, comprising available genomic sequences from 22 and 17 species, respectively (Figure 3-2a). This analysis also supported conservation of the PAX3 target site in the *DCT* gene and confirmed the lack of conservation for *Tyrp1*, *MET* and *RET* (Figure 3-2b). Both the *MITF* and *MYF5* binding sites consist of a PD binding motif (GTTCC or GTCAT) and a juxtaposed HD recognition site (ATTA) at the 5' side. The conservation of these features and their efficient binding by PAX3 suggest that the composite binding motif may be more functionally relevant than the PD-only motif in the remaining target genes.

3.2.2 *In silico* analyses of PAX3 CASTing libraries

To further investigate the nature of PAX3 binding motifs, we mined pre-existing PAX3 DNA-binding data. Previous studies aimed at the identification of PAX3 regulated genes led to the generation of PAX3 CASTing libraries using a Cyclic Amplification and Selection of Targets technique (Barber *et al*, 2002). These libraries contain human and mouse genomic fragments that were bound *in vitro* by human and mouse PAX3 protein. We therefore expected that nucleotide

51

sequences recovered with high frequency in the libraries would coincide with the

highest affinity sites for PAX3. To this end, we retrieved 1203 human and 1260

mouse genomic fragments from the PAX3 CASTing libraries and subjected them

to analyses to identify over-represented motifs. Each library was masked using

RepeatMasker (http://repeatmasker.org) to avoid the selection of repetitive

elements, which are themselves over-represented in the genome. Two motif

finding programs, Weeder (Pavesi *et al*, 2004) and MEME (Bailey and Elkan

1994), which use distinct algorithms, were employed in parallel to search for

potential PAX3 binding sequences. MEME gives a statistical p-value to indicate

the confidence level of each over-represented motif, while Weeder uses a different

scoring system. Because it was not possible to evaluate the statistical significance

across the two data sets, the top ten over-represented sequences identified using

Weeder were collected for both mouse and human datasets, and compared with

the most statistically significant motifs (p-value < 0.05) found by MEME (Table

3-1).

Upon analysis of the mouse CASTing library, the over-represented motif

identified by both Weeder and MEME contained a composite-type PAX3 binding

site, with the exception of the last motif found by MEME (Table 3-1a). This can

be illustrated using WebLogo to depict the consensus of these sequences (Figure

3-3a). For reference, a numbering scheme was used such that the HD and PD sites

are denoted by -4 to -1 and +1 to +5, respectively; bases before the HD and after

the PD motif are numbered in sequence as negative or positive. In addition to the

clear presence of HD (ATTA) and PD (GTCAT) core binding motifs, positions -5,

52

+6 and +7 also showed evidence of selection. In contrast to the mouse library,

analysis of human PAX3 CASTing sequences did not identify a clear consensus

(Table 3-1b; Figure 3-3b). Although the sequences recovered by MEME analysis

had highly significant p-values, the scores assigned by Weeder were lower than

those obtained using the mouse library. The latter, however, was, characterized by

multiple HD half sites (TAAT or ATTA). Nevertheless, using a manual search of

composite binding motifs in the human CASTing library found they occurred at

similar frequencies to the motifs identified by MEME (Table 3-2). The failure to

select this motif may reflect its shorter length when compared to the longer motifs

identified by MEME. Alternatively, the human CASTing library may be

characterized by greater background noise, possibly due to higher level of non-

specific binding *in vitro*. Together with the previous analyses, these results further

solidify the significance of the composite type of PAX3 binding sites.

3.2.3 Identification of potential PAX3 target genes *in silico*

The phylogenetic and genomic sequence analyses described above both

support the importance of the composite-type PAX3-binding motif and provide a

rationale to use the underlying sequence features to predict potential PAX3

binding sites in the genome. Thus, a Perl script was written to search for

composite-type PAX3 target sequences in genomic DNA and output these

sequences in FASTA format (see Appendix I for the Perl script). For this, we

focused on the *Myf5*-like binding element because only *Myf5* target site has been

demonstrated to be bound by PAX3 *in vivo* (Bajard *et al*, 2006). It also displayed

53

Table 3-1. Over-represented motifs in PAX3-CASTing libraries

a. Over-represented motifs among sequences from the mouse PAX3-CASTing library

Weeder (score)	MEME (p-value)
ACCATGACTAAT (0.78)	ATCAATTAGTCATGG (5.30E-10)
AGCATGACTAAT (0.73)	AACCATGACTAAT (6.30E-07)
ATTAGTCATGGT (0.71)	GGGGGTGGGGGTGGG (1.30E-04)
CCATGACTAATT (0.71)	-
AATTAGTCATGG (0.7)	-
CAATTAGTCATG (0.68)	-
TAATTAGTCATG (0.66)	-
ATTAGTCATGGA (0.66)	-
TCAATTAGTCAT (0.66)	-
TCCATGACTAAT (0.66)	-

b. Over-represented motifs among sequences from the human PAX3-CASTing library

Weeder (score)	MEME (p-value)
TTAATTGCTATT (0.31)	TGGTGTAACCGAGCATCTAC (1.8E-616)
AAATCAATCATT (0.31)	CGTTTGGGAACCGTGGGCAC (2.1E-502)
AATCATTTACTT (0.29)	ACTCAGGCCGTCCTGGTTAC (1.9E-543)
CTAATTAGAAAT (0.27)	ACGTAACCGGGTGTTGGTGG (1.3E-513)
AATCAATCATTT (0.26)	TGAAGTAATTCAGACAGTGA (2.4E-508)
ACAAATTAATTA (0.26)	ATCTATCGTCACCAGAACCG (8.4E-446)
TTATTACTAATT (0.24)	GCTTGGTCGTCCCCAGAAGT (1.7E-402)
TCAATTAATTGT (0.24)	TGTTGGTTTATGACTGTAGT (1.8E-373)
ATCATTTACTTT (0.24)	GTCTAGCTTTAAAAGGGCTT (5.9E-373)
TGTCATGGTTTT (0.24)	GGAAGCCAACGTGGCCTGAT (1.0E-344)

54
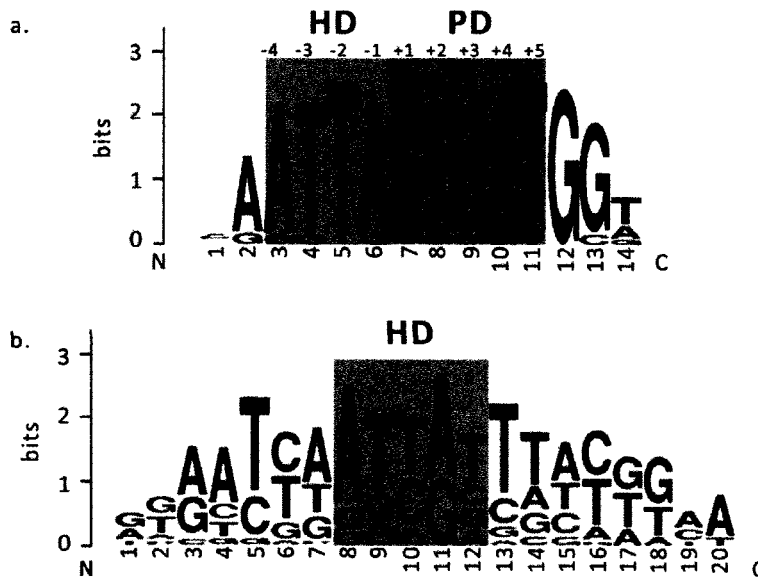
Figure 3-3. Consensus motifs of over-represented PAX3 binding sequences.
Consensus logo of over-represented (a) mouse and (b) human motifs given by
Weeder and MEME. HD- and PD-recognition regions are highlighted in red and
blue, respectively. The nucleotides in the HD and PD motifs are numbered with
negative and positive integers, respectively.

Table 3-2. Frequencies of over-represented sequences in human and mouse CASTing libraries using manual motif search

	Motifs	Human	Mouse
Composite-type motif used for the Motif Scan	ATT.GTCAC	6.98%	2.22%
	GTGAC.AAT	6.90%	3.17%
	ATT.GTCAT	2.00%	3.17%
	ATGAC.AAT	1.83%	3.97%
Over-represented motifs in human CASTing library given by MEME	TGGTGTAACCGAGCATCTAC	7.07%	0.00%
	CGTTTGGGAACCGTGGGCAC	5.90%	0.00%
	ACTCAGGCCGTCCTGGTTAC	5.99%	0.00%
	ACGTAACCGGGTGTTGGTGG	5.99%	0.00%
	TGAAGTAATTCAGACAGTGA	6.07%	0.00%
	ATCTATCGTCACCAGAACCG	5.40%	0.00%
	GCTTGGTCGTCCCCAGAAGT	4.90%	0.00%
	TGTTGGTTTATGACTGTAGT	3.08%	0.00%
	GTCTAGCTTTAAAAGGGCTT	4.57%	0.00%
	GGAAGCCAACGTGGCCTGAT	3.82%	0.00%
Over-represented motifs in human CASTing library given by Weeder	TTAATTGCTATT	0.17%	0.00%
	AAATCAATCATT	0.08%	0.00%
	AATCATTTACTT	0.42%	0.00%
	CTAATTAGAAAT	0.33%	0.00%
	AATCAATCATTT	0.08%	0.00%
	ACAAATTAATTA	0.08%	0.00%
	TTATTACTAATT	0.25%	0.00%
	TCAATTAATTGT	0.17%	0.00%
	ATCATTTACTTT	0.50%	0.08%
	TGTCATGGTTTT	0.58%	0.00%

56

the greatest level of phylogenetic conservation, and it resembled the consensus

motif derived from analysis of the mouse PAX3 CASTing library (Figure 3-3a).

Previous experience has indicated that position -1 does not influence DNA

binding so any nucleotide was allowed at this position in our genome scan. In

addition, compared to the optimal PD binding sequence GTCAC, both the PD

portion of the *Myf5* promoter and the over-represented motif in the mouse

CASTing library showed a C to T base substitution at position +5 (Figure 3-2a;

Figure 3-3a). Both nucleotides (C and T) were therefore allowed at this position.

Moreover, according to the consensus of the enriched PAX3 binding sequences

from the mouse CASTing library (Table 3-1a; Figure 3-3a), the nucleotides at

positions +6, +7, and +8 were also under selection (GCT or GGT), and these

positions were also included in the scan. Therefore, the motif of

ATT.GTCA[C|T]G[C|G]T and its complement was used for motif scan.

The mouse genome was chosen as the starting material for the motif scan.

The Perl script returns segments of 72 base pairs (bp), which is sufficient to

provide a unique identifier and to enable subsequent phylogenetic comparisons.

Figure 3-4 displays a flowchart that outlines the search for candidate PAX3

binding sites. The motif scan identified a total of 1913 unique sequences across

the mouse genome. These were first filtered using Megablast with default

parameters to limit analyses to sequences that were conserved in both human and

mouse genomes. Amongst the motif-scan hits, 45 sequences had regions in the

human genome with 85% or higher sequence similarity and E-values of no greater

than 1×10^{-9} (Table 3-3; see Appendix II for BLAST hit sequences). In addition to

sequences obtained from the mouse genome motif scan, we reasoned that all

sequences from mouse and human CASTing libraries could contain potential

binding sites because they were recovered *in vitro* by PAX3 binding. As above,

their respective conservation between human and mouse was determined using

BLAST. This identified 74 and 75 segments from the mouse and human CASTing

library, respectively (Table 3-3; see Appendix II for BLAST-hit sequences). All

recovered sequences were then secondarily filtered by examining their

phylogenetic conservation across multiple species. The conservation of sequences

obtained by BLAST was reviewed using the ECR browser and sequences that did

not match to a specific genomic region were automatically eliminated at this stage.

Replicate sequences that corresponded to the same genomic region were

considered as one sequence. In higher eukaryotic genomes, transcription factor

binding sites (TFBSs) can be located upstream, downstream, or within intronic

regions of their regulated genes (Bulyk 2003), and all were considered in

subsequent analyses. Exons were excluded from the candidate list since they are

normally well conserved across species. Because TFBSs are expected to be more

conserved than their surrounding sequence, we also excluded sequences that did

not meet this criterion.

Of the genes identified after these filtering steps, three had a high-level of

overlap with *Pax3* expression in the neural tube during embryogenesis (Kubic *et

al,* 2008; Mennerich *et al,* 1998; Jagla *et al,* 1995; Mulieri *et al,* 2002) or were

implicated in oncogenesis (Reid and Nacheva 2004; Jagla *et al,* 1995; Onder *et al,*

2008; Bailey *et al,* 2009). These are: *BOC* (Brother of Cdo (Cell Adhesion
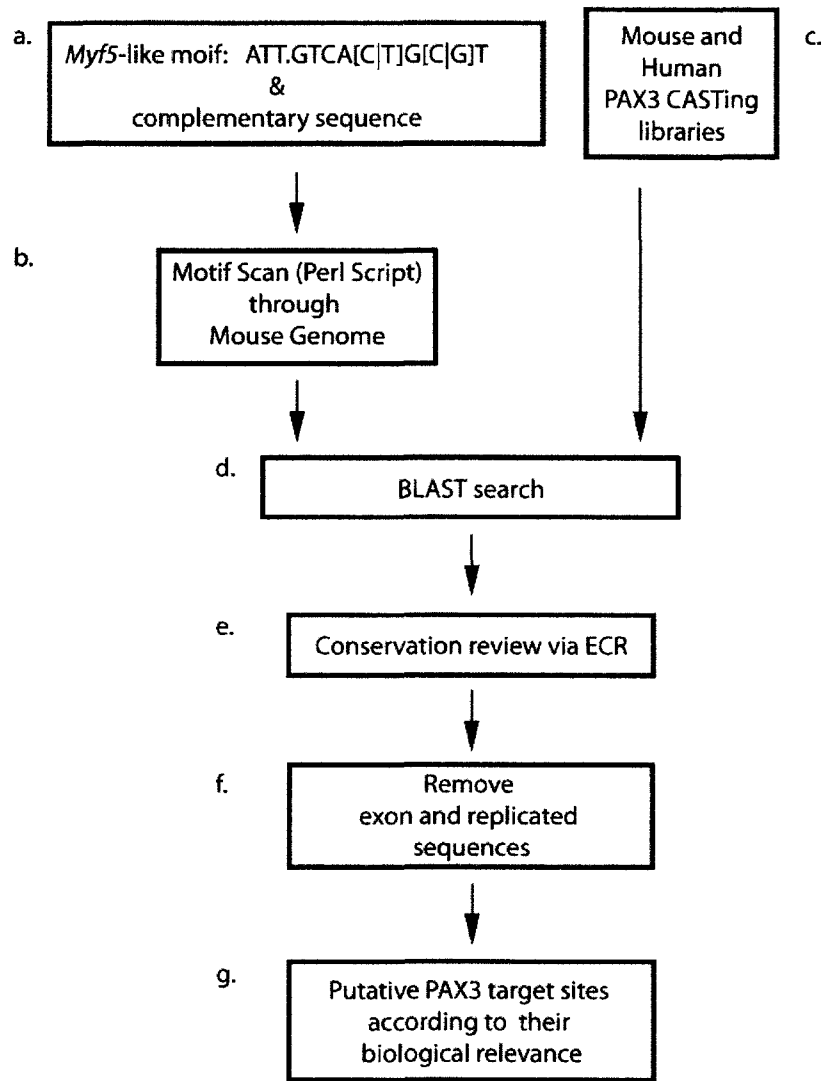
Figure 3-4. Flow chart of phylogenetic identification of potential PAX3 target genes. (a) The composite PAX3 binding motif was used for motif scan. The "." represents any nucleotide (A, T, G or C) at this position; '[C|T]' refers to either C or T; and '[C|G]' refers to either C or G. **(b)** The *Myf5*-like motifs were searched through the Mouse Genome using a motif scan program written in Perl script. **(c)** PAX3 CASTing libraries were also considered as candidate sequences to submit to BLAST. **(d)** The motif-scan results and CASTing library sequences were analyzed using BLAST: human sequences were blasted against the mouse genome and *vice versa*. **(e)** The ECR browser was used for assessing the phylogenetic conservation of each BLAST hit sequence. **(f)** Remove sequences mapping to exons and replicate sequences. **(g)** Potential PAX3 target sites were selected for further analyses based on their biological relevance.

59

Table 3-3. Number of sequences recovered from during the identification of putative PAX3 target sites

Sequence database / library	No. of candidate sequences	No. of BLAST hits	No. After removing exons and replicates	Putative target sites
Mouse genome motif scan for ATT.GTCA[C|T]G[C|G]T	1913	45	40	2
Mouse CASTing library	1260	74	54	0
Human CASTing library	1203	75	23	1

Table 3-4. Putative PAX3 binding sites chosen for further assessment

Putative PAX3 regulated genes	Target sequences (5' to 3')
PRDM12	taaacggtgtccattgcggcgcggcgggcaatttgtcacgctgttaaagtgatcgcattcattcggctccat
LBX1/BTRC	ctctgatctccaaattgttccttttattgcatttgtcacggtgaagattaaagtccctgggtggcctgcact
BOC	ccagaaatgaaaataagaaatagaagggtaattagtcactgtcacacacactgagggggtgaatgcggc

60

Molecule-Related/Down Regulated by Oncogenes), *LBX* (Ladybird homeobox 1

gene), and *PRDM12* (PR domain containing gene 12) (see Table 3-4 for their

sequences; see Figure 3-5 for outputs from the ECR browser).

3.2.4 Assessing phylogenetic conservation and PAX3 binding affinity of the *in silico* determined putative PAX3 binding sites

We first examined sequence conservation of the three putative PAX3

binding sites empirically in more detail. According to the previous ECR results,

they are well conserved across primates, rodents, and other mammals (Figure 3-5).

However, the *PRDM12* sequence in rhesus macaque (*Macaca mulatta*) and *LBX1*

sequence in chimpanzee (*Pan troglodytes*) were not annotated in the ECR browser.

In order to determine if these sequences are indeed not conserved in these species

or if the ECR algorithm failed to detect their conservation, the three putative

PAX3 target sites were blasted against the whole-genome shotgun (WGS)

database. This revealed that the proposed PAX3 target sites were in fact very well

conserved across multiple species (Figure 3-6).

To assess the interaction between PAX3 and the *in silico*-identified targets

described above, we prepared oligonucleotides for each putative PAX3-binding

element (Figure 3-7a) and performed binding assays with the recombinant PAX3

PD or a protein containing the tethered PD and HD (PDHD). Previously

characterized PAX3 targets from the *MYF5* and *MITF* promoters were included as

positive controls and an oligonucleotide spanning the *MYF5* site in which the PD

motif was mutated was used as negative control. As shown in Figure 3-7b (top

panel), the PAX3 PD bound with relatively equivalent affinity to the *MYF5*,

PRDM12, and *LBX1* oligonucleotides. As expected, we observed no binding to

the *MYF5 PDmut* sequence and also detected limited binding by the PD to the

MITF sequence, consistent with prior analyses (Corry and Underhill 2005).

Unexpectedly, we failed to detect PD binding to the *BOC* oligonucleotide, despite

the presence of a consensus 5'-GTCAC motif. On the other hand, the PDHD

protein bound to all oligonucleotides (Figure 3-7b, *bottom*). Interestingly, the

PDHD protein interacted with the *MYF5 PDmut* sequence, likely due to the

presence of a consensus 5'-ATTA HD motif. Together, these results demonstrate

that PAX3 possesses the ability to interact with each of our *in silico*-derived target

sequences, consistent with the possibility that these sequences represent *in vivo*

targets of PAX3.

3.2.5 Comparison of BOC and PAX3 expression in the mid-gestation mouse embryo

Among the three proposed PAX3 regulated genes, we focused on BOC to

establish further proof-of-principle. For this, we carried out co-

immunofluorescence analyses of PAX3 and BOC expression in embryonic day

10.5 (E10.5) mouse embryos. PAX3 expression was observed throughout the

dorsal neural tube (NT), dorsal root ganglia (DRG), dermomyotome (DM), and

migrating myoblasts (Figure 3-8a), which was consistent with previous analyses.

The BOC protein was also observed in the DM and dorsal NT, but not in the

PAX3-positive migratory myoblasts that populate the limb buds (Figure 3-8a).

This may suggest that BOC is down-regulated before PAX3-positive myoblasts

migrate away from the DM; the same may be true for PAX3-expresssing neural

crest cells that form the DRG and other neural-crest derived structures. Closer

views of cells in these regions are shown in Figure 3-8b, c, and d. In these studies,

it is clear that PAX3 and BOC are co-expressed at the cellular level; PAX3

displayed the expected nuclear distribution, whereas BOC was found primarily in

the plasma membrane, which is consistent with its known function as a

membrane-located receptor (Kang *et al*, 2002). In short, the extensive overlap of

BOC and PAX3 in the mouse embryo provides further evidence that *BOC* may be

an authentic PAX3 target gene.
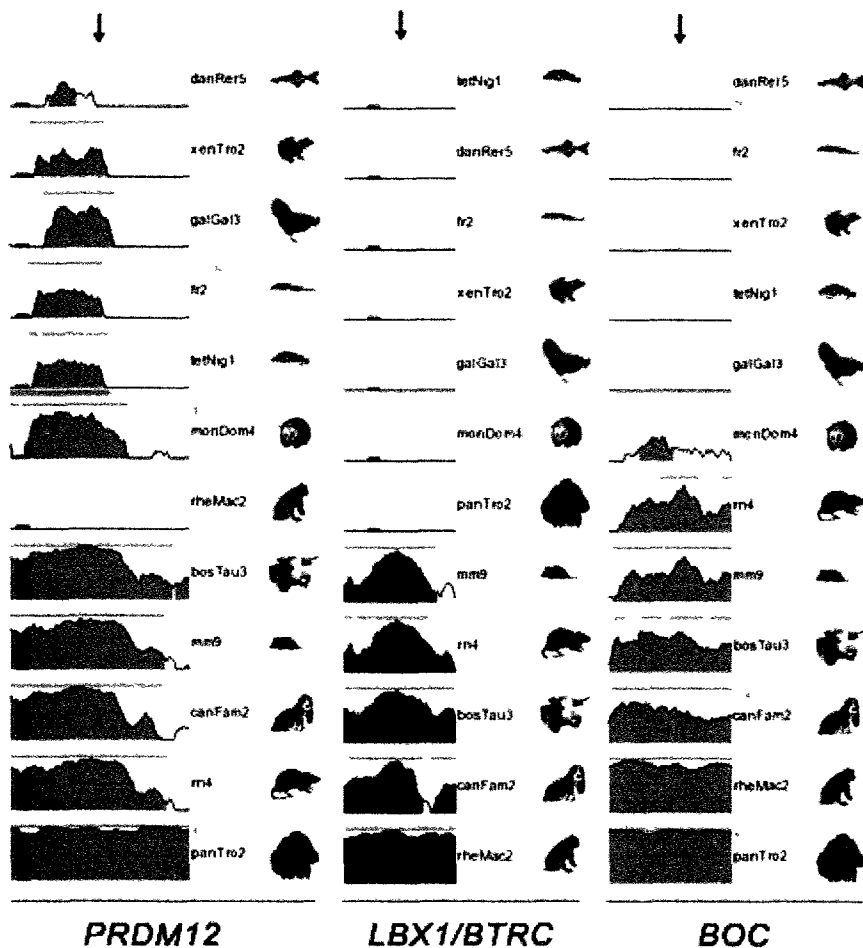
Figure 3-5. Evolutionary conservation of potential PAX3 binding sites.
Conservation of three putative PAX3 binding sites corresponding to genes *PRDM12, LBX2/BTRC* and *BOC* were reviewed using the ECR browser. Black arrows on the top and red vertical lines indicate the location of centers of each PAX3 binding site. The level of conservation is represented graphically according to percentage conservation on the right edge of each ECR output. Intronic regions are highlighted in pink; intergenic regions are highlighted in red; yellow and green regions are UTRs and repetitive elements, respectively. The species are annotated on the right side of each profile: bosTau3, *Bos taurus*; canFam2, *Canis familiaris*; danRer5, *Danio rerio*; fr2, *Takifugu rubripes*; galGal3, *Gallus gallus*; mm9, *Mus musculus*; monDom4, *Monodelphis domestica* ; rheMac2, *Macaca mulatta*; rn4, *Rattus norvegicus*; tetNig1, *Tetraodon nigroviridis*; xenTro2, *Xenopus tropicalis*; and panTro2, *Pan troglodytes*. The corresponding gene names are listed on the bottom.

64

Figure 3-6. Conservation of proposed PAX3 target sequences in higher eukaryotes. Alignment of whole genome shortgun blast hits of proposed PAX3-binding regions of *BOC*, *LBX1/BRTC* and *PRDM12* among higher eukaryotes. Sequences were aligned using ClustalX2 and shared nucleotides are highlighted in black with BoxShade 3.21. Consensus sequences were derived via WebLogo 2.8.2. The HD and PD recognition motifs are highlighted in blue and red boxes, respectively.

65

a.
```
MYF5          taccatgcaatt█████gcttttatgattta
MYF5 PDmut    taccatgcaatt█agc█gcttttatgattta
MITF          ctattcatcttt█████agtagtattaatag
BOC           agaagggtaatt█████tgtcacacacactg
PRDM12        cggcgggcaattt████gctgttaaagtgat
LBX1/BTRC     ttttattgcattt████ggtgaagattaaag
```

b.
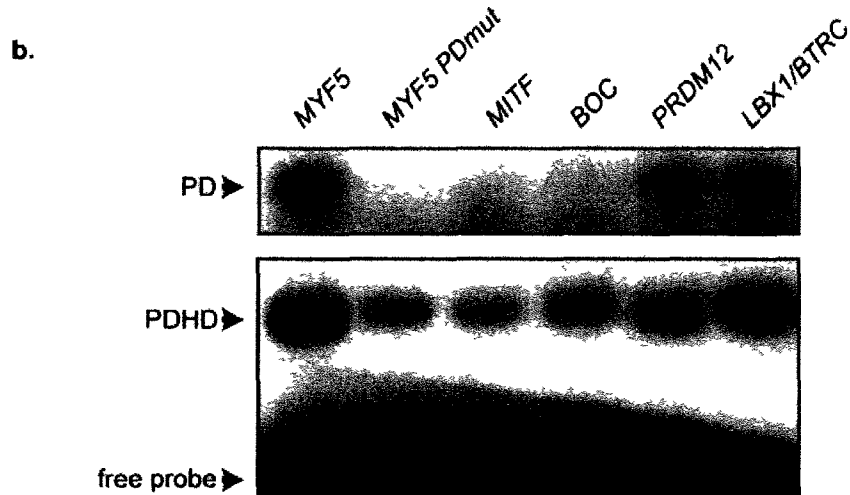
Figure 3-7. EMSA analysis of proposed PAX3 binding sites. (a) Alignment of proposed PAX3 binding sites. Black and gray highlighted regions are PD and HD motifs, respectively. (b) EMSAs of PAX3 interaction with previously characterized (*MITF, MYF5*) and *in silico*-derived putative PAX3 binding sites. Top panel, binding of the PAX3 PD; bottom panel, binding of the PAX3 PDHD protein, which contains the PAX3 PD and HD tethered by the PAX3 linker. Unbound probe is shown for the bottom panel only.

66
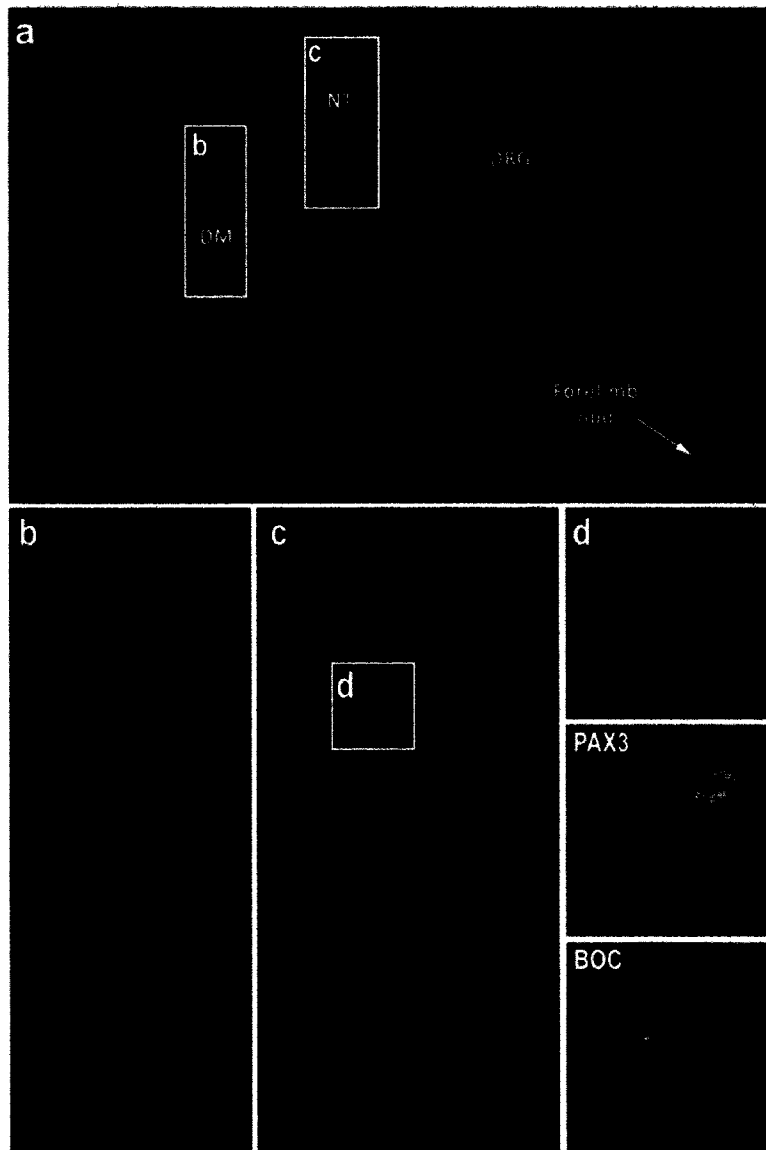
Figure 3-8. Expression of PAX3 and BOC during mid-gestation mouse embryogenesis. Immunofluorescence was used to examine the localization of endogenous BOC (*green*) and PAX3 (*red*) in 10.5 dpc mouse embryos. (**a**) Overview PAX3 and BOC distribution at E10.5. Locations of the dermomyotome (DM), dorsal root ganglia (DRG), neural tube (NT), and forelimb bud are indicated. (**b**) and (**c**) are enlarged views of the dermomyotome and neural tube. (**d**) Depicts the co-expression of PAX3 and BOC in individual cells derived from the neural tube (panel *c*).

67

3.4 DISCUSSION

The importance of evolutionary conservation of functional regulatory elements in genomes has been well established (Dermitzakis and Clark 2002; Prabhakar et al, 2006). With the increasing availability of genome sequences from diverse species, this has provided an important tool to identify functionally significant genomic regions. In the case of PAX3 target genes, this type of analysis has not been consistently used and many genes were reported when there was limited genome sequence available. After assessing conservation of published PAX3 target gene sequences, our results revealed that only *MITF* and *MYF5*, both of which have a composite-type binding site, were conserved across species and efficiently bound by PAX3. It is worth noting that two new putative binding sites were identified in the *MET* gene in RMS cell lines in a recent PAX3-FKHR ChIP experiment (Cao et al, 2010). However, the published *MET* target site was not detected, leading Cao et al to suggest that PAX3 might regulate *MET* through these two elements (Cao et al, 2010). This indicates that previous PAX3 target sites still require further authentication.

The significance of the composite-type of PAX3 binding site was further investigated using pre-existing PAX3 CASTing libraries (Barber et al, 2002). Sequences from the mouse CASTing library showed an abundance of composite-type motifs when analyzed using two independent motif-finding algorithms (Weeder and MEME). Notably, the last motif reported by MEME GGGGGTGGGGGTGGG (Table 3-1a) has not been linked to PAX3 binding so far. This kind of tandem repeat has recently been suggested to have potential

68

influence from analyses of evolutionary selection (Mularoni *et al*, 2010). Since this sequence was not over-represented in the human CASTing library, it may reflect a speciation effect that is not relevant to PAX3 binding. Analyses of the human PAX3 CASTing library gave discordant results using Weeder and MEME. Nevertheless, although neither method revealed enrichment for composite motifs or those comprising a PD recognition element, Weeder did identify multiple homeodomain binding sites (TAAT or ATTA) (Table 3-1). Moreover, while the enriched motifs given by MEME had extremely small p-values, scores assigned by Weeder were less significant than those observed in mouse. Nevertheless, manual analysis indicated that the composite-type of PAX3 binding site was enriched in the human CASTing library when compared to the over-represented motifs given by Weeder and MEME. The different results given by the two analysis packages could indicate that the human PAX3 CASTing experiment was prone to a higher level of non-specific binding.

Identification of transcription factor binding sites (TFBSs) has been a challenge due to the fact that regulatory sequences can be found upstream, downstream and within intronic regions of genes, and can also be distal to the regulated gene. Although known promoter regions of genes have been widely used for identification of TFBSs (McCue *et al*, 2001), this approach can only identify a small subset of TFBSs. For instance, the PAX3 binding site for *Myf5* is located about 57kb upstream of the transcription start site (Bajard *et al*, 2006). Thus, the entire mouse genome was considered in our study for the identification of putative PAX3 binding sites. The PAX3-binding site of *PRDM12* is located in

the intronic region between exons 1 and 2, and the PAX3 regulatory element of

BOC is also found in an intronic region (between exons 9 and 10); whereas the

proposed PAX3 target site for *LBX1* is located about 70-75kb upstream from its

transcription start site in the intergenic region with *BTRC*. Although this site

seems quite distal to the *LBX1* gene, it still has the potential to regulate gene

expression given the example of the PAX3-binding site for *Myf5*. Our findings are

also consistent with recent genome-wide analyses of PAX3-FKHR in RMS where

the majority of sites were identified in distal enhancer regions (Cao *et al*, 2010).

An advantage of our method is that it is not based on lineage-specific analyses and

therefore has the capacity to identify potential targets regardless of the cell type in

which they are expressed. It would be useful to combine the *in vivo* binding data

of Cao *et al* with our approach to provide a modified search string to carry out

additional *in silico* analyses.

Another method for the computational identification of TFBSs involves

motif matrix databases such as TRANSFAC (Kel *et al*, 2003). The publicly

available TRANSFAC matrix for the PAX3 binding site was generated using

artificial sequences derived *in vitro* and the matrix only specifies determinants for

the PD binding site using the consensus sequence TCGTCACTCTTHM.

Moreover, this TRANSFAC database has not been updated since 2005 and the

TRANSFAC matrix for PAX3 was last updated in January 2000. As a result,

many PAX3 target sites published after January 2000 were not included in this

matrix, especially the *Myf5*-like composite motif, which is the focus of this study.

Although our study used a relatively simple "word search" approach rather than

implementing a nucleotide-frequency matrix for the motif scan, the composite-type PAX3 binding site used was derived from multiple experimentally identified PAX3 binding sites that were highly conserved in diverse species. The potential limitation of our study is that we excluded many sequences obtained from the motif scan because they had not been implicated in any biological-relevant processes according to our knowledge. For future motif scan experiments, it would be the best to generate a frequency matrix for this kind of motif scan if possible.

After assessing thousand of sequences obtained by our motif scan, *PRDM12*, *LBX1* and *BOC* were chosen for further studies because of their potential functional overlap with PAX3 in oncogenesis and myogenesis. All three target sites are evolutionary conserved across multiple species and they were also efficiently bound by the PAX3 protein *in vitro*. The potential biological relevance of each of these three putative PAX3 regulated genes is discussed below.

PRDM12

PRDM12 is a member of the positive regulatory (PR)-domain containing (PRDM) family of proteins, many of which have been proposed to function in cancer and neurogenesis by interacting with the Notch-Hes pathway (for review, see Kinameri *et al*, 2008). In particular. *PRDM12* is located within the deleted region on human chromosome 9 that is observed in approximate 15% of chronic myeloid leukaemia (CML) patients (Kolomietz *et al*, 2003). Although little is known about its biological role, the results of recent immunohistochemical studies

71

in mouse and zebrafish embryos have suggested potential roles in brain development and neurogenesis (Kinameri *et al*, 2008; Sun *et al*, 2008). The expression of *Prdm12* is weakly observed in the caudal forebrain and midbrain at E9.5 in mouse embryos, which are signalling centres involved in brain patterning, and at higher levels in the same regions through E10.5 to E11.5 (Kinameri *et al*, 2008). In zebrafish, its expression is also found in brain, but is restricted to the olfactory placode, tegmentum, cerebellum and hindbrain (Sun *et al*, 2008). Besides the central nervous system (CNS), *Prdm12* is also expressed in both dorsal root and cranial ganglia in mouse embryos (Kinameri *et al*, 2008). Importantly, this spatial- and temporal-expression pattern of *Prdm12* overlaps with that of *Pax3* during embryogenesis (Goulding *et al*, 1991). Histological studies of *Pax3* deficient *Splotch* (*Sp*) mice have revealed abnormalities in the brain lumen (Moase and Trasler 1992). Thus, it is possible that *Prdm12* could be regulated by PAX3 during embryogenesis, particularly in the developing CNS and PNS.

LBX1/BTRC

LBX1 is a homeodomain protein encoded by the *Ladybird Homeobox 1* (*LBX1*) gene that maps at q24 on human chromosome 10 (Jagla *et al*, 1995). Because *LBX1* is within the same region as the proto-oncogene *HOX11*, as well as two translocation breakpoints observed in some T cell leukemia patients, it has been implicated in tumorigenesis (Jagla *et al*, 1995). *Lbx1* is expressed in the central nervous system (dorsal part of spinal cord and hindbrain) and migrating

72

muscle precursor cells in a pattern that is overlapping with *Pax3* (Jagla *et al,*

1995). Importantly, there is a complete absence of *Lbx1* expression in the somites

and limbs of *Pax3* deficient mouse embryos at E9.5 and E11.5 (Jagla *et al,* 1995;

Mennerich *et al,* 1998). In contrast, the expression of the proposed *Pax3* target

gene *Met* was not completely eliminated in these two tissues (Mennerich *et al,*

1998). These studies suggest the possibility that PAX3 regulates the expression of

Lbx1 more directly than *Met* in the development of skeletal muscle. It was also

demonstrated that *Lbx1* and *Pax3* could induce the expression of each other in

chicken fibroblasts, but this positive regulatory loop was not detected in murine

fibroblasts (Mennerich and Braun 2001) and may be species or tissue-specific.

Considering its overlapping expression pattern with *Pax3* and that it appears to

function in an auto-regulatory loop with PAX3, *Lbx1* warrants further analysis as

a potential PAX3 target gene and that may be involved in both normal and

pathogenic aspects of PAX3 biology.

BTRC (beta-transducin repeat containing) is a member of the F-box

protein family (Fujiwara *et al,* 1999). It is located on human chromosome 10q24-

q25, part of the long arm of chromosome 10 that has been associated with several

cancers (Fujiwara *et al,* 1999). It is a subunit of ubiquitin ligase complex

(Maniatis 1999). Notably, PAX3 has been shown to be degraded via the ubiquitin

pathway (Boutet *et al,* 2007) and could therefore function in a feedback loop with

BTRC. However, limited embryonic studies have been done for BTRC and the

extensive overlap in expression of PAX3 and LBX1 led us prioritize *Lbx1* as the

potential PAX3 target gene. It would nevertheless be worth evaluating *Btrc* and *Pax3* expression during development.

BOC

BOC (brother of CDO) is a receptor-like protein that belongs to the immunoglobulin/fibronectin type III repeat family of cell-surface molecules (Mulieri *et al*, 2002). Similar to the other two putative PAX3 target genes, *Boc* expression also overlaps with that of *Pax3* in the spinal cord and during development of the nervous system at early stages of embryogenesis. This is most notable in the dorsal neural tube, dermomyotome and limb bud in mid-gestation mouse embryos (Kang *et al*, 2002; Mulieri *et al*, 2002; Figure 3-8a). In addition, BOC has been shown to act as receptor for sonic hedgehog (Shh) via binding to one of its fibronectin repeats and this is necessary for enhancing Shh signalling, which is essential for neural tube patterning in embryogenesis (Tenzen *et al*, 2006). Moreover, knock down of *Boc* expression in zebrafish embryos revealed defects in dorsoventral axon tracts in the brain (Connor *et al*, 2005). BOC has also been implicated in myogenesis by positively regulating myogenic differentiation together with CDO in the C2C12 myoblast cell line (Kang *et al*, 2002). Given the high-level of overlap between *Boc* and *Pax3* expression during embryogenesis, and that both proteins have overlapping functions in neurogenesis and myogenesis, we believe *Boc* is a strong candidate for being a direct target of PAX3. Intriguingly, the BOC protein was not detected in the PAX3-positive migratory myoblasts that populate the limb buds, but was present in those cells with less

74

mobility around the more mobile myoblasts (Figure 3-8a). This latter population

likely reflects cells that are differentiating to form limb musculature. Because

BOC is involved in cell-cell contact (Krauss *et al*, 2005), this observation suggests

that *Boc* expressions needs to be down-regulated to facilitate delamination of

myoblast precursors from the dermomyotome so that they can enter the migratory

pathway. This may also occur as cells enter the neural crest pathway from the

neural tube.

Overall, the three putative PAX3 target genes identified in this study show

considerable overlap in their expression with that of *Pax3* during mouse

embryogenesis and they all have potential roles in cancer, either as oncogenic

proteins (LBX1) or tumor suppressor proteins (PRDM12 and BOC).

CHAPTER 4 – CHARACTERIZATION OF PAX3

IN MELANOMA

4.1 BACKGROUND

PAX3 is first expressed in neural crest and myogenic precursors during embryogenesis (Goulding *et al*, 1991). In these lineages, PAX3 plays important roles in the specification of a number of cell types, such as melanocytes and muscle cells (Tajbakhsh and Buckingham 2000; Hornyak *et al*, 2001). Upon differentiation of these cell types, *PAX3* expression is generally turned off, although *PAX3* expression persists in some stem cell populations where it maintains undifferentiated state (Lang *et al*, 2005). Haploinsufficiency of *PAX3* leads to a reduction in proliferation and survival of melanocyte precursors (Hornyak *et al*, 2001) and causes Waardenburg syndrome in humans and the *Splotch* phenotype in mice (Tassabehji *et al*, 1992; Epstein *et al*, 1991). On the other hand, deregulation of PAX3 activity has been associated with malignancy, notably in alveolar rhabdomyosarcoma and melanoma (Barr *et al*, 1995; Biegel *et al*, 1995; van den Broeke *et al*, 2006; Scholl *et al*, 2001).

PAX3 is an important regulator of melanocyte development and it is now thought that PAX3 also has important roles in melanoma pathogenesis. In this regard, previous studies have established that PAX3 is broadly expressed in melanoma cell lines, in primary cell lines derived from melanoma, and in tumor samples (Barr *et al*, 1999; Scholl *et al*, 2001; He *et al*, 2005; Plummer *et al*, 2008; Medic and Ziman 2010). Initial studies showed that reduction of PAX3 levels in melanoma cells resulted in increased apoptosis (Scholl *et al*, 2001; Muratovska *et al*, 2003), suggesting that the PAX3 may promote melanoma cell survival, possibly through up-regulation of $BCL-X_L$ (Margue *et al*, 2000), or down-

77

regulation of caspase 3 or the p53 tumor suppressor (He *et al*, 2005). Subsequent

studies have revealed that decreasing PAX3 levels results in a dose-dependent

reduction of proliferation of melanoma cells (He *et al*, 2005). The high level of

PAX3 expression in advanced melanoma tissues was also suggested to promote

metastasis by up-regulating the receptor tyrosine kinase *MET* in melanoma

(Mascarenhas *et al*, 2010). Moreover, PAX3 was shown to regulate TGFß directly

during embryonic neural crest migration, which may play an important role in the

context of melanoma metastasis by affecting cell-cell adhesion, growth and

migration (Mayanil *et al*, 2006). Thus, PAX3 expression in melanoma can be

linked to control of cell survival, proliferation and metastasis.

We previously assessed several human and mouse melanoma cell lines and

observed widespread PAX3 expression by immunofluorescence (Figure 4-1). In

collaboration with Dr. Victor Tron, PAX3 expression was also observed in patient

tissues when assessed using a melanoma tissue microarray (TMA) (unpublished

data), which is consistent with previous studies (Scholl *et al*, 2001; Plummer *et al*,

2008; Medic and Ziman 2010). Taken together with published studies, these

results provided further rationale for characterizing the role of PAX3 in melanoma.
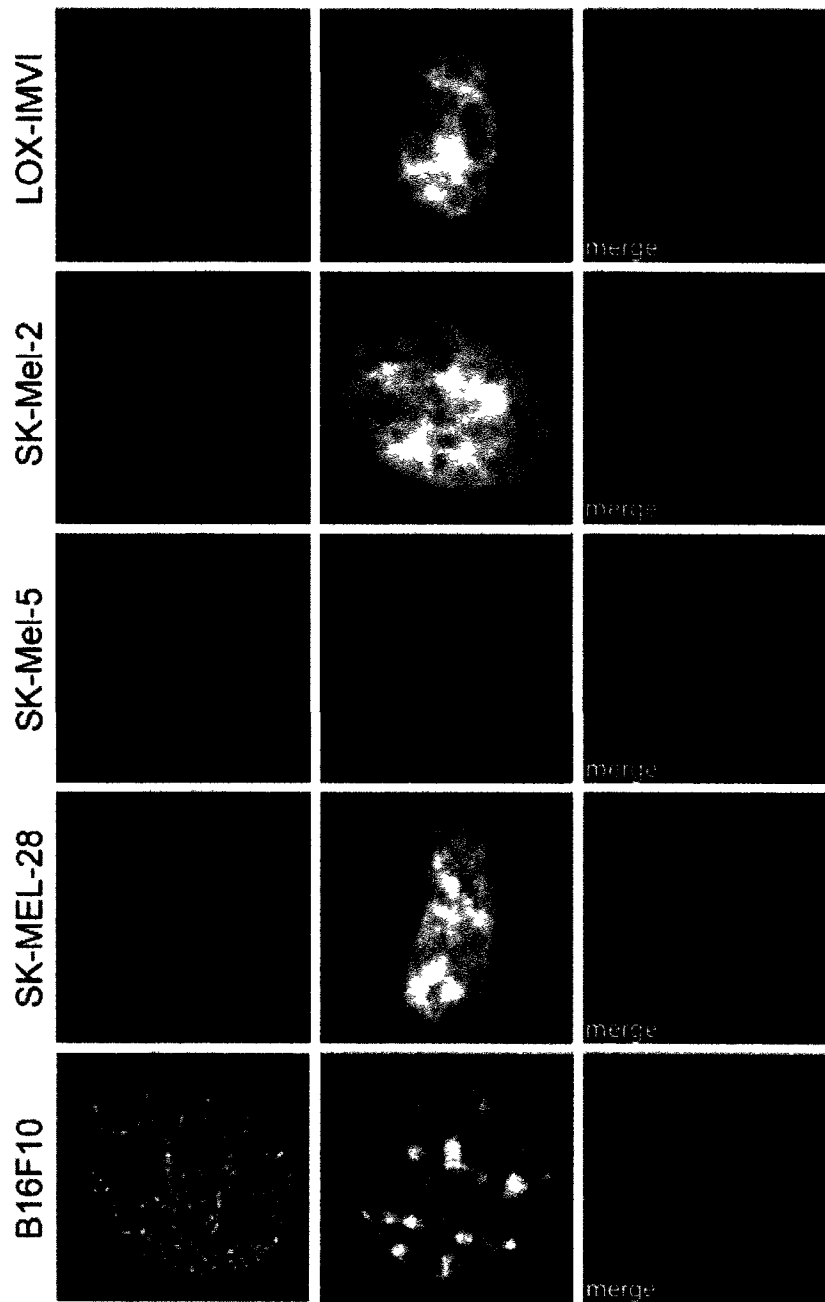
Figure 4-1. PAX3 expression in melanoma cell lines. Four human melanoma cell lines (LOX-IMVI, SK-Mel-2, SK-Mel-5 and SK-Mel-28) and one mouse melanoma cell line (B16F10) were stained with a PAX3 antibody and Hoechst and visualized by immunofluorescence (Corry and Underhill, unpublished).

4.2 RESULTS

4.2.1 Characterization of PAX3 expression during the cell cycle

To further characterize the potential role of PAX3 in melanoma, we made

use of the B16F10 mouse melanoma cell line, which has been extensively

characterized (Fidler 1975) and expresses high levels of endogenous PAX3

(Figure 4-1). By immunofluorescence staining of B16F10 cells with a PAX3

antibody, we found that its levels varied in unsynchronized cells (Figure 4-2). It

also appeared that PAX3 fluorescent intensity was related to altered morphology

of Hoechst stain in the nucleus (Figure 4-2). Specifically, PAX3 expression was

higher when foci of pericentromeric heterochromatin were prominent, suggesting

its levels may be elevated in the G2 phase of the cell cycle. We therefore

hypothesized that PAX3 expression is altered during the cell cycle. The use of cell

synchronization together with western blotting is a common approach to

determine if protein expression is altered during the cell cycle (Ito *et al*, 1999;

Bilican and Goding 2006; Lim *et al*, 2006). Fluorescence activated cell sorting

(FACS) is also commonly used for this purpose, but allows detecting protein

expression at multiple cell cycle stages in an unsynchronized cell population

(Sturzu *et al*, 2009). Hence, B16F10 cells were co-stained with the DNA dye

Propidium Iodide (PI) and PAX3 antibody, and their corresponding intensities

during the cell cycle were measured using FACS. The result revealed that PAX3

levels were lowest at the G0/G1 stage and increased through S to G2 and mitosis

(Figure 4-3). The intensities were significantly different between G1 and G2/M

with a p-value < 0.01. This suggests that PAX3 protein levels are cell-cycle

Figure 4-2. Endogenous PAX3 levels vary in B16F10 cells.
Immunofluorescence staining of unsynchronized B16F10 cells with a PAX3
antibody (FITC conjugated mouse secondary antibody) and Hoechst.

PAX3 expression in B16F10 along cell-cycle phases
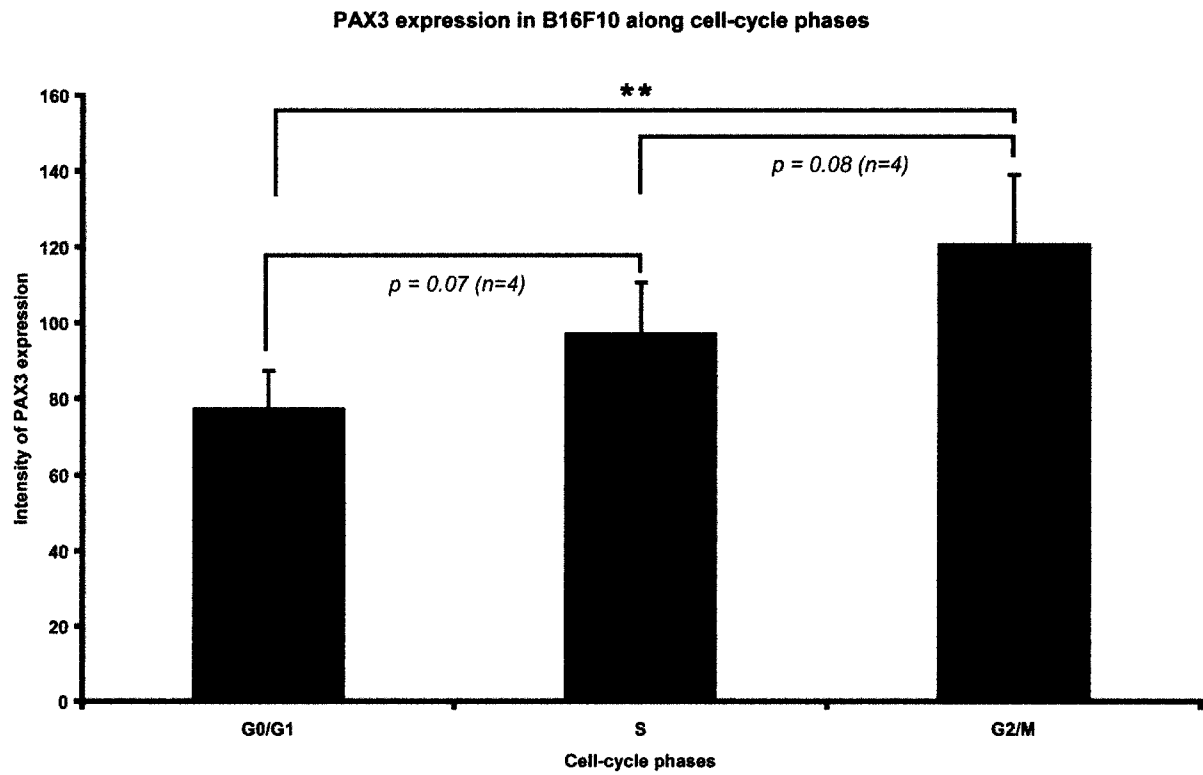
Figure 4-3. PAX3 expression in B16F10 cells during the cell cycle. PAX3 expression levels in terms of cell cycle stages were analyzed using fluorescence activated cell sorting (FACS). PAX3 protein was recognized using a FITC-conjugated PAX3 antibody and the DNA content determined by staining with propidium iodide (PI). Standard errors were calculated based on three independent experiments.

regulated and that it may have a functional role in cell cycle progression.

4.2.2 Characterization of the effect of attenuated PAX3 activity on cell cycle progression in B16F10 melanoma cells

To determine the potential role of PAX3 during cell cycle progression, we used the RNA interference (RNAi) technique to knock down PAX3 levels in the B16F10 melanoma cell line. One of the complicating factors is how to remove the multiple PAX3 isoforms expressed in mammalian cells. Previous studies have indicated that *PAX3c* and *PAX3d* are the two predominant *PAX3* transcripts expressed in melanoma cells (Barr *et al*, 1999; Parker *et al*, 2004). Because both *PAX3c* and *PAX3d* contain exon 8, a small interfering RNA (siRNA) targeting the end of exon 8 was employed to knockdown both transcripts specifically. In both immunofluorescence and western analysis, we observed that PAX3 levels were reduced efficiently over 72 hours compared to control cells (Figure 4-4).

Upon reduction of PAX3 levels in B16F10 cells, we first observed cell growth was delayed when compared to cells that were not transfected or transfected with a negative control siRNA, which is not targeted to a known genomic sequence. Cell numbers were quantified using a hemocytometer and the results were supportive of this observation (Figure 4-5). Statistical analysis (student t-test) revealed that cells grew significantly slower in PAX3-knockdown wells compared to controls, and this was consistently observed over a 72 hour period post-transfection. In addition, more floating cells were also observed in the PAX3-knockdown well and these reached statistical significance at 48 and 72hrs

83
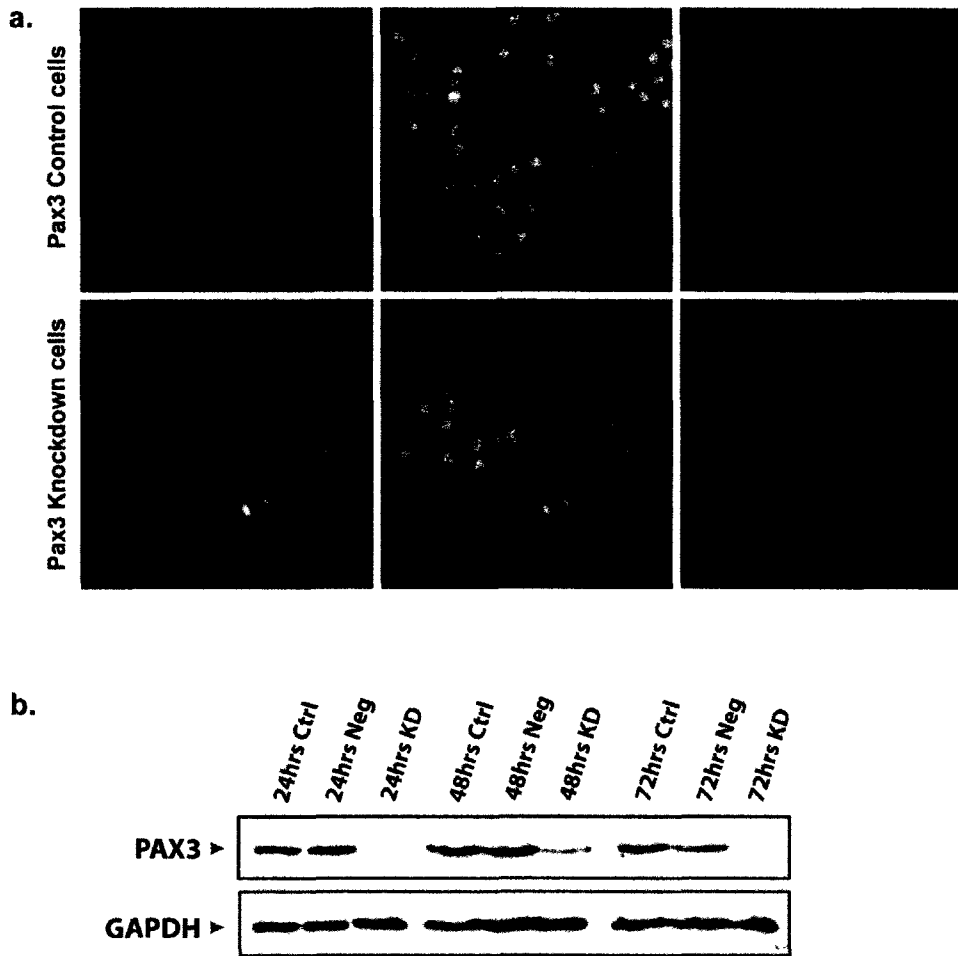
Figure 4-4. Examination of Pax3 knockdown efficiency in B16F10 cell line. (a) Immunofluoresence analysis before and after the knockdown of PAX3. The PAX3 antibody was detected with FITC-conjugated mouse secondary antibody; Hoechst was used to label DNA in cells. (b) Western blot of PAX3 protein levels in untransfected and transfected B16F10 cells over 72 hours. The control was transfected without siRNA. Negative control was transfected with Qiagen Allstars negative siRNA, which does not target a specific sequence in the mouse genome. PAX3 was knocked down using Qiagen mouse *Pax3* siRNA#4 and cells were harvested 24 hours, 48 hours or 72 hours post-transfection. GAPDH was used as a loading control.

Cell counts for PAX3 knockdown experiment (n=3)
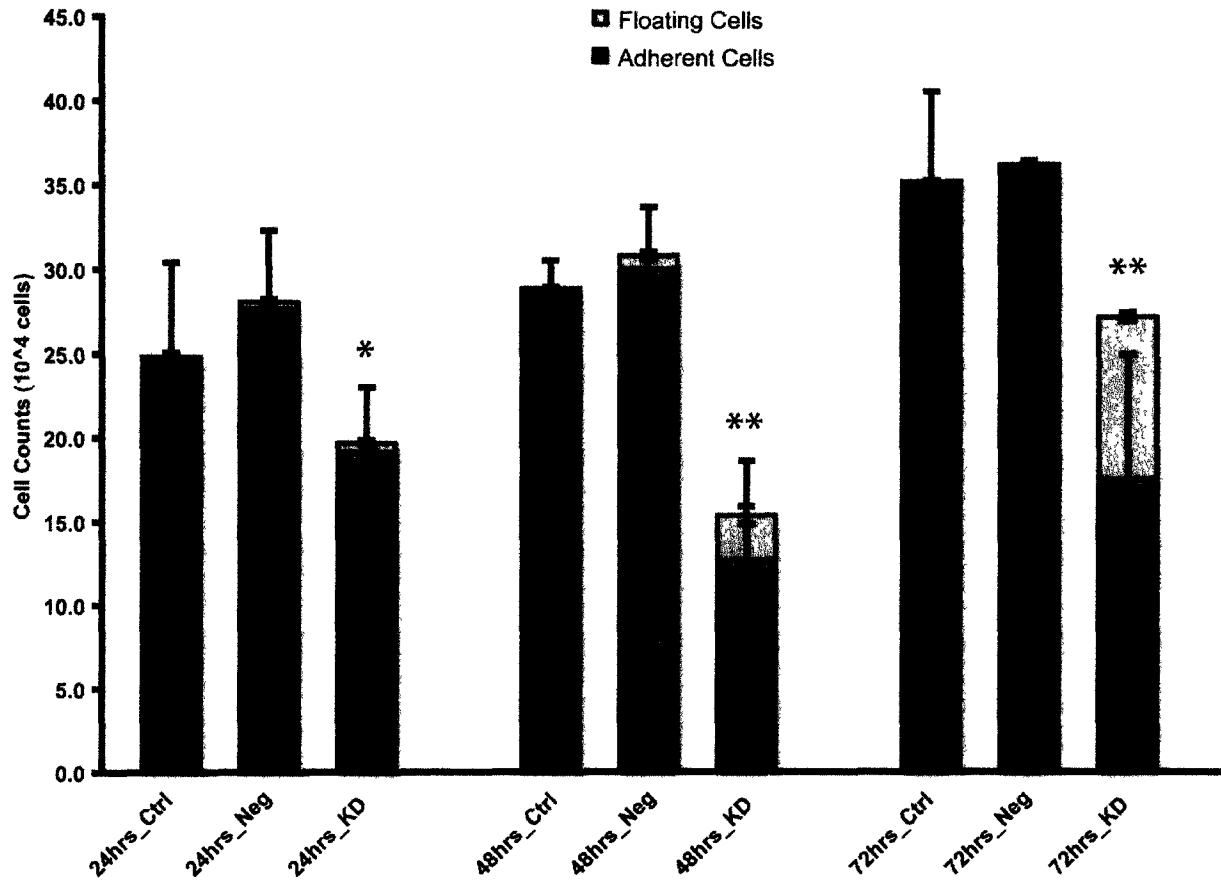
Figure 4-5. Effect of PAX3 knockdown on B16F10 cell growth. Aliquots of B16F10 cells were counted from untransfected and transfected wells using a Hemocytometer. Standard errors were calculated from 3 independent experiments. The * sign indicates a p-value < 0.05 and ** indicates a p-value < 0.001.

post-transfection. Because PAX3 has been reported to contribute to melanoma cell survival in a previous study (Scholl *et al*, 2001), the reduction in cell numbers in the PAX3-attenuated cells could potentially be due to cell death. However, according to the results shown in Figure 4-5, even with the inclusion of floating cells to the total cell count, there was still a large reduction in cell numbers in the PAX3-knockdown population at all time points. Thus, although the reduced cell numbers in the PAX3-knockdown may be partially attributed to cell death, they appear to be largely a consequence of a cell growth defect.

The cell cycle distribution upon PAX3 knockdown was examined using FACS analyses of cellular DNA content with PI. Upon comparing FACS results from PAX3-knockdown cells and controls, we found that attenuation of PAX3 levels in B16F10 cells led to a significant increase in the proportion of G1 cells (Figure 4-6a). Concomitant with the G0/G1 accumulation, we observed that cells in S and G2/M phases were significantly reduced at 24-hours post-transfection but recovered over time (Figure 4-6b and Figure 4-6c). The G0/G1 accumulation was not as pronounced at 48 hours post-transfection and was even smaller at 72hr. Similarly, although the percentage of S-phase cells was reduced at 48hr post-transfection, it was not at 72hr. The G2/M cells followed the same trend at 48 and 72-hours post-transfection, and continued to show a significant difference at the 48-hour time point. These results indicate that reducing PAX3 levels in mouse B16F10 melanoma cells is associated with an increase in the G0/G1 population, which could reflect a delay in entering S-phase.

a.

24hrs after transfection (n=4)

Legend: Ctrl, Neg, PAX3_KD

Y-axis: % of Gated cells

X-axis: G0/G1, S, G2/M

b.

48hrs after transfection (n=4)

Legend: Ctrl, Neg, PAX3_KD

Y-axis: % of Gated Cells

X-axis: G0/G1, S, G2/M
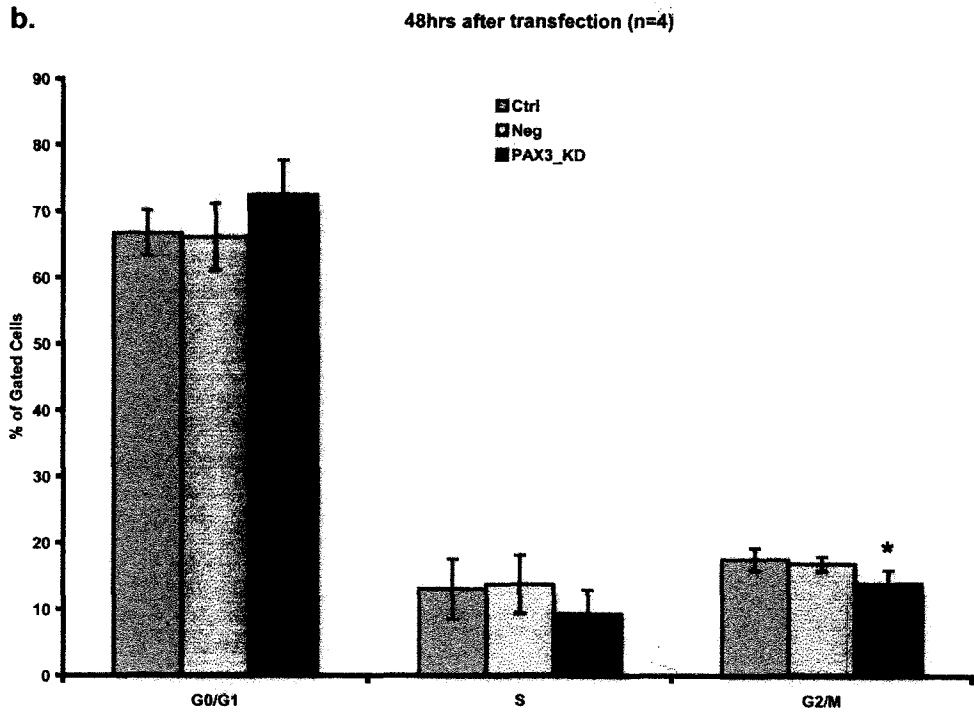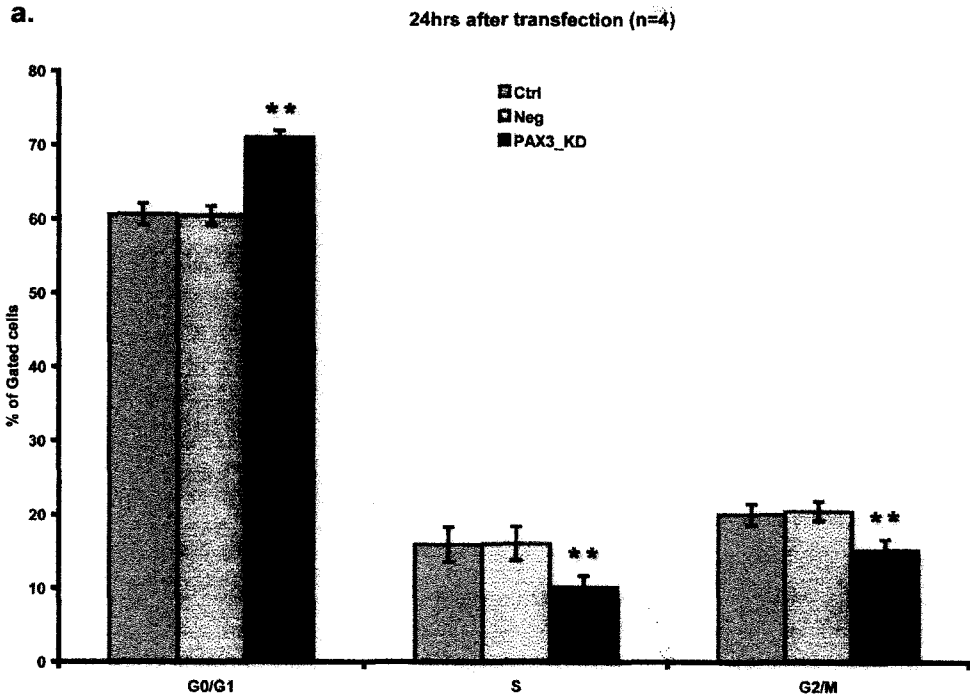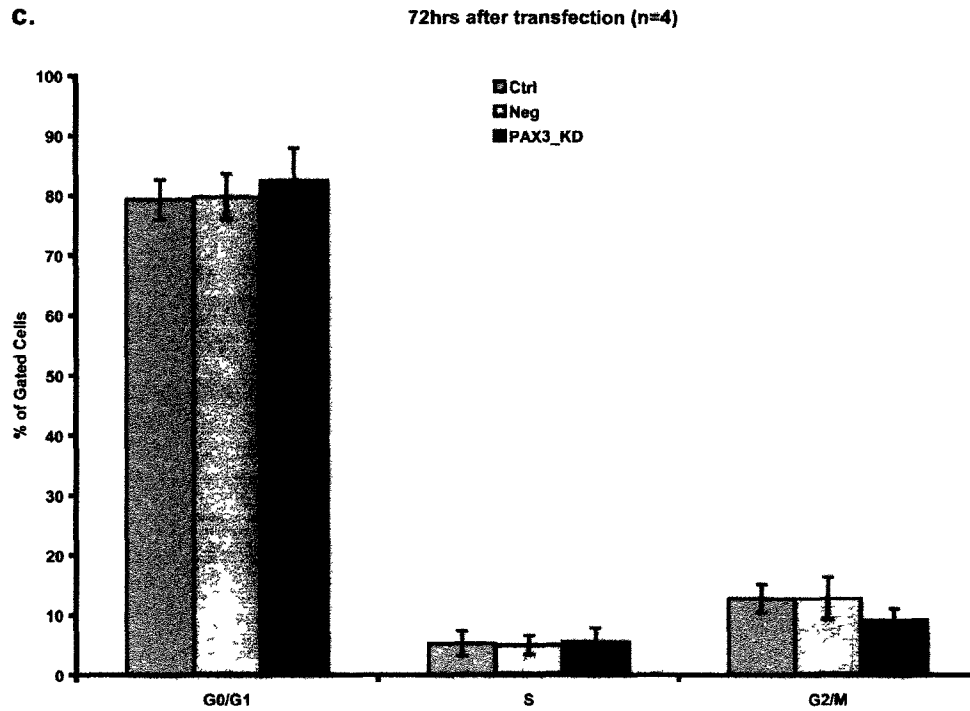
Figure 4-6. Effect of PAX3 knockdown on B16F10 cell cycle progression.
B16F10 cells treated with transfection reagent were used as cell control and marked as purple bars. Cells transfected with negative control siRNA or Pax3 siRNA #4 are marked in yellow and blue, respectively. Cells were collected, fixed and stained with PI after (a) 24 hours, (b) 48 hours and (c) 72 hours of transfection followed by FACS. Standard errors were calculated from 4 independent experiments and the "*" and "**" signs indicate a p-value of <=0.05 and <= 0.001, respectively.

89
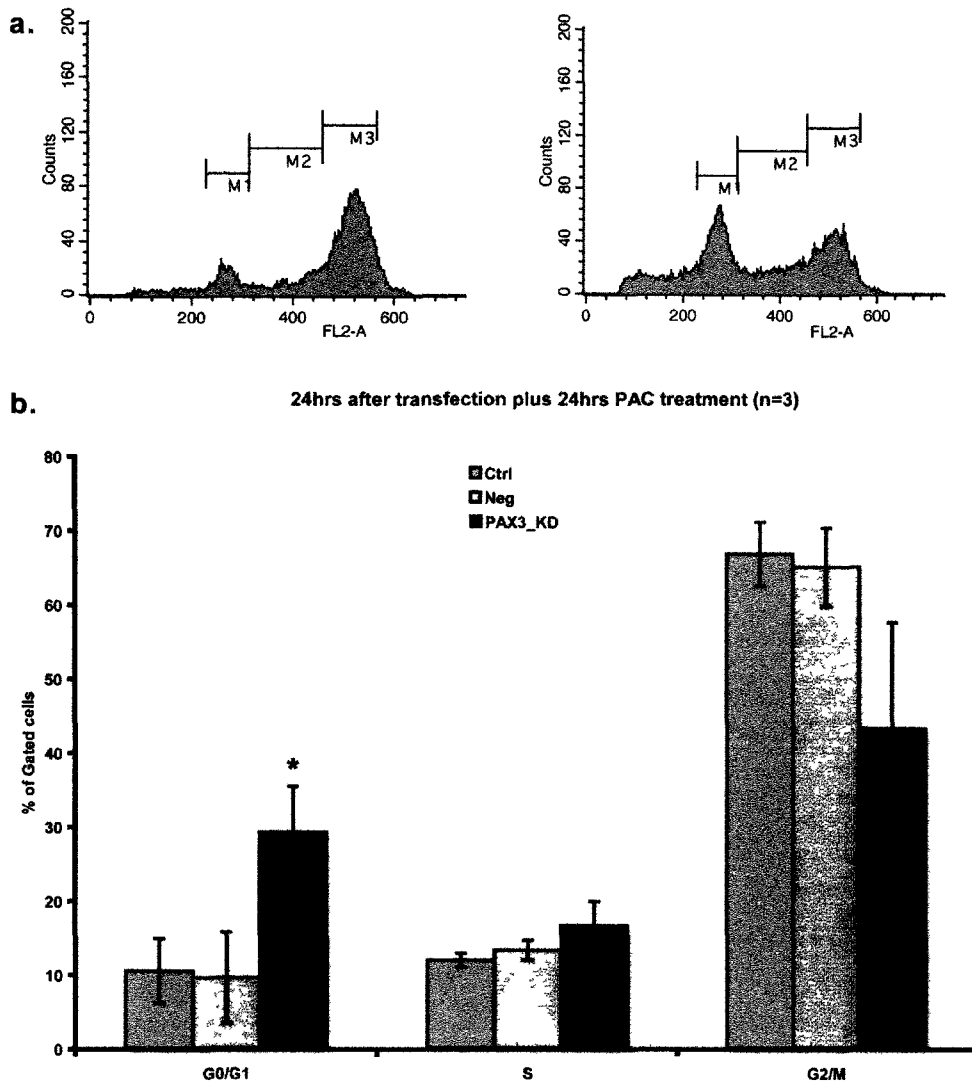
Figure 4-7. Combined effect of PAX3 knockdown and PAC treatment on B16F10 cell cycle progression. (a) FACS output of B16F10 cells with 24-hour PAC treatment after transfection without Pax3-siRNA (left) and with Pax3-siRNA (right). Gates M1, M2 and M3 indicate the areas of G1, S and G2/M, respectively. (b) Cell cycle distribution at 24-hour after addition of PAC. Cells were first treated with only transfection reagent (purple), negative control siRNA (yellow) or Pax3 siRNA (blue) and followed by a 24-hour PAC treatment. Standard errors were calculated from 3 sets of independent experiments and the * sign indicates the p-value < 0.05.

Although the G0/G1-accumulation following PAX3 knockdown was statistically significant, it only represented a 10% increase. In order to validate this effect, we treated cells with Paclitaxel (PAC), a drug that causes cells to arrest in mitosis (Samadi *et al*, 2009; Figure 4-7a). With the addition of PAC, the proportion of cells in G0/G1 doubled from 10% to approximately 20% (Figure 4-7b). This implies that PAX3 is important for controlling the rate at which cells move from G0/G1 to S-phase in cell cycle progression. Nevertheless, this effect again recovered over time (data not shown). Thus, the growth inhibitory effect of PAX3 knockdown may be bypassed or possibly diluted by rapidly dividing cells that were not transfected with the siRNA.

4.2.3 Examination of the specificity of PAX3 knockdown and its affect on cell cycle progression

In order to verify that the accumulation of cells in G0/G1 was specific to the reduction of PAX3 levels, we employed two siRNAs (siRNA#1 and siRNA#3) that target different locations on the *PAX3* transcript (Figure 4-8). The siRNA#3 targets the 3' end of *Pax3d*, while siRNA#1 targets a sequence in intron 8. As PAX3d was thought to be the predominant PAX3 isoform in melanoma cells, we expected siRNA#3 would knockdown PAX3 to a significant degree and likely induce G1 accumulation. On the other hand, siRNA#1 should not affect PAX3 levels because intron 8 is absent *Pax3d* transcripts.

Similar to the previous knockdown experiments with siRNA#4 (Figures 4-4 to 4-7), western blotting was done to examine the levels of PAX3 protein after
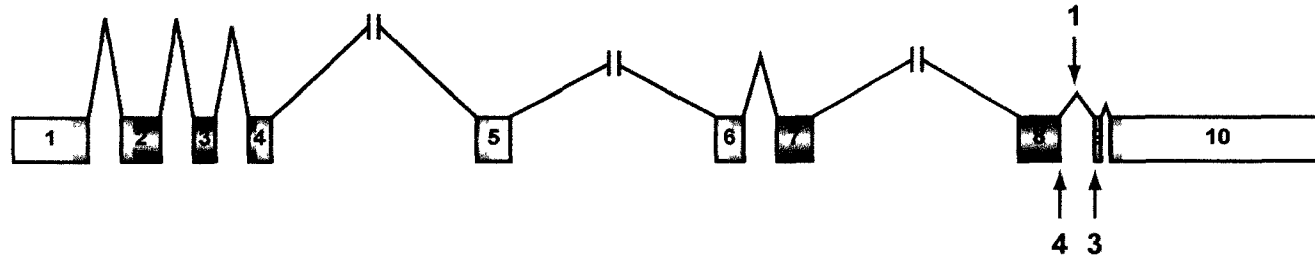
Figure 4-8. Validation of *PAX3* knockdown using additional siRNAs. The regions targeted by each siRNA are numbered 1, 3 and 4 (in red) according to Qiagen's nomenclature. Positions of siRNAs are marked by red arrows. Exons are presented as boxes and introns are shown as lines. Transcribed exons are highlighted in blue and UTRs are left open.
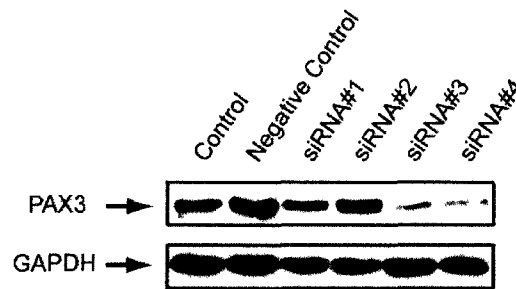
Figure 4-9. Western blot analyses of PAX3 knockdown efficiencies in B16F10 cells using different *PAX3* siRNAs. The control was treated with transfection reagent but no siRNA. The negative control was transfected with Qiagen Allstars negative siRNA. B16F10 cells were transfected with either Qiagen mouse *Pax3* siRNA #1, #3 or #4 to knockdown PAX3 expression. Cells were harvested 24 hours post-transfection. GAPDH was used as loading control.

PAX3 knockdown FACS experiment (n=3)
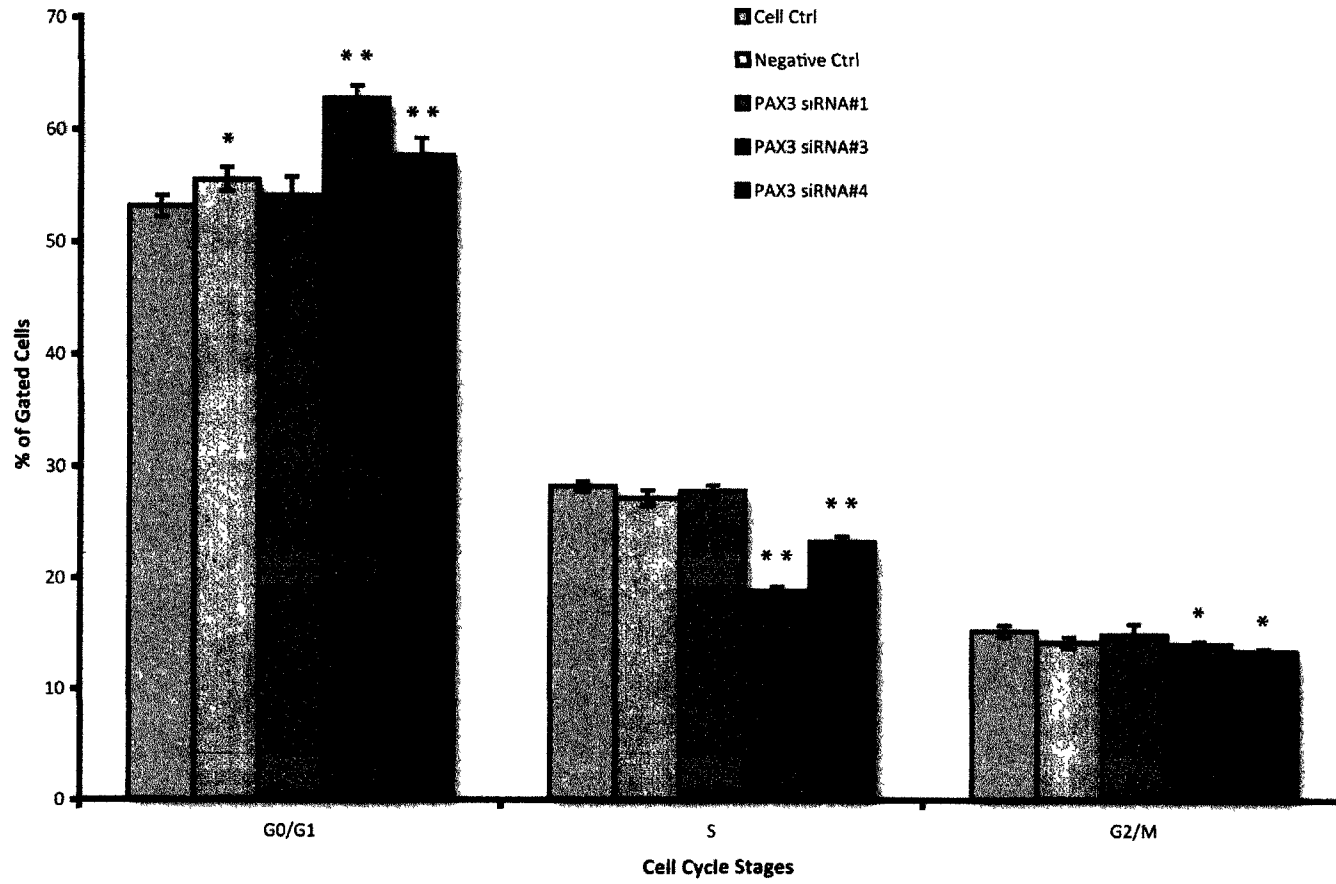
Figure 4-10. Analysis of B16F10 cell cycle distribution using different PAX3 siRNAs. B16F10 cells were treated with only transfection reagent (purple), negative control siRNA (orange), Pax3 siRNA#1 (olive), Pax3 siRNA#3 (light blue) or Pax3 siRNA#4 (dark blue). Cells were collected, fixed and stained with PI after 24 hours of transfection and subjected to FACS analysis. Standard errors were calculated from 3 independent experiments. The "*" and "**" signs indicate p-value of < 0.05 and $<= 0.001$, respectively.

treatment with each siRNA (Figure 4-9). As expected, PAX3 levels were markedly reduced upon introduction of siRNA#3 and siRNA#4, while siRNA#1 was largely ineffective. At the same time, FACS analysis was carried out and the results were consistent with those described above (Figure 4-10). Specifically, the percentage of G1 cells was elevated in the B16F10 population transfected with either *Pax3* siRNA#3 or siRNA#4, but not with siRNA#1. This result reinforced the idea that PAX3 attenuation induced G1 accumulation in B16F10 melanoma cells. In addition, the proportion of G1 cells was significantly increased in the negative control siRNA population when compared to the untreated control. Nevertheless, the associated p-value (0.04) was just slightly below the conventional threshold of 0.05 and was still significantly higher than observed for siRNA#3 and siRNA#4. As well, no specific effect was seen with this siRNA in either the S or G2/M fractions. This G1 affect could be due to the small variation between the three experimental replicates for the negative control samples and may therefore not suggest any biological significance.

4.2.4 Verification of the specificity of PAX3 function in cell cycle progression

Thus far, we have found that reducing PAX3 levels in B16F10 cells leads to a significant increase in the number of cells in the G0/G1-phase of the cell cycle. In order to further verify that this was a direct result of PAX3 activity, we made an RNAi-resistant PAX3 expression plasmid to rescue the PAX3 knockdown effect (Figure 4-11). To this end, we introduced multiple base

95

substitutions in the sequence complementary to siRNA#4 in a Pax3 expression

plasmid (noted as RNAiR-Pax3). The mutations did not affect the coding potential

of the Pax3 expression plasmid. Consequently, the Pax3-RNAiR plasmid was

transfected in to B16F10 cells together with Pax3 siRNA#4. Two different

quantities of RNAiR-Pax3 plasmid were added (100ng and 1000ng) and PAX3

levels were determined via western blotting (Figure 4-12). Importantly, PAX3

levels were restored in a dose-dependent manner by the introduction of the

RNAiR-Pax3 plasmid. FACS analysis also revealed that G1 accumulation was

reduced (Figure 4-13) in proportion to the amount of RNAiR-Pax3 plasmid and is

consistent with a direct role for PAX3 in regulation of cell cycle progression.
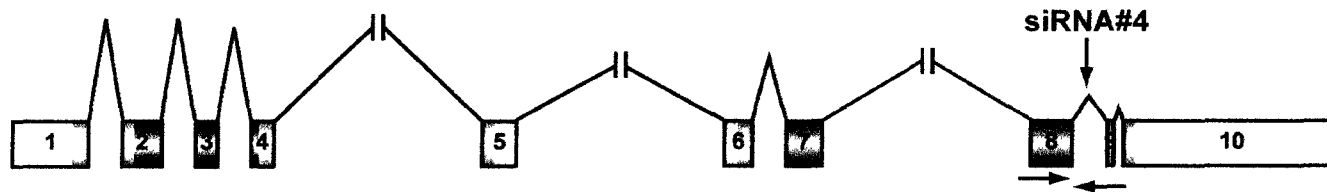
Figure 4-11. Strategy for creation of an RNAi-resistant PAX3 expression plasmid. Two primers were designed for mutagenesis and indicated as red arrows towards the 3' end of exon 8. The centre of the region is where siRNA#4 targets the mouse *Pax3* transcript. Exons are presented as boxes and introns are shown as lines. Transcribed exons are highlighted in blue and UTRs are left open.
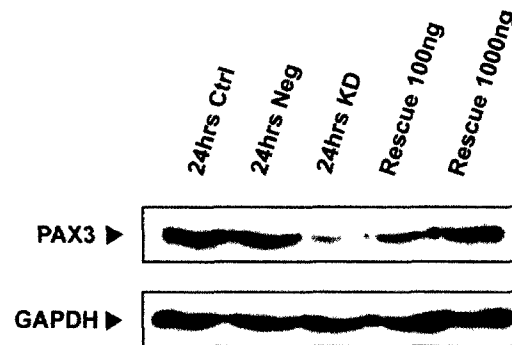
Figure 4-12. Western blot analysis of RNAiR PAX3. The control was transfected without siRNA. Negative control was transfected with Qiagen Allstars negative siRNA targeting no sequence in mouse genome. B16F10 cells were transfected with Qiagen mouse Pax3 siRNA #4 or Pax3 siRNA #4 with RNAi-resistant plasmid to knockdown or rescue PAX3 expression, respectively. Cells were harvested 24 hours post-transfection. GAPDH was used as loading control.
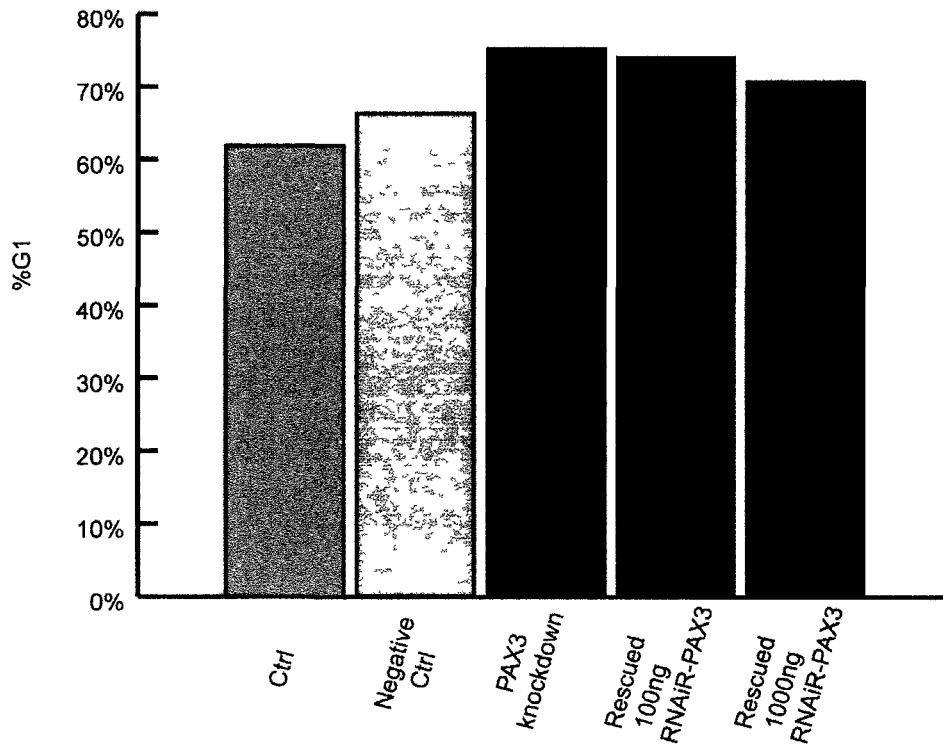
Figure 4-13. FACS analysis of G1 phase upon PAX3 rescue. B16F10 cells were treated with only transfection reagent (light blue), negative control siRNA (orange), Pax3 siRNA#4 (bright blue), Pax3 siRNA#4 with 100ng RNAiR-resistant Pax3 plasmid (dark blue) or Pax3 siRNA#4 with 1000ng RNAiR-resistant Pax3 plasmid (purple). Cells were collected, fixed and stained with PI, followed by FACS analysis.

98

4.3 DISCUSSION

In recent years, PAX3 has been suggested to be involved in the pathology of malignant melanoma (Scholl *et al*, 2001; He *et al*, 2005; Plummer *et al*, 2008; Medic and Ziman 2010). However, when our studies were initiated the role of PAX3 in melanoma at the cellular level had not been established. Previous studies in the lab reinforced the idea that PAX3 is widely expressed in melanoma cell lines and tumor samples. We therefore carried out a series of experiments directed at understanding how PAX3 deregulation affects melanoma cells using the mouse B16F10 cell line as a model system.

We first observed that PAX3 levels were altered during cell cycle progression in B16F10 cells. A similar phenotype was also observed in the human A2058 metastatic melanoma cell line, but the levels of another melanocyte transcription factor, MITF, were found to be uniform when assessed by co-immunofluorescence (Medic and Ziman 2010). In general, eukaryotic genes are expressed at a relatively constant level throughout the cell cycle, although exceptions include the histone genes and cyclins (Ewen 2000; Malumbres and Barbacid 2009). Nevertheless, another transcription factor implicated in melanoma, T-box 2, also displayed variation during the cell-cycle in melanoma cells, but its profile differed from that of PAX3 (Bilican and Goding 2006). Thus, we reasoned that altered PAX3 levels might have an important and distinct impact on melanoma cell cycle progression.

Upon attenuation of PAX3 levels in mouse B16F10 cells, we observed a significant reduction in cell number that became more prominent over the course

99

of 3-days (Figure 4-5). FACS results then revealed that there was a significantly greater percentage of cells in G0/G1 upon knockdown of PAX3 in B16F10 melanoma cells, which was more pronounced with the addition of PAC. Although PAX3 has previously been implicated in melanoma cell survival (Scholl *et al*, 2001; Muratovska *et al*, 2003), our analyses suggest a more significant role in regulating cell cycle progression and this is consistent with a number of recently published studies (Yang *et al*, 2008; He *et al*, 2010a). Importantly, this is consistent with the role of PAX3 in melanocyte development and stem cells (Hornyak *et al*, 2001; Yang *et al*, 2008; Fenby *et al*, 2008). Although the PAX3 target gene *Mitf* is involved in cell cycle regulation, it does not likely account for the effect of PAX3. First, MITF was shown to induce G1 arrest in melanocytes and melanoma cells by up-regulating expression of the cyclin-dependent kinase inhibitor $p21^{Cip1}$ (Carreira *et al*, 2005). The potential failure to activate *Mitf* upon PAX3 knockdown would therefore be expected to cause an increase in cell proliferation. Moreover, recent analyses in multiple melanoma cell lines find that MITF levels are not affected by PAX3 knockdown (He *et al*, 2010a). These findings suggest that the relationship between PAX3 and MITF may be deregulated in melanoma, and that there are likely other targets of PAX3 that are important for controlling cell cycle progression.

The G0/G1 accumulation in PAX3 knockdown cells was lost over time (about 72 hours). I believe this may reflect the incomplete knockdown of PAX3 expression. There were still a small number of B16F10 cells in the population with normal PAX3 activity that presumably proliferate at a normal rate, which is

100

approximately 14 hours per cell cycle in B16F10 (see Appendix III). As a result, it is possible that the progeny produced from these cells could overcome the significant growth retardation caused by reduced PAX3 levels. For instance, the small number of untransfected cells would have divided approximately 3.5 times after 48 hours of culture and approximately 5 times after 72 hours. This is consistent with what we observed in the FACS analyses (see Figure 4-6b and Figure 4-6c) where the elevation in G0/G1 cells was progressively diminished at 48 and even 72 hours post-transfection. This might also explain why the PAC treated cells resembled the control populations when treated at the 48-hour time point following Pax3 knockdown. Alternatively, B16F10 cells may adapt to the reduction in PAX3 levels throughout a yet to be defined mechanism.

The previous work done by Wang and colleagues established that PAX3c and PAX3d promoted cell proliferation, transformation, and migration of melanocytes, suggesting they function as oncogenes (Wang et al, 2006). PAX3c and PAX3d are the dominant isoforms expressed in melanoma and is therefore consistent with my observation that knocking down these two PAX3 isoforms would reduce cell proliferation. Taken together, we believe that the widespread expression of PAX3 in melanoma tissues contributes to proliferation and could therefore promote tumorigenesis, Recently, however, it has been suggested that PAX3 expression is not simply deregulated in melanoma, but is also widely expressed in normal skin and melanocytes (Medic and Ziman 2010; He et al, 2010b). Nevertheless, Medic and Ziman note that their samples were probably exposed to more UV than those used in other studies (Scholl et al, 2001; Plummer

101

et al, 2008) and that PAX3 expression coincided with an 'undifferentiated plastic

state' (Medic and Ziman 2010). Ultra-violet irradiation has been demonstrated to

upregulate *PAX3* in individual melanocytes in normal human skin via inhibiting

TGFß, which is secreted by keratinocytes to block melanocyte proliferation (Yang

et al, 2008). Plummer *et al* also found that PAX3 expression was more promiment

in younger patients with low levels of sun damage (Plummer *et al*, 2008).

Although the solar elastosis grade of each sample in Medic and He's studies are

not available, their samples (derived from Australian patients) potentially had

higher level of sun exposure compared to patients from North America. It is

therefore possible that PAX3 expression in these skin tissues was induced by UV

irradiation. In addition, Medic and Ziman used an antigen retrieval step in their

analyses that may provide a higher sensitivity for detecting PAX3 in normal and

pathogenic samples. Moreover, it was important to note that there are multiple cell

subtypes existing in normal skin epidermis, and PAX3 may have different roles in

these distinct populations, possibly involving cell survival, cell migration, cell

proliferation, and maintenance of the stem cell phenotype (Medic and Ziman

2010).

CHAPTER 5 – CONCLUSIONS AND FUTURE DIRECTIONS

5.1 CONCLUSIONS

Despite the essential roles of PAX proteins during embryogenesis, after birth and in multiple diseases, we have only a limited picture of genes under their direct regulation. Only PAX5, and more recently the PAX3-FKHR fusion protein (Schebesta *et al*, 2007; Cao *et al*, 2010), have been subjected to methods that identify binding sites on a genome wide scale. As a result, reported PAX target genes have typically been identified by a candidate approach. We therefore investigated a number of published PAX3 target sites and found that the majority either lacked conservation or had low or undetectable binding to PAX3 *in vitro*. Only target sites containing a composite-type binding motif (as in *MITF* and *MYF5*) met the criteria of efficient binding to PAX3 and phylogenetic conservation. Analyses of pre-existing PAX3-binding sequences (CASTing libraries) also revealed enrichment for composite-type motifs. These results provided the rationale for us to undertake an *in silico* search for putative PAX3 targets and three potential genes were identified: *LBX1*, *BOC* and *PRDM12*.

Deregulation of PAX3 activity has been reported to occur in malignant melanoma. Previous studies have revealed PAX3 is widely expressed in various melanoma cell lines and patient tumor samples. In our study, we found that PAX3 levels were regulated during cell cycle progression in B16F10 mouse melanoma cells, increasing as cells progressed from G1 to S to G2/M. Attenuation of PAX3 level in these cells induced a significant reduction in cell number and accumulation in G0/G1. These findings suggest that the increasing level of PAX3 could facilitate proliferation of melanoma cells by altering cell cycle dynamics.

104

Overall, both studies have provided important information to improve our understanding of PAX3 target gene regulation and its potential roles in melanoma pathology. Future studies will focus on directly relating these two areas.

5.2 FUTURE DIRECTIONS

5.2.1 Examination of potential PAX3 target genes *in vivo*

To further characterize the putative target sites we identified *in silico*, the next step will be to examine these sites *in vivo*. Chromatin immunoprecipitation (ChIP) has been widely used for *de novo* identification of transcription factor binding sites (TFBSs) and validation of endogenous binding for *in vitro* identified sites. In recent years, high-throughput ChIP-based technologies such as ChIP-chip (ChIP coupled with microarray) and ChIP-Seq (ChIP coupled with next-generation sequencing) have been developed to facilitate the identification of transcription factor binding sites on a genome-wide scale. However, the success of these experiments is highly dependent on the accessibility of the chromatin region to the transcription complex and the antibody used for its immunoprecipitation (Alexander *et al*, 2010). In the case of PAX3, although considerable effort has been focused on characterization of optimal PAX3 DNA-binding motifs and PAX3 downstream target sites *in vitro*, only the target in the *Myf5* gene has been characterized *in vivo* by ChIP (Bajard *et al*, 2006). A recent PAX3-FKHR genome-wide ChIP study in rhabdomyosarcoma revealed that most PAX3-FKHR target sites were located distal to transcription start sites and only a small fraction were associated with promoter regions of known genes (Cao *et al*,

2010). This is consistent with the location of potential binding sites in the *BOC*,

LBX1 and *PRDM12* genes and with the lack of conservation of the majority of

promoter associated PAX3 binding sites. A priority for future work will be to

validate PAX3 occupancy of these putative target sites. In addition, as these target

sites are likely to be cell type specific (Barski *et al*, 2007), selection of a proper

cell type for these studies will also be a key factor.

5.2.2 Control of apoptosis or proliferation

In this study, we achieved a robust knockdown of PAX3 protein in

B16F10 cells using RNAi. In these cells, we observed a significant reduction in

cell numbers and a greater percentage of dead (floating) cells. This could reflect

decreased cell proliferation, induction of cell death, or both. In this regard, PAX3

has been shown to be involved in both melanoma survival and proliferation

(Muratovska *et al*, 2003; Carreira *et al*, 2005). However, there is still a need to

characterize how PAX3 participates in these processes at the molecular level.

PAX3 has been shown to play roles in cell survival during neural tube

development and myogenesis by inhibiting pro-apoptotic factors p53 and PTEN,

respectively (Pani *et al*, 2002; Li *et al*, 2007). In addition, PAX3 can activate

transcription of the anti-apoptotic protein BCL-XL in RMS cells (Margue *et al*,

2000). PAX3 has also been suggested to contribute to control of cell proliferation

in melanocytes and melanoma via its downstream target *MITF*, which activates

cyclin-dependent kinase *p21* ^{Cip1} and causes G1 arrest (Carreira *et al*, 2005). As

indicated in chapter 4, this relationship appears to be altered in melanoma. PAX3

106

may also be involved in control of cell proliferation through interactions with cell-cycle regulators, including *Rb*, *Myc* and *p21* (Hsieh *et al*, 2006). In order to elucidate the molecular mechanism of cell reduction induced by PAX3 knockdown, it will be necessary to determine if expression levels of these genes are altered when PAX3 activity is attenuated. This can be done by measuring relative levels of their transcripts by real-time quantitative PCR or assessing protein levels by western blotting. Ideally, it would be preferable to assess global changes in gene expression using microarray analyses, which could also be used in conjunction with ChIP methods to identify direct PAX3 target genes.

5.2.3 Defining how PAX3 regulates the cell cycle progression

We observed that PAX3 levels were regulated during the cell cycle in B16F10 mouse melanoma cells and that PAX3 was required for progression from G1. However, the molecular basis of this effect was not further investigated. Previous studies of PAX3 disease-causing mutations in Waardenburg syndrome have identified two classes of mutants. Class I mutants (N47H, G81A and V265F) display normal DNA-binding but defects in localization and mobility, whereas class II mutants (Sp^d, F45L, S84F, Y90H, R271G) have intermediate affects on mobility and resemble wild type PAX3 with respect to distribution, but have severe defects in DNA binding (Corry and Underhill 2005; Corry *et al*, 2008). By re-introducing these PAX3 mutants into PAX3-knockdown B16F10 mouse melanoma cells, we can determine how PAX3 regulates cell cycle progression by assessing their recruitment to active transcription sites and the requirement for

107

DNA binding. This should determine if PAX3 DNA-binding and target gene occupancy are required to regulate cell cycle progression. On the other hand, PAX3 is known to interact with the HIR histone cell cycle regulation defective homolog A (HIRA) protein and the RB1 tumor suppressor, both of which are important cell cycle regulators (Lorain *et al*, 1998; Zhang *et al*, 2005; Wiggan *et al*, 1998). If PAX3 contributes to cell cycle regulation through these interactions, then both classes of disease mutants will likely be able to substitute for the wild-type protein. Alternatively, a series of PAX3 deletion constructs can be transfected into PAX3-knockdown B16F10 cells to identify the region(s) that are necessary for cell cycle regulation.

Although the rate of cell proliferation is clearly related to the cell cycle time span, time in G1 has been suggested to influence cell fate and the development of the nervous system (Salomoni and Calegari 2010). It was proposed that increasing the length of G1 caused cells to switch from a proliferative or self-renewal state to one that induced differentiation. This was observed in neural precursors and is believed to be present in other systems as well (Salomoni and Calegari 2010). With this in mind, PAX3 is essential for the differentiation of neural crest precursors in early development (Sato *et al*, 2005) and the subsequent expansion of committed melanoblasts (Hornyak *et al*, 2001). In addition, PAX3 plays important roles in melanocyte and skeletal muscle stem cells. PAX3 is also expressed in benign nevi (Plummer *et al*, 2008; He *et al*, 2010b), which are comprised of senescent cells that have permanently exited the cell cycle. Because these cell populations vary in their proliferative state, it would

108

be worthwhile to explore if the G0/G1 accumulation we observed in our study was a consequence of extending the G1 phase of the cell cycle. To this end, DNA labelling during S phase using BrdU can be employed to determine the length of cell-cycle parameters. Alternatively, we have determined cell cycle length in untreated B16F10 cells using a double-thymidine block to synchronize cells (see Appendix III) and then monitoring cell cycle progression by FACS. We could therefore apply this approach to determine the nature of cell cycle changes that occur upon PAX3-knockdown.

5.2.4 Significance of further characterization of PAX3 in melanoma

In the model of Salomoni and Calegari, control of G1 length was seen as a critical determinant of cell state (Salomoni and Calegari 2010). This is also likely to be relevant to cancer progression. Recent literature has revealed the existence of cancer stem cells within established tumours that resemble normal stem cells and are resistant to cancer therapy (for review, see O'Brien et al, 2009). Given that PAX3 is suggested to maintain a quiescent state in melanocyte stem cells (Nishikawa and Osawa 2007), it could act to regulate the decision between quiescence and proliferation by controlling G1 length. In melanoma, PAX3 may function to accelerate G1 length and contribute to tumor progression. If this model is correct, reduction of PAX3 levels may contribute to the appearance of markers that signify a more differentiated state, which is readily testable.

BIBLIOGRAPHY

Alexander, R.P., G. Fang, J. Rozowsky, M. Snyder, and M.B. Gerstein. 2010. Annotating non-coding regions of the genome. *Nat Rev Genet* **11**: 559-571.

Anderson, J., T. Gordon, A. McManus, T. Mapp, S. Gould, A. Kelsey, H. McDowell, R. Pinkerton, J. Shipley, and K. Pritchard-Jones. 2001. Detection of the PAX3-FKHR fusion gene in paediatric rhabdomyosarcoma: a reproducible predictor of outcome? *Br J Cancer* **85**: 831-835.

Bailey, J.M., A.M. Mohr, and M.A. Hollingsworth. 2009. Sonic hedgehog paracrine signaling regulates metastasis and lymphangiogenesis in pancreatic cancer. *Oncogene* **28**: 3513-3525.

Bailey, T.L. and C. Elkan. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36.

Bajard, L., F. Relaix, M. Lagha, D. Rocancourt, P. Daubas, and M.E. Buckingham. 2006. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev* **20**: 2450-2464.

Balch, C.M., A.C. Buzaid, S.J. Soong, M.B. Atkins, N. Cascinelli, D.G. Coit, I.D. Fleming, J.E. Gershenwald, A. Houghton, Jr., J.M. Kirkwood et al. 2001. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. *J Clin Oncol* **19**: 3635-3648.

Balczarek, K.A., Z.C. Lai, and S. Kumar. 1997. Evolution of functional diversification of the paired box (Pax) DNA-binding domains. *Mol Biol Evol* **14**: 829-842.

Baldwin, C.T., C.F. Hoth, J.A. Amos, E.O. da-Silva, and A. Milunsky. 1992. An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* **355**: 637-638.

Balling, R., U. Deutsch, and P. Gruss. 1988. undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1. *Cell* **55**: 531-535.

Bang, A.G., N. Papalopulu, M.D. Goulding, and C. Kintner. 1999. Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm. *Dev Biol* **212**: 366-380.

Barber, T.D., M.C. Barber, T.E. Cloutier, and T.B. Friedman. 1999. PAX3 gene structure, alternative splicing and evolution. *Gene* **237:** 311-319.

Barber, T.D., M.C. Barber, O. Tomescu, F.G. Barr, S. Ruben, and T.B. Friedman. 2002. Identification of target genes regulated by PAX3 and PAX3-FKHR in embryogenesis and alveolar rhabdomyosarcoma. *Genomics* **79:** 278-284.

Barr, F.G., J. Chatten, C.M. D'Cruz, A.E. Wilson, L.E. Nauta, L.M. Nycum, J.A. Biegel, and R.B. Womer. 1995. Molecular assays for chromosomal translocations in the diagnosis of pediatric soft tissue sarcomas. *JAMA* **273:** 553-557.

Barr, F.G., J.C. Fitzgerald, J.P. Ginsberg, M.L. Vanella, R.J. Davis, and J.L. Bennicelli. 1999. Predominant expression of alternative PAX3 and PAX7 forms in myogenic and neural tumor cell lines. *Cancer Res* **59:** 5443-5448.

Barski, A., S. Cuddapah, K. Cui, T.Y. Roh, D.E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129:** 823-837.

Bernasconi, M., A. Remppis, W.J. Fredericks, F.J. Rauscher, 3rd, and B.W. Schafer. 1996. Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins. *Proc Natl Acad Sci U S A* **93:** 13164-13169.

Biegel, J.A., L.M. Nycum, V. Valentine, F.G. Barr, and D.N. Shapiro. 1995. Detection of the t(2;13)(q35;q14) and PAX3-FKHR fusion in alveolar rhabdomyosarcoma by fluorescence in situ hybridization. *Genes Chromosomes Cancer* **12:** 186-192.

Bilican, B. and C.R. Goding. 2006. Cell cycle regulation of the T-box transcription factor tbx2. *Exp Cell Res* **312:** 2358-2366.

Bladt, F., D. Riethmacher, S. Isenmann, A. Aguzzi, and C. Birchmeier. 1995. Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud. *Nature* **376:** 768-771.

Blake, J.A. and M.R. Ziman. 2005. Pax3 transcripts in melanoblast development. *Dev Growth Differ* **47:** 627-635.

Bober, E., T. Franz, H.H. Arnold, P. Gruss, and P. Tremblay. 1994. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120:** 603-612.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986.
Conservation of a large protein domain in the segmentation gene paired
and in functionally related genes of Drosophila. *Cell* **47:** 1033-1040.

Bopp, D., E. Jamet, S. Baumgartner, M. Burri, and M. Noll. 1989. Isolation of
two tissue-specific Drosophila paired box genes, Pox meso and Pox neuro.
EMBO J **8:** 3447-3457.

Bouchard, M., A. Souabni, M. Mandler, A. Neubuser, and M. Busslinger. 2002.
Nephric lineage specification by Pax2 and Pax8. *Genes Dev* **16:** 2958-
2970.

Boutet, S.C., M.H. Disatnik, L.S. Chan, K. Iori, and T.A. Rando. 2007.
Regulation of Pax3 by proteasomal degradation of monoubiquitinated
protein in skeletal muscle progenitors. *Cell* **130:** 349-362.

Buchberger, A., N. Nomokonova, and H.H. Arnold. 2003. Myf5 expression in
somites and limb buds of mouse embryos is controlled by two distinct
distal enhancer activities. *Development* **130:** 3297-3307.

Buckingham, M. 2006. Myogenic progenitor cells and skeletal myogenesis in
vertebrates. *Curr Opin Genet Dev* **16:** 525-532.

Buckingham, M. and D. Montarras. 2008. Skeletal muscle stem cells. *Curr Opin
Genet Dev* **18:** 330-336.

Buckingham, M. and F. Relaix. 2007. The role of Pax genes in the development
of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell
functions. *Annu Rev Cell Dev Biol* **23:** 645-673.

Bulyk, M.L. 2003. Computational prediction of transcription-factor binding site
locations. *Genome Biol* **5:** 201.

Burri, M., Y. Tromvoukis, D. Bopp, G. Frigerio, and M. Noll. 1989. Conservation
of the paired domain in metazoans and its structure in three isolated human
genes. *EMBO J* **8:** 1183-1190.

Cao, L., Y. Yu, S. Bilke, R.L. Walker, L.H. Mayeenuddin, D.O. Azorsa, F. Yang,
M. Pineda, L.J. Helman, and P.S. Meltzer. 2010. Genome-wide
identification of PAX3-FKHR binding sites in rhabdomyosarcoma reveals
candidate target genes important for development and cancer. *Cancer Res*
70: 6497-6508.

Carreira, S., J. Goodall, I. Aksan, S.A. La Rocca, M.D. Galibert, L. Denat, L.
Larue, and C.R. Goding. 2005. Mitf cooperates with Rb1 and activates

p21Cip1 expression to regulate cell cycle progression. *Nature* **433:** 764-769.

Chalepakis, G., M. Goulding, A. Read, T. Strachan, and P. Gruss. 1994.
Molecular basis of splotch and Waardenburg Pax-3 mutations. *Proc Natl Acad Sci U S A* **91:** 3685-3689.

Chalepakis, G. and P. Gruss. 1995. Identification of DNA recognition sequences for the Pax3 paired domain. *Gene* **162:** 267-270.

Chi, N. and J.A. Epstein. 2002. Getting your Pax straight: Pax proteins in development and disease. *Trends Genet* **18:** 41-47.

Christ, B. and C.P. Ordahl. 1995. Early stages of chick somite development. *Anat Embryol (Berl)* **191:** 381-396.

Connor, R.M., C.L. Allen, C.A. Devine, C. Claxton, and B. Key. 2005. BOC, brother of CDO, is a dorsoventral axon-guidance molecule in the embryonic vertebrate brain. *J Comp Neurol* **485:** 32-42.

Conway, S.J., D.J. Henderson, M.L. Kirby, R.H. Anderson, and A.J. Copp. 1997. Development of a lethal congenital heart defect in the splotch (Pax3) mutant mouse. *Cardiovasc Res* **36:** 163-173.

Corry, G.N., M.J. Hendzel, and D.A. Underhill. 2008. Subnuclear localization and mobility are key indicators of PAX3 dysfunction in Waardenburg syndrome. *Hum Mol Genet* **17:** 1825-1837.

Corry, G.N. and D.A. Underhill. 2005. Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. *Pigment Cell Res* **18:** 427-438.

Dahl, E., H. Koseki, and R. Balling. 1997. Pax genes and organogenesis. *Bioessays* **19:** 755-765.

Dermitzakis, E.T. and A.G. Clark. 2002. Evolution of transcription factor binding sites in Mammalian gene regulatory regions: conservation and turnover. *Mol Biol Evol* **19:** 1114-1121.

Deutsch, U., G.R. Dressler, and P. Gruss. 1988. Pax 1, a member of a paired box homologous murine gene family, is expressed in segmented structures during development. *Cell* **53:** 617-625.

Dietrich, S., F. Abou-Rebyeh, H. Brohmann, F. Bladt, E. Sonnenberg-Riethmacher, T. Yamaai, A. Lumsden, B. Brand-Saberi, and C.

Birchmeier. 1999. The role of SF/HGF and c-Met in the development of skeletal muscle. *Development* **126:** 1621-1629.

Durbec, P.L., L.B. Larsson-Blomberg, A. Schuchardt, F. Costantini, and V. Pachnis. 1996. Common origin and developmental dependence on c-ret of subsets of enteric and sympathetic neuroblasts. *Development* **122:** 349-358.

Epstein, D.J., M. Vekemans, and P. Gros. 1991. Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67:** 767-774.

Epstein, J.A., D.N. Shapiro, J. Cheng, P.Y. Lam, and R.L. Maas. 1996. Pax3 modulates expression of the c-Met receptor during limb muscle development. *Proc Natl Acad Sci U S A* **93:** 4213-4218.

Erickson, C.A. 1993. From the crest to the periphery: control of pigment cell migration and lineage segregation. *Pigment Cell Res* **6:** 336-347.

Ewen, M.E. 2000. Where the cell cycle and histones meet. *Genes Dev* **14:** 2265-2270.

Fenby, B.T., V. Fotaki, and J.O. Mason. 2008. Pax3 regulates Wnt1 expression via a conserved binding site in the 5' proximal promoter. *Biochim Biophys Acta* **1779:** 115-121.

Fidler, I.J. 1975. Biological behavior of malignant melanoma cells correlated to their survival in vivo. *Cancer Res* **35:** 218-224.

Fredericks, W.J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F.G. Barr, and F.J. Rauscher, 3rd. 1995. The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol Cell Biol* **15:** 1522-1535.

Fujiwara, T., M. Suzuki, A. Tanigami, T. Ikenoue, M. Omata, T. Chiba, and K. Tanaka. 1999. The BTRC gene, encoding a human F-box/WD40-repeat protein, maps to chromosome 10q24-q25. *Genomics* **58:** 104-105.

Galibert, M.D., U. Yavuzer, T.J. Dexter, and C.R. Goding. 1999. Pax3 and regulation of the melanocyte-specific tyrosinase-related protein-1 promoter. *J Biol Chem* **274:** 26894-26900.

Galili, N., R.J. Davis, W.J. Fredericks, S. Mukhopadhyay, F.J. Rauscher, 3rd, B.S. Emanuel, G. Rovera, and F.G. Barr. 1993. Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. *Nat Genet* **5:** 230-235.

Ginsberg, J.P., R.J. Davis, J.L. Bennicelli, L.E. Nauta, and F.G. Barr. 1998. Up-regulation of MET but not neural cell adhesion molecule expression by the PAX3-FKHR fusion protein in alveolar rhabdomyosarcoma. *Cancer Res* **58:** 3542-3546.

Goulding, M., A. Lumsden, and A.J. Paquette. 1994. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development* **120:** 957-971.

Goulding, M.D., G. Chalepakis, U. Deutsch, J.R. Erselius, and P. Gruss. 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *EMBO J* **10:** 1135-1147.

He, S., C.G. Li, L. Slobbe, A. Glover, E. Marshall, B.C. Baguley, and M.R. Eccles. 2010a. PAX3 knockdown in metastatic melanoma cell lines does not reduce MITF expression. *Melanoma Res*.

He, S., H.S. Yoon, B.J. Suh, and M.R. Eccles. 2010b. PAX3 Is extensively expressed in benign and malignant tissues of the melanocytic lineage in humans. *J Invest Dermatol* **130:** 1465-1468.

He, S.J., G. Stevens, A.W. Braithwaite, and M.R. Eccles. 2005. Transfection of melanoma cells with antisense PAX3 oligonucleotides additively complements cisplatin-induced cytotoxicity. *Mol Cancer Ther* **4:** 996-1003.

Hill, R.E., J. Favor, B.L. Hogan, C.C. Ton, G.F. Saunders, I.M. Hanson, J. Prosser, T. Jordan, N.D. Hastie, and V. van Heyningen. 1991. Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354:** 522-525.

Hodgkinson, C.A., K.J. Moore, A. Nakayama, E. Steingrimsson, N.G. Copeland, N.A. Jenkins, and H. Arnheiter. 1993. Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell* **74:** 395-404.

Hornyak, T.J., D.J. Hayes, L.Y. Chiu, and E.B. Ziff. 2001. Transcription factors in melanocyte development: distinct roles for Pax-3 and Mitf. *Mech Dev* **101:** 47-59.

Hsieh, M.J., Y.L. Yao, I.L. Lai, and W.M. Yang. 2006. Transcriptional repression activity of PAX3 is modulated by competition between corepressor KAP1 and heterochromatin protein 1. *Biochem Biophys Res Commun* **349:** 573-581.

Ito, M., H. Kodama, and A. Komamine. 1999. Gene Expression and Its
    Regulation during the Cell Cycle of Higher Plants in Synchronous
Cell Culture Systems. *In Vitro Cellular & Developmental Biology. Plant* **35**.

Jackson, I.J., D.M. Chambers, K. Tsukamoto, N.G. Copeland, D.J. Gilbert, N.A.
    Jenkins, and V. Hearing. 1992. A second tyrosinase-related protein, TRP-2,
    maps to and is mutated at the mouse slaty locus. *EMBO J* **11**: 527-535.

Jagla, K., P. Dolle, M.G. Mattei, T. Jagla, B. Schuhbaur, G. Dretzen, F. Bellard,
    and M. Bellard. 1995. Mouse Lbx1 and human LBX1 define a novel
    mammalian homeobox gene family related to the Drosophila lady bird
    genes. *Mech Dev* **53**: 345-356.

Jiao, Z., R. Mollaaghababa, W.J. Pavan, A. Antonellis, E.D. Green, and T.J.
    Hornyak. 2004. Direct interaction of Sox10 with the promoter of murine
    Dopachrome Tautomerase (Dct) and synergistic activation of Dct
    expression with Mitf. *Pigment Cell Res* **17**: 352-362.

Jordan, T., I. Hanson, D. Zaletayev, S. Hodgson, J. Prosser, A. Seawright, N.
    Hastie, and V. van Heyningen. 1992. The human PAX6 gene is mutated in
    two patients with aniridia. *Nat Genet* **1**: 328-332.

Jun, S. and C. Desplan. 1996. Cooperative interactions between paired domain
    and homeodomain. *Development* **122**: 2639-2650.

Kang, J.S., P.J. Mulieri, Y. Hu, L. Taliana, and R.S. Krauss. 2002. BOC, an Ig
    superfamily member, associates with CDO to positively regulate
    myogenic differentiation. *EMBO J* **21**: 114-124.

Kassar-Duchossoy, L., E. Giacone, B. Gayraud-Morel, A. Jory, D. Gomes, and S.
    Tajbakhsh. 2005. Pax3/Pax7 mark a novel population of primitive
    myogenic cells during development. *Genes Dev* **19**: 1426-1431.

Kel, A.E., E. Gossling, I. Reuter, E. Cheremushkin, O.V. Kel-Margoulis, and E.
    Wingender. 2003. MATCH: A tool for searching transcription factor
    binding sites in DNA sequences. *Nucleic Acids Res* **31**: 3576-3579.

Keller, C., B.R. Arenkiel, C.M. Coffin, N. El-Bardeesy, R.A. DePinho, and M.R.
    Capecchi. 2004b. Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr
    mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev*
    18: 2614-2626.

Keller, C., M.S. Hansen, C.M. Coffin, and M.R. Capecchi. 2004a. Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. *Genes Dev* **18:** 2608-2613.

Kinameri, E., T. Inoue, J. Aruga, I. Imayoshi, R. Kageyama, T. Shimogori, and A.W. Moore. 2008. Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. *PLoS One* **3:** e3859.

Koblar, S.A., M. Murphy, G.L. Barrett, A. Underhill, P. Gros, and P.F. Bartlett. 1999. Pax-3 regulates neurogenesis in neural crest-derived precursor cells. *J Neurosci Res* **56:** 518-530.

Kolomietz, E., P. Marrano, K. Yee, B. Thai, I. Braude, A. Kolomietz, K. Chun, S. Minkin, S. Kamel-Reid, M. Minden et al. 2003. Quantitative PCR identifies a minimal deleted region of 120 kb extending from the Philadelphia chromosome ABL translocation breakpoint in chronic myeloid leukemia with poor outcome. *Leukemia* **17:** 1313-1323.

Krauss, R.S., F. Cole, U. Gaio, G. Takaesu, W. Zhang, and J.S. Kang. 2005. Close encounters: regulation of vertebrate skeletal myogenesis by cell-cell contact. *J Cell Sci* **118:** 2355-2362.

Kubic, J.D., K.P. Young, R.S. Plummer, A.E. Ludvik, and D. Lang. 2008. Pigmentation PAX-ways: the role of Pax3 in melanogenesis, melanocyte stem cell maintenance, and disease. *Pigment Cell Melanoma Res* **21:** 627-645.

Lagha, M., T. Sato. L. Bajard, P. Daubas, M. Esner, D. Montarras, F. Relaix, and M. Buckingham. 2008. Regulation of skeletal muscle stem cell behavior by Pax3 and Pax7. *Cold Spring Harb Symp Quant Biol* **73:** 307-315.

Lagutina, I., S.J. Conway, J. Sublett, and G.C. Grosveld. 2002. Pax3-FKHR knock-in mice show developmental aberrations but do not develop tumors. *Mol Cell Biol* **22:** 7204-7216.

Lang, D., F. Chen, R. Milewski, J. Li, M.M. Lu, and J.A. Epstein. 2000. Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. *J Clin Invest* **106:** 963-971.

Lang, D., M.M. Lu, L. Huang, K.A. Engleka, M. Zhang, E.Y. Chu, S. Lipner, A. Skoultchi. S.E. Millar, and J.A. Epstein. 2005. Pax3 functions at a nodal point in melanocyte stem cell differentiation. *Nature* **433:** 884-887.

Lang, D., S.K. Powell, R.S. Plummer, K.P. Young, and B.A. Ruggeri. 2007. PAX genes: roles in development, pathophysiology, and cancer. *Biochem Pharmacol* **73:** 1-14.

Li, H.G., Q. Wang, H.M. Li, S. Kumar, C. Parker, M. Slevin, and P. Kumar. 2007. PAX3 and PAX3-FKHR promote rhabdomyosarcoma cell survival through downregulation of PTEN. *Cancer Lett* **253:** 215-223.

Lim, C.B., D. Zhang, and C.G. Lee. 2006. FAT10, a gene up-regulated in various cancers, is cell-cycle regulated. *Cell Div* **1:** 20.

Lorain, S., J.P. Quivy. F. Monier-Gavelle, C. Scamps, Y. Lecluse, G. Almouzni, and M. Lipinski. 1998. Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. *Mol Cell Biol* **18:** 5546-5556.

Ludwig, A., S. Rehberg, and M. Wegner. 2004. Melanocyte-specific expression of dopachrome tautomerase is dependent on synergistic gene activation by the Sox10 and Mitf transcription factors. *FEBS Lett* **556:** 236-244.

Mackenzie, M.A., S.A. Jordan, P.S. Budd, and I.J. Jackson. 1997. Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev Biol* **192:** 99-107.

Malumbres, M. and M. Barbacid. 2009. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* **9:** 153-166.

Maniatis, T. 1999. A ubiquitin ligase complex essential for the NF-kappaB, Wnt/Wingless, and Hedgehog signaling pathways. *Genes Dev* **13:** 505-510.

Margue, C.M., M. Bernasconi, F.G. Barr, and B.W. Schafer. 2000. Transcriptional modulation of the anti-apoptotic protein BCL-XL by the paired box transcription factors PAX3 and PAX3/FKHR. *Oncogene* **19:** 2921-2929.

Marquardt, T., R. Ashery-Padan, N. Andrejewski, R. Scardigli, F. Guillemot, and P. Gruss. 2001. Pax6 is required for the multipotent state of retinal progenitor cells. *Cell* **105:** 43-55.

Mascarenhas, J.B., E.L. Littlejohn, R.J. Wolsky, K.P. Young, M. Nelson, R. Salgia, and D. Lang. 2010. PAX3 and SOX10 activate MET receptor expression in melanoma. *Pigment Cell Melanoma Res* **23:** 225-237.

Mayanil, C.S., A. Pool. H. Nakazaki, A.C. Reddy, B. Mania-Farnell, B. Yun, D. George, D.G. McLone, and E.G. Bremer. 2006. Regulation of murine

118

TGFbeta2 by Pax3 during early embryonic development. *J Biol Chem* **281:** 24544-24552.

McCue, L., W. Thompson, C. Carmack, M.P. Ryan, J.S. Liu, V. Derbyshire, and C.E. Lawrence. 2001. Phylogenetic footprinting of transcription factor binding sites in proteobacterial genomes. *Nucleic Acids Res* **29:** 774-782.

Medic, S. and M. Ziman. 2010. PAX3 expression in normal skin melanocytes and melanocytic lesions (naevi and melanomas). *PLoS One* **5:** e9977.

Megeney, L.A., B. Kablar, K. Garrett, J.E. Anderson, and M.A. Rudnicki. 1996. MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev* **10:** 1173-1183.

Mennerich, D. and T. Braun. 2001. Activation of myogenesis by the homeobox gene Lbx1 requires cell proliferation. *EMBO J* **20:** 7174-7183.

Mennerich, D., K. Schafer, and T. Braun. 1998. Pax-3 is necessary but not sufficient for lbx1 expression in myogenic precursor cells of the limb. *Mech Dev* **73:** 147-158.

Mise, T., M. Iijima, K. Inohaya, A. Kudo, and H. Wada. 2008. Function of Pax1 and Pax9 in the sclerotome of medaka fish. *Genesis* **46:** 185-192.

Mitra, D. and D.E. Fisher. 2009. Transcriptional regulation in melanoma. *Hematol Oncol Clin North Am* **23:** 447-465, viii.

Moase, C.E. and D.G. Trasler. 1992. Splotch locus mouse mutants: models for neural tube defects and Waardenburg syndrome type I in humans. *J Med Genet* **29:** 145-151.

Montarras, D., C. Lindon, C. Pinset, and P. Domeyne. 2000. Cultured myf5 null and myoD null muscle precursor cells display distinct growth defects. *Biol Cell* **92:** 565-572.

Mularoni, L., A. Ledda, M. Toll-Riera, and M.M. Alba. 2010. Natural selection drives the accumulation of amino acid tandem repeats in human proteins. *Genome Res* **20:** 745-754.

Mulieri, P.J., J.S. Kang, D.A. Sassoon, and R.S. Krauss. 2002. Expression of the boc gene during murine embryogenesis. *Dev Dyn* **223:** 379-388.

Muratovska, A., C. Zhou, S. He, P. Goodyer, and M.R. Eccles. 2003. Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival. *Oncogene* **22:** 7989-7997.

Neubuser, A., H. Koseki, and R. Balling. 1995. Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1. *Dev Biol* **170:** 701-716.

Nishikawa, S. and M. Osawa. 2007. Generating quiescent stem cells. *Pigment Cell Res* **20:** 263-270.

Nishimura, E.K., S.A. Jordan, H. Oshima, H. Yoshida, M. Osawa, M. Moriyama, I.J. Jackson, Y. Barrandon, Y. Miyachi, and S. Nishikawa. 2002. Dominant role of the niche in melanocyte stem-cell fate determination. *Nature* **416:** 854-860.

Nobukuni, Y., A. Watanabe, K. Takeda, H. Skarka, and M. Tachibana. 1996. Analyses of loss-of-function mutations of the MITF gene suggest that haploinsufficiency is a cause of Waardenburg syndrome type 2A. *Am J Hum Genet* **59:** 76-83.

Norbury, C. and P. Nurse. 1992. Animal cell cycles and their control. *Annu Rev Biochem* **61:** 441-470.

O'Brien, C.A., A. Kreso, and J.E. Dick. 2009. Cancer stem cells in solid tumors: an overview. *Semin Radiat Oncol* **19:** 71-77.

Onder, T.T., P.B. Gupta, S.A. Mani, J. Yang, E.S. Lander, and R.A. Weinberg. 2008. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* **68:** 3645-3654.

Opdecamp, K., A. Nakayama, M.T. Nguyen, C.A. Hodgkinson, W.J. Pavan, and H. Arnheiter. 1997. Melanocyte development in vivo and in neural crest cell cultures: crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription factor. *Development* **124:** 2377-2386.

Osawa, M., G. Egawa, S.S. Mak, M. Moriyama, R. Freter, S. Yonetani, F. Beermann, and S. Nishikawa. 2005. Molecular characterization of melanocyte stem cells in their niche. *Development* **132:** 5589-5599.

Ovcharenko, I., M.A. Nobrega, G.G. Loots, and L. Stubbs. 2004. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res* **32:** W280-286.

Pani, L., M. Horal, and M.R. Loeken. 2002. Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3-dependent development and tumorigenesis. *Genes Dev* **16:** 676-680.

120

Parker, C.J., S.G. Shawcross, H. Li, Q.Y. Wang, C.S. Herrington, S. Kumar, R.M. MacKie, W. Prime, I.G. Rennie, K. Sisley et al. 2004. Expression of PAX 3 alternatively spliced transcripts and identification of two new isoforms in human tumors of neural crest origin. *Int J Cancer* **108:** 314-320.

Pavesi, G., G. Mauri, and G. Pesole. 2004. In silico representation and discovery of transcription factor binding sites. *Brief Bioinform* **5:** 217-236.

Pingault, V., D. Ente, F. Dastot-Le Moal, M. Goossens, S. Marlin, and N. Bondurand. 2010. Review and update of mutations causing Waardenburg syndrome. *Hum Mutat* **31:** 391-406.

Plachov, D., K. Chowdhury, C. Walther, D. Simon, J.L. Guenet, and P. Gruss. 1990. Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development* **110:** 643-651.

Plummer, R.S., C.R. Shea, M. Nelson, S.K. Powell, D.M. Freeman, C.P. Dan, and D. Lang. 2008. PAX3 expression in primary melanomas and nevi. *Mod Pathol* **21:** 525-530.

Poleev, A., H. Fickenscher, S. Mundlos, A. Winterpacht, B. Zabel, A. Fidler, P. Gruss, and D. Plachov. 1992. PAX8, a human paired box gene: isolation and expression in developing thyroid, kidney and Wilms' tumors. *Development* **116:** 611-623.

Prabhakar, S., F. Poulin, M. Shoukry, V. Afzal, E.M. Rubin, O. Couronne, and L.A. Pennacchio. 2006. Close sequence comparisons are sufficient to identify human cis-regulatory elements. *Genome Res* **16:** 855-863.

Pritchard, C., G. Grosveld, and A.D. Hollenbach. 2003. Alternative splicing of Pax3 produces a transcriptionally inactive protein. *Gene* **305:** 61-69.

Read, A.P. and V.E. Newton. 1997. Waardenburg syndrome. *J Med Genet* **34:** 656-665.

Reid, A.G. and E.P. Nacheva. 2004. A potential role for PRDM12 in the pathogenesis of chronic myeloid leukaemia with derivative chromosome 9 deletion. *Leukemia* **18:** 178-180.

Relaix, F. 2006. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J. Cell Biol.* **172:** 91-102.

Relaix, F., D. Rocancourt, A. Mansouri, and M.A. Buckingham. 2005. Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435:** 948-953.

Rudnicki, M.A., T. Braun, S. Hinuma, and R. Jaenisch. 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell* **71**: 383-390.

Rudnicki, M.A., P.N. Schnegelsberg, R.H. Stead, T. Braun, H.H. Arnold, and R. Jaenisch. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* **75**: 1351-1359.

Ryu, B., D.S. Kim, A.M. Deluca, and R.M. Alani. 2007. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PLoS One* **2**: e594.

Salomoni, P. and F. Calegari. 2010. Cell cycle control of mammalian neural stem cells: putting a speed limit on G1. *Trends Cell Biol* **20**: 233-243.

Samadi, N., C. Gaetano, I.S. Goping, and D.N. Brindley. 2009. Autotaxin protects MCF-7 breast cancer and MDA-MB-435 melanoma cells against Taxol-induced apoptosis. *Oncogene* **28**: 1028-1039.

Sanyanusin, P., L.A. McNoe, M.J. Sullivan, R.G. Weaver, and M.R. Eccles. 1995a. Mutation of PAX2 in two siblings with renal-coloboma syndrome. *Hum Mol Genet* **4**: 2183-2184.

Sanyanusin, P., L.A. Schimmenti, L.A. McNoe, T.A. Ward, M.E. Pierpont, M.J. Sullivan, W.B. Dobyns, and M.R. Eccles. 1995b. Mutation of the PAX2 gene in a family with optic nerve colobomas, renal anomalies and vesicoureteral reflux. *Nat Genet* **9**: 358-364.

Sato, T., N. Sasai, and Y. Sasai. 2005. Neural crest determination by co-activation of Pax3 and Zic1 genes in Xenopus ectoderm. *Development* **132**: 2355-2363.

Schebesta, A., S. McManus, G. Salvagiotto, A. Delogu, G.A. Busslinger, and M. Busslinger. 2007. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity* **27**: 49-63.

Scholl, F.A., J. Kamarashev, O.V. Murmann, R. Geertsen, R. Dummer, and B.W. Schafer. 2001. PAX3 is expressed in human melanomas and contributes to tumor cell survival. *Cancer Res* **61**: 823-826.

Smith, A.P., K. Hoek, and D. Becker. 2005. Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular

differences between nevi/melanoma in situ and advanced-stage melanomas. *Cancer Biol Ther* **4:** 1018-1029.

Sosa-Pineda, B., K. Chowdhury, M. Torres, G. Oliver, and P. Gruss. 1997. The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. *Nature* **386:** 399-402.

St-Onge, L., B. Sosa-Pineda, K. Chowdhury, A. Mansouri, and P. Gruss. 1997. Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature* **387:** 406-409.

Steel, K.P., D.R. Davidson, and I.J. Jackson. 1992. TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development* **115:** 1111-1119.

Sturzu, A., U. Klose, H. Echner, A. Beck, A. Gharabaghi, H. Kalbacher, and S. Heckl. 2009. Cellular uptake of cationic gadolinium-DOTA peptide conjugates with and without N-terminal myristoylation. *Amino Acids* **37:** 249-255.

Sun, X.J., P.F. Xu, T. Zhou, M. Hu, C.T. Fu, Y. Zhang, Y. Jin, Y. Chen, S.J. Chen, Q.H. Huang et al. 2008. Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. *PLoS One* **3:** e1499.

Tajbakhsh, S. and M. Buckingham. 2000. The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. *Curr Top Dev Biol* **48:** 225-268.

Tajbakhsh, S., D. Rocancourt, and M. Buckingham. 1996. Muscle progenitor cells failing to respond to positional cues adopt non-myogenic fates in myf-5 null mice. *Nature* **384:** 266-270.

Tajbakhsh, S., D. Rocancourt, G. Cossu, and M. Buckingham. 1997. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89:** 127-138.

Tassabehji, M., V.E. Newton, and A.P. Read. 1994. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. *Nat Genet* **8:** 251-255.

Tassabehji, M., A.P. Read, V.E. Newton, R. Harris, R. Balling, P. Gruss, and T. Strachan. 1992. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355:** 635-636.

Tenzen, T., B.L. Allen, F. Cole, J.S. Kang, R.S. Krauss, and A.P. McMahon. 2006. The cell surface membrane proteins Cdo and Boc are components and targets of the Hedgehog signaling pathway and feedback network in mice. *Dev Cell* **10:** 647-656.

Torres, M., E. Gomez-Pardo, G.R. Dressler, and P. Gruss. 1995. Pax-2 controls multiple steps of urogenital development. *Development* **121:** 4057-4065.

Treisman, J., E. Harris, and C. Desplan. 1991. The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes Dev* **5:** 594-604.

Tremblay, P. and P. Gruss. 1994. Pax: genes for mice and men. *Pharmacol Ther* **61:** 205-226.

Tremblay, P., M. Kessel, and P. Gruss. 1995. A transgenic neuroanatomical marker identifies cranial neural crest deficiencies associated with the Pax3 mutant Splotch. *Dev Biol* **171:** 317-329.

Tsukamoto, K., Y. Nakamura, and N. Niikawa. 1994. Isolation of two isoforms of the PAX3 gene transcripts and their tissue-specific alternative expression in human adult tissues. *Hum Genet* **93:** 270-274.

Tsukamoto, K., T. Tohma, T. Ohta, K. Yamakawa, Y. Fukushima, Y. Nakamura, and N. Niikawa. 1992. Cloning and characterization of the inversion breakpoint at chromosome 2q35 in a patient with Waardenburg syndrome type I. *Hum Mol Genet* **1:** 315-317.

Underhill, D.A. 2000. Genetic and biochemical diversity in the Pax gene family. *Biochem Cell Biol* **78:** 629-638.

Underhill, D.A., K.J. Vogan, and P. Gros. 1995. Analysis of the mouse Splotch-delayed mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity. *Proc Natl Acad Sci U S A* **92:** 3692-3696.

Urbanek, P., Z.Q. Wang, I. Fetka, E.F. Wagner, and M. Busslinger. 1994. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* **79:** 901-912.

van den Broeke, L.T., C.D. Pendleton, C. Mackall, L.J. Helman, and J.A. Berzofsky. 2006. Identification and epitope enhancement of a PAX-FKHR fusion protein breakpoint epitope in alveolar rhabdomyosarcoma cells created by a tumorigenic chromosomal translocation inducing CTL capable of lysing human tumors. *Cancer Res* **66:** 1818-1823.

Waardenburg, P.J. 1951. A new syndrome combining developmental anomalies of the eyelids, eyebrows and nose root with pigmentary defects of the iris and head hair and with congenital deafness. *Am J Hum Genet* **3:** 195-253.

Walther, C., J.L. Guenet, D. Simon, U. Deutsch, B. Jostes, M.D. Goulding, D. Plachov, R. Balling, and P. Gruss. 1991. Pax: a murine multigene family of paired box-containing genes. *Genomics* **11:** 424-434.

Wang, Q., W.H. Fang, J. Krupinski, S. Kumar, M. Slevin, and P. Kumar. 2008. Pax genes in embryogenesis and oncogenesis. *J Cell Mol Med* **12:** 2281-2294.

Wang, Q., S. Kumar, M. Slevin, and P. Kumar. 2006. Functional analysis of alternative isoforms of the transcription factor PAX3 in melanocytes in vitro. *Cancer Res* **66:** 8574-8580.

Watanabe, A., K. Takeda, B. Ploplis, and M. Tachibana. 1998. Epistatic relationship between Waardenburg syndrome genes MITF and PAX3. *Nat Genet* **18:** 283-286.

Wiggan, O., A. Taniguchi-Sidle, and P.A. Hamel. 1998. Interaction of the pRB-family proteins with factors containing paired-like homeodomains. *Oncogene* **16:** 227-236.

Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan. 1993. Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev* **7:** 2120-2134.

Wright, W.E. and W.D. Funk. 1993. CASTing for multicomponent DNA-binding complexes. *Trends Biochem Sci* **18:** 77-80.

Wu, J., J.P. Saint-Jeannet, and P.S. Klein. 2003. Wnt-frizzled signaling in neural crest formation. *Trends Neurosci* **26:** 40-45.

Wu, M., J. Li, K.A. Engleka, B. Zhou, M.M. Lu, J.B. Plotkin, and J.A. Epstein. 2008. Persistent expression of Pax3 in the neural crest causes cleft palate and defective osteogenesis in mice. *J Clin Invest* **118:** 2076-2087.

Yang, G., Y. Li, E.K. Nishimura, H. Xin, A. Zhou, Y. Guo, L. Dong, M.F. Denning, B.J. Nickoloff, and R. Cui. 2008. Inhibition of PAX3 by TGF-beta modulates melanocyte viability. *Mol Cell* **32:** 554-563.

Yang, X.M., K. Vogan, P. Gros, and M. Park. 1996. Expression of the met receptor tyrosine kinase in muscle progenitor cells in somites and limbs is absent in Splotch mice. *Development* **122:** 2163-2171.

Zhang, L. and C. Wang. 2007. Identification of a new class of PAX3-FKHR target promoters: a role of the Pax3 paired box DNA binding domain. *Oncogene* **26:** 1595-1605.

Zhang, R., M.V. Poustovoitov, X. Ye, H.A. Santos, W. Chen, S.M. Daganzo, J.P. Erzberger, I.G. Serebriiskii, A.A. Canutescu, R.L. Dunbrack et al. 2005. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev Cell* **8:** 19-30.

APPENDIX I. PERL SCRIPT FOR MOTIF SCAN

```perl
#!/usr/bin/perl
#
# This program searches given motifs through fasta sequences (e.g. genome)
# Usage: perl motif_scan.pl
# Author: Yisu Li, Dec. 2007
###################################################################################

use strict;
use warnings;

###
# Declare the input file, hits file, output file and the title of describe lines in the #fasta
output
my $inputfile = "Mus_musculus.NCBIM37.50.dna_rm.chromosome.1.fa";
my $hitsfile = "motif_scan_mm_chr1.tmp";
my $outputfile = "motif_scan_mm_chr1.fa";
my $describe_line = "mm_chr1";

# Open the file containing the DNA sequence data, or exit if the file doesn't exist or
#cannot be opened
open (IN, "$inputfile") || die ("Cannot open the file!\n");
# Read the DNA sequence data from the file and store it into an array
my @dna = <IN>;
# Close the file - we've read all the data into @dna now.
close IN;

# Put the DNA sequence data into a single string, as it's easier to search for a motif #in a
string than in an array and this can avoid missing motifs cross two lines
my $DNA = join( '', @dna);
# Remove whitespace
$DNA =~ s/\s//g;

###
# Search for motifs
# Find the location of each hit and print out the segment including 15nt before and #after
the hit sequence
open (TEMP, ">$hitsfile") || die ("Cannot write the file!\n");
# Call to use the sub 'search's to search for the given motif sequences
# (e.g. ATT.GTCA[C|T]G[C|G]T) and the sub will print out the result sequences
&search("ATT.GTCA[C|T]G[C|G]T","A[C|G]C[A|G]TGAC.AAT");
print TEMP "\n";

###
# Subroutines
#
# This subroutin searches all the matches for requested motifs
# and prints out the segment sequence including 15nt prior and after.
sub search {
        my $motif_seq = $_[0];
```

128

```perl
            my $compli_motif = $_[1];
            print TEMP "$motif_seq composite motifs:\n";
            while ($DNA =~ m/$motif_seq/g) {
                    my $start = pos($DNA)-42;
                    my $seq = substr $DNA, $start, 72;
                    print TEMP "$seq\n";
            }
            while ($DNA =~ m/$compli_motif/g) {
                    my $start = pos($DNA)-42;
                    my $seq = substr $DNA, $start, 72;
                    # if the DNA matches the complimentary sequence of the given motif,
                    # use sub 'complimentary_nt' to convert nt
                    &complimentary_nt($seq);
                    print TEMP "\n";
            }
}

# This subroutin converts nt to its complimentary nt
# Eg: A --> T, G --> C
sub complimentary_nt {
        while ($_[0]) {
                my $current_nt = chop($_[0]);
                if ($current_nt =~ /A/) {print OUT "t";}
                if ($current_nt =~ /T/) {print OUT "a";}
                if ($current_nt =~ /C/) {print OUT "g";}
                if ($current_nt =~ /G/) {print OUT "c";}
                if ($current_nt =~ /N/) {print OUT "n";}
        }
}
close TEMP;

###
# Generate output file in fasta format
# open the file containing hits sequences. or exit if the file doesn't exist or cannot be
#opened
open (TEMP, "$hitsfile") || die("Cannot open the file!\n");
open (OUT, ">$outputfile ") || die ("Cannot write the file!\n");
my $count = 1;
while (<TEMP>) {
        chomp;
        # eliminates the first title line and any empty lines
        if ($_ =~ /[a-zA-Z]/) {
                unless ($_ =~ /composite/) {
                        print OUT ">$describe_line";
                        print OUT "_$count";
                        print OUT "\n";
                        print OUT "$_\n";
                        $count++;
                }
```

```
        }
}
close TEMP;
close OUT;

# remove the tmp file
unlink $hitsfile;

# exit the program
exit;

__END__
```

APPENDIX II. BLAST SEQUENCES FROM

MOUSE MOTIF-SCAN AND PAX3 CASTING

LIBRARIES

NOTE: Mouse motif-scan sequences with lower-case letters are the reverse-complimentary sequences of the original chromosome sequences.

AII.1 Mouse motif-scan sequences
>mouse_chr1_seq16 29135bp_at_5'
side_hypothetical_protein_XP_002342276|338021bp_at_3'_side_hypothetical_pr
otein_XP_002342277
AACATTTAATTTAAACAACAATGCTCTGTAATTTGTCATGGTGCTCCCA
AAGTGAGAAATGTTGCAACTTTG
>mouse_chr1_seq122 582057bp_at_5'_side_carbamoyl-
phosphate_synthetase_1_isoform_c|123503bp_at_3'_side_v-erb-
a_erythroblastic_leukemia_viral_oncogene_homolog_4
gattactttgcagccactctgggaacaccaatttgtcatgctagatcaagataaccacagcctgactagagg
>mouse_chr1_seq126
392812bp_at_5'_side_solute_carrier_family_4_anion_exchanger_member_3_isof
orm_2|1391415bp_at_3'_side_ephrin_receptor_EphA4
tttttacaactgagaaaatccagatttcttattagtcatggtctaaaactgacaaccccctgcctatgcagc
>mouse_chr2_seq15 hypothetical_protein
CTATCTAGGAGCTTGTAAACACATTGGTGAATTTGTCATGCTGATTATA
GTGTTGGCAAACACAATGGGTGA
>mouse_chr2_seq44 127637bp_at_5'_side_gamma-
butyrobetaine_dioxygenase|83667bp_at_3'_side_coiled-
coil_domain_containing_34_isoform_1
AGAAATTAAAAAACAAAAGATTATCCCTTCATTGGTCATGGTTATTCA
AGTTTCTTATATCATGCTAGCCCA
>mouse_chr2_seq82 PR_domain_containing_12
taaacggtgtccattgcggagcggccggcaatttgtcacgctgttaaagtgatcgcattcattcggttccat
>mouse_chr2_seq131
715429bp_at_5'_side_paired_box_1|151570bp_at_3'_side_forkhead_box_A2_isof
orm_1
ctcactgggcaccatgctcgctgcaattcaattagtcatggtcagagggaggcattattattcatacatcac
>mouse_chr2_seq139
155413bp_at_5'_side_transcription_factor_MAFB|184734bp_at_3'_side_DNA_to
poisomerase_I
aaaagactggccaatttacaaattacacgtatttgtcatgctcttgtacctttcctgagctaaccgaaatgg
>mouse_chr3_seq31 7659bp_at_5'_side_hepatitis_B_virus_x-
interacting_protein|35625bp_at_3'_side_prokineticin_1
CATCTTTTTATTCATAAGGGCTTCATTCCCATTTGTCATGGTTGCATTA
ACTGGGCATCTAATAAAGGCTTG
>mouse_chr3_seq33 dihydropyrimidine_dehydrogenase
TGAACACTTGGGGATTTAATTGCTCATTATATTTGTCACGGTGGTGTCA
CAGACCCAAATAAAGACTGTTGC
>mouse_chr4_seq37
135081bp_at_5'_side_APG4_autophagy_4_homolog_C_isoform_7|323748bp_at_

132

3'_side_forkhead_box_D3|1236bp_at_5'_side_hypothetical_protein|135099bp_at_
3'_side_APG4_autophagy_4_homolog_C_isoform_8
AATTGCTCATCCTAATGTGCAGTGATCTGTATTCGTCATGCTCCGAAGA
GCCAGCCATCTCTACGCCGCCA
>mouse_chr4_seq83
zinc_finger_protein_618|4753bp_at_5'_side_zinc_finger_protein_618|367009bp_
at_3'_side_ regulator_of_G-protein_signalling_3_isoform_6
aattctaagcaaatcaatcaaaaattaatgatttgtcatgcttgcacaatggaggaaaaacagctccttggg
>mouse_chr5_seq80 135369bp_at_5'_side_DEAH_(Asp-Glu-Ala-
His)_box_polypeptide_15|79697bp_at_3'_side_superoxide_dismutase_3_extracell
ular_precursor
acatgaagcactgagtagaaggcagcaggcattagtcacgctggaggaagttgcagataaattgtaaagagt
>mouse_chr6_seq84 similar_to_hCG1644697|homeobox_A5
cagatctacccctggatgcgcaagctgcacattagtcacggtaaagccagccgttttctgatccacgcgtcc
>mouse_chr6_seq94
3735bp_at_5'_side_hypothetical_protein_XP_002342252|42659bp_at_3'_side_tet
_oncogene_family_member_3
cactcctgagtctgaagctgatagggtgtcattagtcatgctgctgtcaccatggagatggctgaagagagg
>mouse_chr6_seq97 coiled-coil-helix-coiled-coil-helix_domain_containing_6
gcagctattgcaagcgagtcttctcctgtcatttgtcatggtttggagctttgaaagtgacagtttgctttc
>mouse_chr7_seq10 neuron_navigator_2_isoform_3
TGAGGCTCGGCGGCAGCTGGGGTAACATCAATTGGTCATGCTCAGCTG
ACATTCATTTTAACTTGAACTCTA
>mouse_chr8_seq77
222932bp_at_5'_side_Nedd4_binding_protein_1|446439bp_at_3'_side_cerebellin
_1_precursor
ggaagagtgggttatacaatcaggtgtgtgattagtcatgctaatggcgctttattatactgtgagtagaac
>mouse_chr9_seq12
1324bp_at_5'_side_immunoglobulin_superfamily_member_4D_isoform_1|15350
1bp_at_3'_side_hypothetical_protein_XP_002343162
AGTAATCTCTTGGTCATTTAGCCACATATAATTTGTCATGGTTTGTTTT
AGCTGTTGTGATATTGATATTTA
>mouse_chr9_seq15 289929bp_at_5'_side_transducin-
like_enhancer_protein_3_isoform_c|270256bp_at_3'_side_uveal_autoantigen_wit
h_coiled-coil_domains_and_ankyrin_re…
CTCTGCCCTAATTGAGCAGATCATCAGCCGATTTGTCATGCTTAGCCCA
GACCCCTAACTGCCACCCCTCCC
>mouse_chr9_seq40
158847bp_at_5'_side_RNA_binding_motif_single_stranded_interacting_protein_
3…|444118bp_at_3'_side_transforming_growth_factor_beta_receptor_II_isoform_
A_pr…
CTGTTCGTCTTAATAGCAGTTGCACATCTAATTGGTCATGCTGAAAATA
AGCTCCATTATATATATTTTTTC

133

>mouse_chr10_seq47
21401bp_at_5'_side_RAB21_member_RAS_oncogene_family|32644bp_at_3'_sid
e_TBC1_domain_family_member_15_isoform_1
CTTGTTACAAGATTGTATTTCCATTCACAAATTGGTCACGGTGTATTAA
AACGCAAATGAGGTAGCCATGAT
>mouse_chr11_seq4
464397bp_at_5'_side_hypothetical_protein_XP_002342262|237344bp_at_3'_side
_ETAA16_protein
AGCTCTTTCTAATTTGTTTAGCTAAACTGTATTAGTCATGCTGAAAGTT
TTATTATCAGAGTCAAATACAGG
>mouse_chr11_seq27
14968bp_at_5'_side_mitochondrial_rRNA_methyltransferase_1_homolog|315606
bp_at_3'_side_LIM_homeobox_protein_1
GACCCCTAATTAGTTGCTGCTGATAACTATATTTGTCATGCTCTCCAAT
TATTCCACTGTAAATTTCAGAGT
>mouse_chr11_seq74 31614bp_at_5'_side_Cdc2-related_kinase_arginine/serine-
rich_isoform_2|42450bp_at_3'_siden_eurogenic_differentiation
2284270bp_at_5'_side_ETAA16_protein|348534bp_at_3'_side_nuclear_DNA-
binding_protein
ttttcccaatgatttttacatgctccagccattagtcatgctatattttatcagggaaatgagtttgcttaa
>mouse_chr12_seq22
474218bp_at_5'_side_forkhead_box_G1|334200bp_at_3'_side_protein_kinase_D
1
ACTGTTGCTTCCCACTGAGCCTTTATCTCTATTTGTCATGGTGCCAGAA
CTGCCTGCATAGCACCTTTCATG
>mouse_chr12_seq23
685453bp_at_5'_side_forkhead_box_G1|122967bp_at_3'_side_protein_kinase_D
1
GGCAAAGGGATGTTTTTTTGCCCTGAGGGTATTAGTCATGGTCCATAG
TTCGTGGGAAAAACTAAGGAAGTT
>mouse_chr13_seq29
40853bp_at_5'_side_hypothetical_protein_XP_002342573|1469616bp_at_3'_side
_centrin_3
CTGGGCTGAAAAGATATGGCTTGATGGGTAATTTGTCATGCTGTGCAA
TAGAGAACAACAGTAAAGAGACTT
>mouse_chr13_seq67
206853bp_at_5'_side_hypothetical_protein_XP_002342573|1303616bp_at_3'_sid
e_centrin_3
aatttcaaacacccagttcataaattatttattggtcatggtcaaatgaatgttagctatttcataattatc
>mouse_chr14_seq36
3540681bp_at_5'_side_sprouty_2|969bp_at_3'_side_slit_and_trk_like_1_protein
GGTCGCAGCCTAAGCACTAGAGTGACATTAATTGGTCATGCTTAACTG
CCCCAGAATCTTTTTTTCTTAAAT
>mouse_chr14_seq62 ecto-NOX_disulfide-thiol_exchanger_1
tttcatggtaggttgctaggaaaatcataaattggtcatggttaaaatgtcaaaagactcaaaattgtgagg

134

>mouse_chr14_seq65 diaphanous_homolog_3_isoform_a
acacctttgtgcagttggtggaatgctaaaatttgtcatgcttttattattgttgccttgaattaactctgc
>mouse_chr14_seq71
1098655bp_at_5'_side_hypothetical_protein|1371587bp_at_3'_side_kelch-
like_1_protein
aataattaaagtgtgcatgtaccacaacctatttgtcatggtaactagacagcataagacattaacatgatt
>mouse_chr14_seq76
536944bp_at_5'_side_dachshund_homolog_1_isoform_a|306538bp_at_3'_side_h
ypothetical_protein_LOC440145
atttattcacatatctatttaaaaaattccattggtcatgctaaaaaaggttgcaagccctttgacttaagt
>mouse_chr14_seq79
339305bp_at_5'_side_ring_finger_protein_219|322131bp_at_3'_side_RNA_bindi
ng_motif_protein_26
tcaagaattgcaagctgcattattaaagtcatttgtcatggtttgaatatttctgcctgattcaccattaaa
>mouse_chr15_seq59
11524bp_at_5'_side_homeobox_C12|5082bp_at_3'_side_homeobox_C11
cttgtgcgtgtacttgtgtgttgggtgagaatttgtcacggtgttctttgttctgggcattttaagtgttcc
>mouse_chr16_seq8
kalirin_RhoGEF_kinase_isoform_1|kalirin_RhoGEF_kinase_isoform_2
AGAAGTTAAAGAGCTTTAAATGACTGATCCATTTGTCACGGTGATCAA
TACATTACTGTCTTTAAGGCACTT
>mouse_chr17_seq31
29506bp_at_5'_side_similar_to_hCG2040074|59391bp_at_3'_side_suppressor_of
_cytokine_signaling_5
GAGTCCCGCCCACCTGGAGCCCCGCCTGCCATTGGTCACGGTAGGCTG
CCCTTCAGAGAGGCGTGGCCCAGC
>mouse_chr19_seq8
71161bp_at_5'_side_hypothetical_protein_isoform_1|50495bp_at_3'_side_beta-
transducin_repeat_containing_protein_isoform_1|66317bp_at_3'_side_hypothetic
al_protein_isoform_2
CTCTGATCTCCAAATTGTTCCTTTTATTGCATTTGTCACGGTGAAGATT
AAAGTCCCTGGGTGGCCTGCACT
>mouse_chrX_seq70 fragile_X_mental_retardation_2
ggctgatcgggagctgcttttgcatgtgtaattagtcacggtgcacctgtacctaactgcattaacaattaa

AII.2 Mouse PAX3-CASTing sequences
>embl|AZ755438|AZ755438 382986bp_at_5'_side_peroxisome_proliferator-
activated_receptor_gamma,_coactiv...|197554bp_at_3'_side_hypothetical_protein
gatctaattcacaggagtgatcatgcagacccaggtacagtattcattattaacttaaaaagtatgccaaatcaattaaga
agcattttaaagggaaggaataaaatttagttctacctaactgcggctcagctctaatttatctcccagccactttcaagg
gattttacttggttcttaatctaacattcattggaaagggtaattctcatgaaattacaatgcatctcttgcatcgtttaatgca
aatgcccctcttcttaggtaatggagagcggcttctccaataaagactccnttcaaagtgcaaagtgggggagccctg
ggagctaccgtgaagaacaaaatactgcatttctgaagctgctaagacttaaaac
>embl|AZ755470|AZ755470 MACRO_domain_containing_1

gatcaaaggagcaccgggtgggaagcagccccgtgtgtgtcccctgtctcagtgtgctgggtgaggactatttgatg
cctgcagntgtgggggattgggcctattacagaggctcggaggtttatccaatgccttttccaacgccccctttctccac
ctgttttcatatactccaagaagaccaagggaatcttcaaagtccagactgagccaagagcagaaattctttaattagc
tttctcttggattggggccctggcgaatggattagtcatggtgggggggcaggaggaagggatc
>embl|AZ755478|AZ755478 RAD54-like_2
gatctccaatgtcaacatgtagttttaaatcttttctgtgggtaagagaaagaacctgggctaggtgctggagtgtaaga
aaacccctaactataagtgcttggttacagcgcactggaaaaggcaaagggcacgatgacacttgatggtcttgctct
gtagctactgcactcgactgtcctacagatattaagcatcttctcacctaaatacaagaatcttgtccccgagcttcacgc
tttctttaatcaggtggaaaagcagtaccatcttgggagagttctccaggactccagtctgataattagttcagaagctcc
ttggcctgcaaaagagaagcaaggagattggggcaggtc
>embl|AZ755485|AZ755485 FAT_tumor_suppressor_1_precursor
gatcctcacgccaatgatttcctcaacgtccgtcacggaagaattaatcttatgcagaagttgctttgttgaaacgtgggt
gccgcctgacctctcgacaaagagtaagacgtccagatgcgaatgggggttcggagggctgcaggctgacgatctgt
atgtcattcttccggacgcccaggatgttgcgtagggctcgctggaagttgcgccagtagtcaccaacagactcctcc
ggggtgagattcgcgaanatgggncgacggtgtggttcagcatctcctgggtcacttgttggatgtgcacggtgatgt
tcggccaccgtggtgaatttcccatctgtcacgctgacattgagaaggtactggcctatgtccagctttc
>embl|AZ755506|AZ755506
11183bp_at_5'_side_nuclear_receptor_subfamily_4,_group_A,_member_2|13455
2bp_at_3'_side_glycerol-3-phosphate_dehydrogenase_2,_mitochondrial
gatccggagccctccgctcccaccctggctctcccgagcccactgtcgccctggcaggtactttttagggacttgttgg
ccccggagccagccttgagcccggtccttccggctgctccgcccncttcgcacgtggcaggcaagcaggctgcgg
agaactttgctgcgtcttctaggaaattgcaatcgatttgtaacgcactttcgctaattgggagttcagaattaaactttaat
aaagcatttgtgttgtctcctctaattgtgaaggctcccattcgtcctggtagatagactctaagcctggagttaatgggc
gctaagtagatc
>embl|AZ755513|AZ755513
ADAM_metallopeptidase_with_thrombospondin_type_1_motif,_2...
gatcactcacattgcctcgtcaaagttggaagttaacagtcgtgagagggttaatttgcagttacacaccaattatttaca
taacagtaattagtgccacgggaagtaggaggaggcttggctgcatggtccagcaggaacttgaattattaacagatg
aaggtggctgggggccaggaggggggctggtgggactgtcagttgggagtgctacattggtcacggctaaggacccca
ggtgtgtcaccctctggtgtaaaggccgcgctgtgggaaccccttgtgggaagtctccacaaaacagtcagcgccaa
ctggccaggtgggcatcagagggggggacctttgtcagagtgcctgctaagtcccttcacgcacagaaggctcagg
ggggaagcaaggccatctggatttagaaacctacccttgggcaaaaggtgggcaaaatgggt
>embl|AZ755325|AZ755325 early_B-cell_factor_2
gatccactcaggtcctcttgtttacatggccagccctctaccaatggcgccatctccccagcctcacatgctctgaatttc
atctagaggtatctgggttttaagagctccgaagatacctcacatgttgctaaggctcaaacagagtcatttgtcctgg
gaggttttgttgatattgttgctggattattttttcttctgcggtcagttttattttcttgacccacccctnacaaaacagaggtc
actgcggcacgacctttccctagtggaatggaactgctcaaggagtccacatggccaagtttgtatccggggggtcc
acatttgtttcctctactcttttaatcagaatgctcatttaattctatcccagca
>embl|AZ755331|AZ755331 atrophin-1_like_protein_isoform_a
gatctntattaccattaagggaaaatgtaataatgaattagtcacactgagaagcttcaagacctaggtgtcttgtctctg
agatgagcagttgcagggttgcagcagggttccgatctgagggcctgggctggnggcaggggattctgctgttagtc
ctgttcctgctgcctgcctgtccaccgttgcagggcttttggtcgtaggtatgatgtgcacttgcaggcacctgaaacct
gttcctaacgctgttagattttcctttcagagtaagagggaccatctactcatgaacgtcaaatggtactaccgtcagtct
gaagtcccagattctgtctatcagcatttggttcaggatcggcataatgaaaacggtactgtggcccttgtgatgatgta
cttgtaactcatagcttgcaaagttcttttaaaaacatcagttaagtgga

>embl|AZ755333|AZ755333 cadherin_22_precursor
gatcagaggggaaatatttatacacgctgcaggctctacgagcccctctgaggatgtcgccctcagccagagcagcc
gcctgtcacgctggattacctcaaaggcttgggcccacctagtatgctggctcctctctcccgaacctctgagtcatcc
atccttcccatgcaggcgcgccttccactgacattatctttaacgttttgtaatcatgaaatattaatgaccaatattgaattt
aaagtacaatttgtgtctggatc
>embl|AZ755342|AZ755342 nel-like_1_isoform_1_precursor|nel-
like_1_isoform_2_precursor
gatccagccattgtccacactgcgccgctcagcctcgaaataccagagctgtccaggcaagtttttctgatgtcatagg
caatattatcagccaggcaggggtcactgacacagcgggggacgacactctccttcaaagatggctgtgtattcacag
ctcaaactagggcaagctagaggccagcagtctgcctctccttcctgtggagagaaagatgaaataagttgtcctcat
ccttacacagttttatattcccagtgtcccttgggggaaggaaagtgttacaaatctcattattttcttactgaatgcttcca
gccggaaccctaatgtgtatgatctcccatggacacagggcaacaagaggaaaacagg
>embl|AZ755349|AZ755349 SEC31_homolog_A
gatctgacgtcctcacacacacatacatgcaaagcaccaattcacataaaaacaccaatgcacaaaaataaaaatagt
gatttttaaaagaagactgtagctagaatctatgttccaatagagaaagattaattttaaagattaaggaaatgagcctgg
gtaatgttcctgacagagttctgatggattagtatctgtcaccttcctccaatggttttaattataattatgtgcttaatatcca
tgtggaaaattcttattcaggtagaagtaagatattccaagggctgacaggctgagaagcacccctcctcctacccctc
ctgtgttagggttagagcagagaatcttagcatcttttccacaggcttaacaacaactcaggattc
>embl|AZ755356|AZ755356 akt_substrate_AS250
gatccgtgtataaattattcagaaattcggtaaataaacatgcaaacttctagccgggactatcgctactgcattattcag
cccaataaggagaatttaatttaatggatattcagcaacattgttcggctatcaacaaaaggcatactgtcttttatgaata
actgattttctcggggcattctgaggagagtggggtcctcgttaaccctagaatggctagctaaatgtaattgaagcgcg
acc
>embl|AZ755411|AZ755411 82875bp_at_5'_side_BarH-
like_homeobox_1|26465bp_at_3'_side_protein_tyrosine_phosphatase_domain_co
ntaining_1_protein_...|
27320bp_at_5'_side_aryl_hydrocarbon_receptor_nuclear_translocator_isoform_2|
23574bp_at_3'_side_SET_domain,_bifurcated_1_isoform_1
gatcattttgttccagggatgtacttttgggattccccaaatactccatcagggtatcctctccccaggtgatgcctttgt
tcttgttggcatctgtgtaagagaatccagcagcctggcctgtcttccgcccgaacagaccgtggagatttggtccagt
cttatgcttgcctcccttttccacagtgtggcactgggcatacttctgaacaaaaatcttcttgccttttttcaacatcactca
tttttaattcgttccgggctggtcaacaccgcgaacgttcggactcgaagacagacgtcccgcgtggtaaacaacatct
taactattcacatcgaatgaaatgtttctattgcactgcaaattaaatacttgaaattaggttgtta
>embl|AZ755199|AZ755199
814928bp_at_5'_side_antigen_identified_by_monoclonal_antibody_Ki-
67_isoform_1|526463bp_at_3'_side_O-6-methylguanine-DNA_methyltransferase
gatcttcatgtttggagaactgcttgacatattctgaggactttggaagacagggaggtgtccctggggctagtggttg
catcagggactgcctctctctgggaacccaggcaactcctactgtaacgtacagcacacattctcctgatgcctggttc
tgtgtctttaaggcaataattgctgaattgaaatttgagactgttatcttaataaccaattactcaaataactttgaaattggc
ttaaattggctcgtcatgctacatttacctggatc
>embl|AZ755208|AZ755208
79419bp_at_5'_side_echinoderm_microtubule_associated_protein_like_1_isofor
m_a|45676bp_at_3'_side_Enah/Vasp-like
gatcaaaccaattatccatgctgcacttccatgcagcagacatccctgaccctcaccacctcgataaaccctaccgaat
gcagacaagaacccctttttgtgtatagagccagtcagagcctaagcagtcattacacagcaaaaacgtagcgagaa

137

caaaaacccagcattagactgagggaaaacccagaacatacctgaattcttcaaatgcgaacatgacctagtcagag
gcaatttccttggattgtatggcagcctttttggtaccgaggacaaccatttagtcagtcagatgctacaaagtgcagaa
aaaaagaggtgaggtttccacacccccaatattctgaagta

>embl|AZ755213|AZ755213 basonuclin_2
gatcaggcttgtttagtgatgacactgttcctgtaataatttgacaagctactgtgaatttgccaggttattaaacaacctg
cagccttcacttacggagtcagagggacggggtccacagcagcagccactgccagttaaatgcagcacgagggtg
aaccagaagccaccccttgaagtctatctatcaccacaagcagatttatatgcctgtgactttggagaacttaattaatgt
ggcatggagatatcttgaaaacttttgtttattgtatgatgaacctactgggacaaagctgatc

>embl|AZ755234|AZ755234 deleted_in_colorectal_carcinoma
gatcctagtcagagtgaggaagatgaaatacctcagcagtcctaaatgcatggtaagaatccacacctctcgccttcg
gtggcccctcttatttaaaccgttgtgctgaggatacctgcctctcctctgaggatgccccacaaccatctctatttctgg
ctccatctgttcttgcctatttccagctgctgttcagaggaaaggccatgga

>embl|AZ755239|AZ755239
37966bp_at_5'_side_transmembrane_9_superfamily_member_2|171782bp_at_3'_
side_citrate_lyase_beta_like
gatctcagagccaagttctatctgtgccactgcttaaaaaccattccacagaacttctaataggcttctttctaagtgcgc
cattctgaagtttgaaaccagtgacaccggcccgctcagcctttgagactcaacttttggtgagaaaaccctacatccgg
ccagacagcaaaaggaaggaaagaatggattgtgataaagagccttTggggttcctaaccagagaataaagggctt
cattttcctaatgggtctgcccttTcccactggaaagttatttaaatgttaccgcaaacagtgaggttgatcagatattttca
gcgggcgaatcctaggaggtgttctgtagaaccaatacaaagctagatagtataagacaaacaatgaaggttaattga
gttttcacgataattacaagggccatgtggtgaaaatataactgcacttctatctaggacaaatgtatttgtgaaacaatg
g

>embl|AZ755103|AZ755103 n/a
gatcaaaagcaccgcgtncaacagcataacgcccgcttcaggctttcccgtgccgggagggcgcccatcgcagct
ctgtttagacagcggtggctgctacaaagtgcaacccaaacattaaacactggtgagcgctccacgctcaattaagaa
acactgctgtcagctacttcctgagcacaggggctattaattcagcactggagaacatgctctggtgagtcctctgtgg
gattatggctgcacatatgtcacagagcactgtctcaattgtcaccattagtcaaatgaaacctcaccaaatacctgatg
caacaaatgaagctacactggtttgcagaccctcagaaagctcagaggcaattacactctaccttctaaaaacaaacct
tgacaggggaaatgtggaaatgttcaaagaaactcaggattgacatatcagtcattgaagactccatcaactctgtgtt
gagagcatcaatatcttgttggattaactttgtctggaaaaaaaaaattaaatggtaaattttttccccctcactccaagaat
tagttcaaaaatggctaatttttttctttatgtttggctcaagaaataaatattgttacccgttt

>embl|AZ755122|AZ755122
252454bp_at_5'_side_hypothetical_protein_LOC285237|698504bp_at_3'_side_ep
hrin_receptor_EphA3_isoform_a_precursor
gatctgaataggtacactgtctaagcccatctgaagtctaggcatcttatttgaatatctagcaaatttctgcaagtggatg
cttccaggtttctcagaagggaggggagttgagtgggaggacagagacggaggggagatgaacacaagcagaga
ctaagatggagcagggaactgctgaagcatggctaatcagacaggacacttcatgcatattatggacccttTgtgataa
taaaattggaatcaaattgcctatccaagccactcgaggcatcctgcatctttaattttTaaagttatgaagacacttgaca
cagaaggggatc

>embl|AZ755159|AZ755159 phosphatidylinositol-3,4,5-trisphosphate-
dependent_Rac_ex...
gatccaacacagcaggacaaagagccaaatgacgagcagagggangacccagccccccaagacactgaagccc
attcacattggaggctggccatttccccagggagctgggaacaatattccctttTaccacccttcttgtctcctacttccg
catttaattgcaagttcatgggtattaagtagctcacacagctgacctagctccctctcccagccccatcccccaaaaa
gagtggcacagcgtggcctctgttcccagcccctctgtggtcttcccagggatttctgcctagagatgagaaaacagc

138

aatcccaaaaacccacgtaggtccccagactgttttctcacctagccagccaggctgctcctgtcctccccacccctg
ctcaaaggtgactcaatggcagcctctgctctgacctcccatgcctttctgttgccccactgactcccagggaaaacta
catcctgcctggcagcagaacccaggctcagagnatcagcatctnttgcctgcctggtttgcttccct
>embl|AZ755183|AZ755183
422038bp_at_5'_side_ring_finger_and_WD_repeat_domain_2_isoform_d24|1164
71bp_at_3'_side_tenascin_R
gatccagatatagaactctcaactacctctccaataccatctgcccatgtgttccctcccttcagccgcttccctttccca
gagagatgcaaactgtgtggcttgcagaacatattttaaaaattatttccctgaaacctggagagctgtgaagttaggc
attccgggataataaaagtagtggggaaaataaaaatttttatcagcttgtttacacaattatgtaaatgatgagcacttat
acatggctgattcctcaattattcatggtttgaagctcttgaaaattcaattcaactaaataacctttctctcagtgctcacg
gtgagcaaggtactgtaatttgcaaaggggggagagagcgatacaatttgatc
>embl|AZ754981|AZ754981
20906bp_at_5'_side_phosphatidylinositol_transfer_protein,_membrane-
associated_2|101123bp_at_3'_side_M-phase_phosphoprotein_9
gatcactataaaccgaggaactatattgaagggttgcaggattaggagactgaaaaccactgctctagggcctggaa
cctcctcaaagcagtcactggtccaggggtcctcaaacactgccagctttgcccccctccntcccccattgtacatg
gcactttgtagactgggtgctaaacagatacaatacttggaactaaaataacacccggcttgatgacatcttgatgcag
ctttgtgggaacagcaggctcctgtgtacacatcgttccactctgcatgtctatttgtaaattaatcgcacaccagtagac
ttggctcagatgcccatgtccctggaccactttgagggctagaataggggctataggagcc
>embl|AZ755034|AZ755034
129844bp_at_5'_side_fragile_X_mental_retardation_2|361293bp_at_3'_side_idur
onate-2-sulfatase_isoform_a_precursor
gatcactactcagactcaagtcactgggaggccctgctgtcttctgcacagaagttggtgttaatccgtaaaaaattgat
ttatgtcactgccagttatgtgctcatgtggagacgggagcaggcccnatggagatgnagggctgctcntcatgtgat
nttcctctgtgaacttagcacagaattgtcctcatctcctcatctccctttggagaaagacagagtctaggtgtgttcagg
aatagtgtttgcttaatgatacacttaatgaaagtctttgatccgac
>embl|AZ755069|AZ755069
kelch_repeat_and_BTB_(POZ)_domain_containing_9| kelch-like_29_(Drosophila)
gatctnggctnanntaccagcacaagatgagatgtggaactcattttgcttagagtagtgtttaggcacaagcaaagg
aaacttctcatgtctgcatatttaagcacttttattaagtggaacaagctatttctgaacacgaaagtgccaacagccttag
aggaaattaaaatgaaccatgacaatggaaaattgcacaggggcacggcggattctacagagataaatacaacaac
cctttcaacagcattcgggaaagactagagatc
>embl|AZ755070|AZ755070
4398bp_at_5'_side_hypermethylated_in_cancer_2|15881bp_at_3'_side_hypotheti
cal_protein
gatcataanatancantttctcatcactgtacattaagactgcacgcttttgaatggagagattagtcatggataaatcac
tctaacatcggagggctgcagantgcactctgacctgaatgcatccatccaaactctttagggaagtaaactcaaatca
atgtgatgcgtggtaataggggtgaccagtacaggagtgcgatcccagcatccctcccctgctcggggaagggcat
aaaggacgtcatgatccctcttccagttccaattaaatgaaacagataccccatccctcctgaggtatttgcaaatggga
tggata
>embl|AZ754883|AZ754883 32000bp_at_5'_side_mitogen-
inducible_gene_6_protein|277124bp_at_3'_side_DNB5
gatcaacctttgatagcacttgattctgatcgtgtctaatcacgacgatccagatataattgaacaagccctaggcacag
aagtttatgactcaaacctcagtccttccaggaacccaaccagaagtttctccatgtgtctgtagatgctgcacctcagg
atgtctgtgaacactttactgtgtcccactgtgcccttgtgtttgtttcccagccatggaaaatgcccatgagtggacgttg

cgcaatcgccagctgttagtcacttctgtttcctgtaggcccgcacttaatttccctggtagctggcactgactnctcagg
atttaagttgctgaggggtattcgggaaaacccaca

>embl|AZ754932|AZ754932 95406bp_at_5'_side_similar_to_high-
mobility_group_nucleosomal_binding_domain_2|83952bp_at_3'_side_high-
mobility_group_20A

gatcacagacttctcatcagtgtctcagtcaagnagactagtatcactaaagaagaacataagtaattaattcaggagc
aggaaaaggtaataatgagaagagtcaattaattataaacagcatgtctaattacagattttgctaggggctccagatac
ttttaattagttttacctggtatattaattataaaatagaacatctcaagagaatgttctaatcacagaagttaactgagtatat
aaaatcagaatttaagttatctttataatttaataattaaatgttcaataatataggataattactttgaaagtcattgattaaat
atgtactacatattgttcaagttaggacattactaaagaaagaacagagtcacttaaaaagcctaaggaaacaagtgtg
caactgaaactgaatatgtagcttagtaagaaatttgtat

>embl|AZ754759|AZ754759 calmodulin-binding_transcription_activator_1

gatctattaacactttggggttgattcatatctgaagctgtgtgagaaagaggagtctcgcccttggtattaattgcgtcttt
gatggtggttccagagcggtgtgggggggaggcagctgtcaggaccgttccttggcgctctctgagagaagcccgtc
ctctgtcagcgacaatctgttagaactagctggagtgatgtgtgcagaaaccccgggtttcagctcaagagcgccacc
tgtctgctgcacagcgcaccaactcatgctgcaaccaccgtaagtctcctcggaatgggcacgtggctccagaagttc
tcagaatcctctgggattctgagtcacagatgtctcccagtttttggaccagt

>embl|AZ754776|AZ754776 engrailed_homeobox_2

gatctgaaaccctaaacagtcccctttgcagaaaaataataataacataaaattagcataacctctgaaacaccaaaaa
actcaaaaatattttatagtcagattttggattttacaccacctgtaaattatatgtacatagattattgtagaaccatagcta
atacactggattttcacagataatgctggtataatctttatacactggccttttgtacacggttttaaaattattgcattgtttc
gcgcggccctagacatggccggccgcagncccacccgccctgcttcgccctactcgctgtccgacttgccctccttg
gccgtggtggaatggttgtacaggccctgtggccatgaggtgcaccgccaaagtgttcttgttgncccgtggctttctt
gattttgggcc

>embl|AZ754810|AZ754810 mitogen-
activated_protein_kinase_kinase_5_isoform_A| mitogen-
activated_protein_kinase_kinase_5_isoform_B

gatcctacttccggagccaaaactgtcatttagggttnctggtgcatctgtcggcgccaggcttttcactggcatggga
gagtgtgaataacaaattggatgcacctgagattcggatgattagtgtgcgctgtgaatctctgagacttgtaagcggc
gactctggggggaggggcagagggccggagctgagggtgggtgggtctgtgcccacaccttgcaatcggggctt
ggcaagtgctgtgccgacgtcgtcnaggccctggtgcagagatagtaacatatattttcatcgtgccttgtcatgattag
aaaagagccttaattttcatttcaccccatgttgtgattaatttngcattctttgttcagaagttaatggagtttt

>embl|AZ754811|AZ754811
24084bp_at_5'_side_far_upstream_element_(FUSE)_binding_protein_3|3191bp_
at_3'_side_PR_domain_containing_12

gatcgttcccatccgattaccaccctatcctgccctgtcctctacaccccactcgggtcccttcgctggtgggtctgtcg
gccggccgctgcttcccttgtccccactcccgtgggctgggggatggggcacgggcggcggagatgacctccgaa
gagccgcagccgattgattcgttccgcccagagtccgggccgagtgaatggggctcgggcggcggcggcggcgg
canagcagcagggataaagatgttgtccctgaaaaggggggtggggtgggtgaggagcgggaagaggggccatt
cactcctggcccggccccttcgggaaagcctctggcgaagaaagagggccagggttctattcaagccctacgagcc
gcccacaatgggccgatatgctgggggcttctctattaacgcccatgaatattaaag

>embl|AZ754813|AZ754813
1383bp_at_5'_side_hypothetical_protein_LOC283951|3390bp_at_3'_side_coiled-
coil_domain_containing_154

gatctagtgagaggttgcatagcccagtccccccctttctccaggtcctgttgctgcagagatggtttatagcatgcact
gcaggcaggatgtcctctggagaggtgtactccagaactcaaaatagcatgaaatttaattttctgcctatttgacacag
cttggatggggcccccaatcagcagtatagataagggtcctcgggtccagggtgtgggccagctgggcctgaggcaa
acactccatgttctcccggcaggctgcgtgcaggtgaggcagtctttgggtactcccccgcagctctggttgctggg
cgctaaaacaaacgggctctgtttgcccagtctgcgatgacaatgccgttggccaaatatttggcc
>embl|AZ754844|AZ754844
224567bp_at_5'_side_crystallin,_beta_A4|895657bp_at_3'_side_meningioma__1
gatcacctcccatttggagggatttgtactagatgcaaatgggacaagtttctaaacttcaacgtaggttttcatctgcgc
aggtntaactgccgcagacggcccccggttttaaaaccaggagcagaaggtacagtaattagttatgtaaataacgtg
gcattcttgcttataattagttcaagaataaaataaaattcacgcatgtcacaggactaattttcgtatgcaaatgaaggat
c
>embl|AZ754849|AZ754849
531323bp_at_5'_side_solute_carrier_family_16,_member_7|107545bp_at_3'_side
_hypothetical_protein_XP_002343217|107588bp_at_5'_side_similar_to_XAGE-
3_protein
gatcttataattttacatcttgagaaagcataaactcttctgttgagatattaaatttcatttaaatatgaaaacacacccctttt
atctgctattcatttatgatttttccaatagttcaaatctgggcagaaaaagtgtaagatacaacacagtttagcctttctgc
ccgttttatgtcaatccaagtcaagtgcaaagaaccctgtttaggaaatgcactggcaccaacccttactcatcttgtga
caccacagaaaagcaataccatatctaattaaattagctaaatgctgctgtctggatttatctggttgttgtaatttaatcag
agctaaatgagatc
>embl|AZ754642|AZ754642 82875bp_at_5'_side_BarH-
like_homeobox_1|26465bp_at_3'_side_protein_tyrosine_phosphatase_domain_co
ntaining_1_protein_...|
27320bp_at_5'_side_aryl_hydrocarbon_receptor_nuclear_translocator_isoform_2|
23574bp_at_3'_side_SET_domain,_bifurcated_1_isoform_1
gatcattttgttccagggatgtacttttgggattctccaaatactccatcagggtatcctctccccaggtgatgcctttgtt
cttcttggcatctgtgtaagagaatccagcagcctggcctgtcttctacccgaacagaccgtggagatttggtccagtct
tatgcttgcctcccttttccacagtgtggcactgggcacacttctgaacaaaaatcttcttgcctttttcaacatcacccatt
tttaattcgttccgggctggtcaacaccgcgaacgttcggactcgaagacagacgtcccgcgctctagcccaaggac
gcgccggcccgcgcctatcagttcagacattcttgtatgtcttccctatctttatgggctagttggttccgggatgaacttt
acagccttgtcctttggctctaagcccaaaaaa
>embl|AZ754697|AZ754697
7866bp_at_5'_side_hypothetical_protein_XP_002343435|101329bp_at_3'_side_R
h_family,_C_glycoprotein
gatctcaaagatgtcaacacctttcantttncaactagaaacgcgganatgtcgcagaaggtcacacggtaggccact
tggcaggtcagggttgagaacagccctcgttgagaaggaggcttcgaaggcttcttcctccgttgtccaaagaatttca
acactgaactataccggccatatgggctcaacccgaacaatgaaagctggcatccgcagtcctcacggaaacgcac
ggtgcagcgaggccagaagctggtctagactggttcgggagttctcgaccgctgtcggacctcccgggaaatggga
acagagcacgaagcagccgggccaggaggcttcccaggacccagtgcgacgctgaaaagcatgaagggtcgga
a
>embl|AZ754700|AZ754700 PR_domain_containing_12
gatcgctgttcctctcttaatttaccgtgaagactaattccacttccattcacgctatgtcaaccatctaatcctgccttttgt
aaggggaattcctcggcccctttaaacaagtcccctccgcattgagatacaatttactgctacagctttcttccagggct
aatgaatttagaattagcaatttctttcgaatggaaccaaatgaatgcgatcactttaacagcgtgacaaattgccggcc
gctccgcaatggacaccgtttaacccccccctttcagccggccggtccgcggtaattttcccaggtagcttagagggga

accttgttagacatggaggccgcccctgtgtgcctcaacgctcccacccctggccgcagaggcccaccgggattc
cga
>embl|AZ754721|AZ754721
forkhead_box_P4_isoform_1|forkhead_box_P4_isoform_2
gatccacactgactttgagagttcttttcagccagccaacttcattcatgctactggggaaagagtagcctgtggtttctt
cccatctattttgcaccttctccgggtcatctttctctacatcacttgggttacttgatccccacccccacccccttctggtcc
ccatccatagaccatgccggctgctttagacctagggtatccaggcttgggggtgggcgtgtgccaagccagctgcc
atttcttagggacagagctaccatctgccaagtggatggtcctcagcagccttcctgtgtccatcccctcctagggcct
gtgttgtgggcccatccttttgtttacgcagcccccgcccttaattgccgttaatgtttcagcgtaacgaattagttctnctc
gttcacgaatcaggcttcgaaacgagggaaaaaaagc
>embl|AZ754728|AZ754728 717749bp_at_5'_side_CWC22_spliceosome-
associated_protein_homolog|273766bp_at_3'_side_ubiquitin-
conjugating_enzyme_E2E_3
gatctgtgatacatttggacacatcacaggtttcccatggttgagtcactttcaatggggcatgggcacctactgaatgt
cttctgtgccagtgtgtggctaattggccatatgcccagattgctacagccctctttgggtcagtaatagtcctggttcttg
ggtgaagaactgctcacctctcagaggtgaattaactctttctgaacactagggcaccattaagcacactagaaaaaa
agtaatttactcaatggtcctctaaggacaccctggagaagagaacagaagctgtgaaatgctaagctcgcatggaat
tactggcagaatcaattagggaaacaagacaaattagatgggcatgtggctctgttccaattgagacagaaagaatctt
tatggggcaggatgtaattggcctgatggaaatgtgggctaagaaaatttgtgttttctgagaaggacttgaaaggaaa
ga
>embl|AZ754534|AZ754534 stanniocalcin_1_precursor
gatctttggggggagaaaccggtcctgggttttcttgcttacctgaattttgagccgccacccgggattttctggggctc
acagaatcattttgttccgcctcgtgcgctgcagctgcactgatgaccagcgccagaatcactgctgagttttggagca
ttctctgagaagtttccgctaagttgttggctttttttttcctcccccccctttcctctttccctctcccggctagagtgaagatg
tggatc
>embl|AZ754560|AZ754560
307832bp_at_5'_side_akt_substrate_AS250|105361bp_at_3'_side_non-
protein_coding_RNA_153
gatccttgaatgtaatgtatgcacaaactaatttatgatactgtgactaatttccatttcttgtgcttagagttgcctagctaat
tacagagatgttttcttaactttagtaattacgttctatgctttcgacaggcaacgaagcaaatgagcaaggccagattgc
acagtacttgggagacaagcgcgtccaggcgagaggagaatgtgcgctgagcttgcagggccggccagcatcctt
gatcagaagcctcaccagctttgaccagatcccacaagctgcatcggaagcatctggagacatatttgatggaacac
cccataattagaccggctggcttaattaggagaagtttgtgggtagagggttgctgaaggggnctgtgctgggattct
gaaggtttgctngcctcctcccaggaggcctggggcacaaaagctctaacccgctgaagtgtgcaccggttggtg
>embl|AZ754601|AZ754601 392446bp_at_5'_side_BCL-
6_interacting_corepressor_isoform_b|110303bp_at_3'_side_ATPase,_H+_transpo
rting,_lysosomal_accessory_protein_2
gatccacacgccccccctcaaccccctcctcacacactttaattacccattgctggcatgcgctgcataaattaagagct
cgttaatgctgtaatgcgtaaaacaatttcatgcatcttaaatatcgctcttgtactttaattccacattaagtgggcccctg
aaaattttaaatttacaaagtcactgccatgctacatgataaaggtgctaattaagtgttaatcagtagccagataaccttc
ctcgaaccctacccaaggtctccaaaatgacccatagccagagaaggcaacgtgacggccacagaatagagcaca
gcccagacaccacatcccagccaattaaggcagatttaggctattgtacagttctaattaagtgacacccactatggtg
gacacgagg

>embl|AZ754621|AZ754621
107982bp_at_5'_side_neuropilin_2_isoform_6_precursor|152581bp_at_3'_side_I
NO80_complex_subunit_D
gatctgtaagcctcctggccagaccgtgaaccgtgaaaaattcattaagataatgccactttggaaggcctcttggtgg
acggccagccagccatttcctgacaatattgcatggctttgggcccacttttggaatcctgatggtacctaactggactg
agggcaagaaagaagatggtagagagaacaggaagtaaacgggaaaagcattaaaagtggttttcagttagtccac
aactaaaaactgtggtgtgctgaaacatgattaaatctcagtgcttttaaagtggggtgggtacccttgtaggggagcat
ggaaatgtgggatgaggggagatgtgtagtaaagctctgaaagattaatttatcccacaaacttggggatc
>embl|AZ754462|AZ754462
209790bp_at_5'_side_natural_cytotoxicity_triggering_receptor_2|4599bp_at_3'_s
ide_forkhead_box_P4_isoform_1
gatcgcggaggtgttaattggaaaaccctggtattgtgcctgtttgggggacgagaacgtcaataaaaattaattgatg
agttggcagggcgggctgcgcgggttcgcggcgaggcgcagggtgtcgtggcaaatgttacggctcagattaagc
gactgttaattaaaacgcaatggtaattaataccccccacgccatatgggcccgcaaaaaggcacaaaaaggtttgtc
cacgtgtggggctctggtgtcttttccttntccgcttccntaaagcaccccaggctcggagtcccttttggcccttagcta
ggtgcaactcgtcacttccggcagggaggggacagggcggaccctggctggttcagagaaggcgttgctgccttc
gttcccgnctggagtgcgtcatagaggcaaaagcctcagggaaagtgggagcccccagatggccgcagaattccc
caaagcggaagaagcggctgtagctttgacctagcatcttcccgacgatgcctgctggagccccggcctggccatg
gttccttaagtggctggactttaccaagggaatctgaagaatgggagactggggggccttcaaatcctggatc
>embl|AZ754471|AZ754471 3640bp_at_5'_side_glycerol-3-
phosphate_dehydrogenase_2,_mitochondrial|671357bp_at_3'_side_UDP-N-
acetyl-alpha-D-galactosaminepolypeptide_N-acetylga...
gatcatcntacatgtcaagttagtgcagggttgccangcgttgataaatgctctgacanaaaaaaaactgctatatatttt
cattgtcggtaaatcantaaagctgttaccagtttaacactggncaaattttanaatacccatgttattttataantttaattt
gcacatgtacagaganagcnttctntatttgcngntcccaagtgctnccgccaggttgagcgtgctcagaggcgtttg
tgagtttctcatagtccttcggaactgtgtcccaagtgaagccttcttcccagacagacttgggacaataagtcatttgttt
agcttctggaaagaaaaatacatagtttgaaacattttccatagatc
>embl|AZ754480|AZ754480
31809bp_at_5'_side_serine/threonine_kinase_3|91426bp_at_3'_side_odd-
skipped_related_2_isoform_a
gatccagcanttaacgtatcttgaacaccangtcagctcttgctggtggtggttgctgaggccanccttataantttaa
nggacnttgatgcattacttcantaagaacttcacatgctacaaactcttaatgcgccttaagctttntagcacacctgtc
caaaacaaccagcaaacattttaccctcctcgatagaaacttaacactnttaatttgctaataaggcaaataaatcttgctt
tgagagtcgtgacaggttctgtacttgtatttggggtcatagctggaaattctgagtccccaagtaatgatatttttaagttt
gcctctttgaccttcattaattgcaaaataaacagatc
>embl|AZ754300|AZ754300
15813bp_at_5'_side_NADPH_oxidase_3|1306279bp_at_3'_side_AT_rich_interac
tive_domain_1B_(SWI1-like)_isoform_2
gatcattgggaatgtcagagacagtaattttcaagcttggctgattttgaaatcatacggttatgagtattaacagcttaat
ataaaatgtggtgctgtcctcaaagttaaaagaaattatcnttgccataataagctatattaatttaattagttatgtatagg
gttatattttcatttgaagatggntacagagtggtagctgttaaatcaatgctgggactagcaaanctgcccgagagttt
gtnatcggatcccagtgaccatgttgaggntcggccacggaggcaacagagtgcagaaggagggccacagttacc
cacttaacggttgccccaaattacttttgcttttacaagacaatgtgaaacttgaggtgttaatttaaaaaaaaaactttatc
tttgtatcctatcctgctccacactcatcctgcatattgggcatatgtagcctctcttggggaagtctgctttgacatctaca
acttatgctccccagtactgtcaactgttgggacagccagccttgcctacaggtgctgggccagctgatc

143

>embl|AZ754307|AZ754307

18749bp_at_5'_side_transforming_growth_factor,_beta_2_isoform_2|713671bp_a
t_3'_side_lysophospholipase-like_1

gatctctgcctagtcacagcagggtgcacacatanctgtccgatgactcactaccaaatgcttcctaaaaganactcc
ctaagttaccgtaacatttcanaacggtatttatagtgccgatagaaaagaggtggagccagggacttggagcctctc
cgcacccttgctttcctgccagctagatccagctcancgggaggcccttaagcttgacactattactgagtttatggagt
gctcacattaaaggacctatcatgattgacctntgaaaggcntaataagcagcagaaagagtgagatgcagatatttg
cacaaaaccaaaagacagaagctgctgacccctgttgttgagttagggaaaagctgtaagaagctgagaaggaaaa
gaaagaaggagaaggaggaggaggaggaagaggaggaggaggaaggaaaccatgtaggaggaagagtagtct
caattaatctggatc
>embl|AZ754331|AZ754331

162974bp_at_5'_side_spermatogenesis_associated_8|1012617bp_at_3'_side_arres
tin_domain_containing_4

gatcaaggtggtgcccgactggctgatgacttcatacaccgtgtcccggctcactgacccacaaatctcctttgatatg
ccaggcaggcgaaagagcttgaaaggatcaggaggagaagggccgcgatgaagccatcatgtgacaagaatccc
aagcacganctgacgagctggaggagagcgggagcttggcagctgatacagacatcgtcagagccacaggagtg
acctgcgcaggcaggcggtggggaccaccgaagacgcgtccggcacgggccaagcanagccggcaagcaatc
ttcgctaaattaaacaaatctccatgggnattagcatgaaacaatttctaatctcctgcntttcactgtcnacaanggtc
aatgggttaccttcnctcntaatttantaatccccanccacntccngggatccttanncggtcactgggcctggggct
aaggcagcatgtcanactgccgcantcttgggctcttaanc
>embl|AZ754402|AZ754402

33355bp_at_5'_side_nucleolar_protein_with_MIF4G_domain_1|2354bp_at_3'_si
de_homeo_box_HB9

gatctaattggcttaattctaggaacccaaatcttaagactgtaaaaccaaccgttttcttagccgcccctcccccgga
attcccaccatactggcctggggtcncanaggctagagtgtgaacnaagganaggccnaggcccaaggtagatcc
tgntgatggcagagccaagttggtgacccagagtgcaggagcctggagtggggcgcgcggcctccgtgcgcgcta
cgaggactntgggcagttcgttttgctgaatgaatgtgcgccgtgcccggagtaagctaggatgagccagtgattata
ttaaaaattgctataattgatttggggaagagctactgtgacaagaacgaattgatctgcctctgtgacttcaaggccag
ccaaagtgaattccaggacagtcagggctatatagagagatc
>embl|AZ754409|AZ754409 syndecan_3

gatctggaatggccanaggtgtggctacaaagtcaccaggctgctcacagagggcccagtgcctcgccagtgggt
gggtgtcttgggcatctgccccagctggcantgggggaaccttggaangcgtacgancagangagaaactagactg
gtgcagcactagatttacattcagcatctcacacccctgccattcacagacccttctanaattcatgcananccaaaga
gaggtcttatcctggttcctgcccctanccccagcacctgctcnaaagcccccagccaagccagccagggcaggga
gccaagagataccctcatacaacacacacactcacacacacacacacacgcacgcacgtctggaagtaatgaa
ttaatcagtatttactttatacccagtgagtcacggggtggaacaattagtgtaattgcccccatgactaatcactgggtct
tccccctttgtagacatcccctccaactcccatctctgcccagcccacacttcccaggtcctgggctgcaagagtcac
atcttctataagctagaatcctgtggggactttgggcatctgtccctcttgttgtccagccagtagcccgatgtcacacag
caggatggggactaagcagggacctctatcttcctcttctgcttcaaaggaacattntaagagatc
>embl|AZ754280|AZ754280

538140bp_at_5'_side_hypothetical_protein|179549bp_at_3'_side_odz,_odd_Oz/te
n-m_homolog_3

gatcatttataatccttgactcctccgctcagccacttgtcaagatgacaggtgctttatttcttcttcttctttattcaaaaaa
tcagattcccggttttctgccttttttcgctgtcaaatgggatttgccggactgaaccacgtccacaccgccgcccctccc
tcccatccctcctcccagccacgccggtacggggaaagcaccgaggctcccatgctgtgaccaatcagtgtatctac

tactctcgccgtgatttangggtcatccatgaataacacttcatgctgatgaatggctcttgccggctttcggctctgatg
acgggtctgccggccgctgattgcaaacactaatgaacttacagcaccatttgaaattcacacgcaccaacaaagtga
gggggtacgtgcgtgggagtctgtcctttcttttctgggtgaccggggaagaagacctggggaaaggggcgggta
attgccttggatttaagtggaaacacttctcggtgtaatcaaacctctaaccttcttagaaaactactttcttgatc
>embl|AZ754350|AZ754350
Rap_guanine_nucleotide_exchange_factor_(GEF)_4_isoform_a|
Rap_guanine_nucleotide_exchange_factor_(GEF)_4_isoform_b
gatctgagtactcttattttgaacactgacacggaacttaaagacattttaatgaaaggaatcccaaccaagactctgaa
agctgaaccgtgtccccgtagcagcggtacggcttggctgatttccacagaagattgtctaaagcctggagcatcatc
tccagtgagcaccagaggctttcccgacagaccgcggtgtagaacaagccacagagcagtcaagaatttttagcaa
ccagcctcagtgtggccgtgaaattaagctgcagacaaggttttctgattgcacacggtgctctgaggaggccattca
ctcttaccagtgagtcgaattgctctctcgggcaggcaaacaggcgtccattaatggtgagcgtcgtaaatactgaaac
atcattagatttcagcaccaccttttctgtgaacataaaaagcatgagattgcctattaaatccatctgaggtttgctgcgc
tgccctaatcacaccaacaaattgccagcttggtgaattaccggaacatgcaatgcgactgcataaaacactcgagcg
aacggatc

AII.3 Human PAX3-CASTing sequences

>embl|AZ756669|AZ756669 Ena-vasodilator_stimulated_phosphoprotein
gatcacatggtttctagctgcctcccctgttagctgccgcttgtcagcagctggcggactcggaggctgcaggcccct
gtctgcagaaagcctttacttcagaatattgggggggtgtggaaaccccgcctcatttttctgcactttgcaccatctgac
taactaaatggttgtcctccataccaaaaggactaccacacaatctaaggaaattgtctctgactaggtcatgtttgcattt
gaagaattcaggtatgttctgggctttctacatcagtctgctgctggctttttgttctcgctaggtttttgctgtgtaatgaca
gcataggctctggctggttctcacaaaaaaaacagggttctttcagcttgtggtggggtttatcccagtcacggggggt
gggaggtgcagaggtggtgcaggtcaggggtgtttactgtgtgcagaagtgcagcgtggataattggtctgatc
>embl|AZ756703|AZ756703 137269bp_at_5'_side_ST3_beta-galactoside_alpha-
2,3-sialyltransferase_1|833063bp_at_3'_side_zinc_finger_protein_406
gatcatcgcctttaagcaatccactctgcgagcctggacttgacaggcctcactgtgccgtgaaatctaatactatttta
tgcctttaatctatcttttaacagcctccttgtttagaactctgcgccctggattcacggaaagcacgttttgagagtggt
gaggacaaacttggcccgtcttaatcaaaggctcggggagacccaatctggggtccttactcaatcctcaactctccta
atgccttcccccactgcttccttttataatggcctcagcaggaaccttcctggctgcttgacagagcctctgactggcttc
ttctatcttaatagccgctgggctctccaaactaggatc
>embl|AZ756509|AZ756509
10657bp_at_5'_side_forkhead_box_B1|224385bp_at_3'_side_BCL2/adenovirus_
E1B_19kDa-interacting_protein_1,_NIP2_iso...
gatctgcgtctttctttgcagtccggcctgcgcgacccgtgcgccctccacaccgcctagctttggggccctgaagag
caaagaggtctgagaaccgggatcctggccggaccgccgccctgtttcacctggtaatttatgcacaattaattgagat
atcttttatacttgatgtgttgcaaaattaatgcctaatttatgtaattagtcaattaactctaattccagcatatgttcca
>embl|AZ755727|AZ755727
7834bp_at_5'_side_transcription_factor_EB|15880bp_at_3'_side_MyoD_family_i
nhibitor
gatcccagaactgggaggggcacaaaggaagtggggccccctccccacccactgtcccctcttccatctgcctcc
cgctcccccaccttccattccagaagcttccagcaggagggcatttcaggcttctctgaagaggggactaattaaacct
taattgacacagcactaattaacttcttaattacacggaaccacatgattagccacgcagcagacactaattgttcattag
gaggcagcccctctgtcccccaccccctttcctaggacctaccagctaaggtacctgaggggcgcagagttggcaa

145

ccccaccacccagcccagtgcctgagcccagcctgctccagagcctcagctggaatagcactcaccgccctggcct
gaagctagctgcccttccataatcgtgacacaaatgatccggcaac
>embl|AZ755850|AZ755850
112325bp_at_5'_side_eukaryotic_translation_elongation_factor_1_epsilon_1|161
304bp_at_3'_side_solute_carrier_family_35,_member_B3
gatcactgatttgagaaatggttcaagtcttaatgggcgctgagaatgaaggctgatcgattacttccttcctgggctgg
tagtgttgatgtttgctctgctaattgtacttaatttactgtgaagttttttttttttttttttgagac
>embl|AZ756636|AZ756636 developing_brain_homeobox_1
gatctaatccccctcccctgagcgaccctaggcatcccacacaagagggcaccccatgtggtcttgtgaaaagccc
gcaagccctgacatgagtcttaaagagaaattagccccggctggtaggtgtctctcggcggcgtagaaggactatgg
ccggtctgtgtgtggctgcccggggggagctgaccaattggtcatggtgctgaaacctttgtgcgggaatcgtggcca
agtgcactgactctgtgggaaaacggcggcgtgaatggatgccaacagctcctcgccgagaagggaaccgcctca
aattgggaccatgacaattcctgctaatttgtagagcagggaattagtgcaaagtgtcaagcagtcactgcttgtgctca
atagggcatcaaattggcgcgacggattcatgaatgttgtacgcctcgcaacctctgaaccagagcataaccc
>embl|AZ756644|AZ756644 171477bp_at_5'_side_hypoxia-
inducible_factor_1,_alpha_subunit_inhibitor|14570bp_at_3'_side_paired_box_gen
e_2
gatctttctcgccgctacccagcctggcgtgacggattaaacatcaactagttaaataaaaattaatatcaactgcctcct
aacttctaatgtcatgttgtttataactaatgaactgattaggggataaacacatcacgaagacattccactctcctgcctc
ggctccgctgccgcctggtacccactccccagcatggcctgcgggttgccaggccagaaggagttgggggaggg
ggttgctggaggcggaacaggctggagggcctgagggctccctggggcttgctcctcttcctgccttgatc
>embl|AZ756398|AZ756398 977518bp_at_5'_side_potassium_inwardly-
rectifying_channel_J3|534847bp_at_3'_side_nuclear_receptor_subfamily_4,_grou
p_A,_member_2
gatctccacatggggctgttagagggataatgtaccatacaatacacggtttactttttaattaatagaaaaaaatgtgaaa
tgaatattaaatgtagcttaatgtgaaaggaagctcaaagcagtttggttgttatttatttattttaaaaatattcagtcccgc
ttcatgattagaaaatagggaatgcattaccggactggaagccttgctggagacttggaaggaaaaaagagcttgga
gaggctttaacaaaccagctaatggtactgggcctgacacaaagaagggatgctttagggagcggtccaaggagct
ctgctctgtgcctattgaatcccagcagagctttctctatattccctgctgagattgtcccctgctctataccatacggatt
gaaggatc
>embl|AZ756440|AZ756440 leucine-rich_repeats_and_immunoglobulin-
like_domains_1
gatcagtgtgtgtgtgaagggagtgtccctcctttgcggctgcacctgaagcatccactgttgtgtttccttcagccatat
cagcagcgacacttcctgtgtgactgccagctgaagtggctgcccccgtggctaattggcaggatgctgcaggccttt
gtgacagccacctgtgcccgcccagaatcactgaagggtcagagcatttcctctgtgccaccagagagtttcgtgtgc
ggtaagactgtctttgtaactttttctctccttcatgaaggacaaggaagccaggaggtactcttcattatgggagtaattc
ctgtgatc
>embl|AZ756471|AZ756471
21774bp_at_5'_side_hypothetical_protein_LOC68115|417655bp_at_3'_side_hypo
thetical_protein_LOC67198
gatccgtaacctcagtttgggtttcctagtgcttggatggatgtcaagatgatttataaagcactctcatcccagtgccat
gggccagtgataccatcagtttagcaactcctgaactgcttacctagagttgcctattgattaaggctgctaccagatgc
agcttggtccacatgctccaaccagcaagccctagtttatgagccagtcatttatgatgagcatcatctcagtcatcattg
cgcgatgtgcttttggcacttttttgattcacatgcatagactgcagtcacgcataattaaatcctgtcacagtcgcatgaa
ttatt

146

>embl|AZ756249|AZ756249

215191bp_at_5'_side_M_phase_phosphoprotein_6|266619bp_at_3'_side_cadheri
n_13

gatcacacttctgaatcaaaattgggaattgcggtttgcccagtattttaatgaccacaagggcagtagggatagggtat
tccgaagtgggaaccttcccattatgtgtgggatggaggaaatccagggcacacaaacctttccattatgtgtgggatg
gaggaaatccagggcgcgcaatcacgacatgcagaagtgaatgctcagagctggtccatctgagctgggccacca
gtaaagggcatgcatttgaggagttaattctgggaccactctgtaccttgtaatcactccttgggctgttataatcaaatg
gaacttaccacaaagttaattcgagattatatttagtgatatgctttatcgttgttgtagtgagtggtaattaggaataagat
gggaaacaattcccataaaagattgtgattagcattgattatgcagacaacccgaggttttccactcgcacctttttgctg
cttctgtctctctaatgag

>embl|AZ756264|AZ756264 brother_of_CDO

gatctttaagacttgttagttccgtaagagaggaactgctttcaagccagaacttctgccttggcaaccccaggctgag
gagaccccaggctggagcctgagagaagccagacgcagcaggggcctgtccggactctcacttcccctgcgccc
catgtaccctgccctcccagcactggcctcacacttccactgggcccagaaatgaaaataagaaatagaagggtaatt
agtcactgtcacacacactgagggggtgaatgcggcctgagcctggcatttccggatc

>embl|AZ756118|AZ756118

761417bp_at_5'_side_carboxypeptidase_4,_cytosolic|664474bp_at_3'_side_neuro
trophic_tyrosine_kinase,_receptor,_type_3_isoform_1

gatctacaaaggcattttaatagtgcatacggtagagggaagagagtgaaaaatgtctttgtttcactctgacgggtcc
gatgaaccatgacacaaatgattgcttgcctcgtttgtaattattcatggtaaagatc

>embl|AZ756146|AZ756146 hypothetical_protein_LOC101694

gatcagggaattagcgctgtaacttcggattagtaacgctcgctaaataagtgcctaattccgacccttcatgcttttcag
cgtcgcggctgagtcctaggaggcctcctggccaggctccctccgcgctctggtcccatttcccgggaggtccggc
cgcggccggcagagaactctccggcctgtggagatccggcctcaggccgcgccgcagcctcctctcggcagcgc
atttcattggcccgggcagaggcccatgcagtcccacggcgttcggagttgggagttcgtggaccagagaaaagga
agccccggatgcctcgcccaccacggtggctgtactgatc

>embl|AZ756158|AZ756158

32062bp_at_5'_side_hypothetical_protein|100348bp_at_3'_side_hypothetical_pro
tein

tctccccgctgaccacgaggcaccggcagcagcccctggggccatggcgctggcaggagcatgccggtcactca
agaggggccgacccactccgaggacagggaccccccttgcacaaatgagaagagccagcaaacggttaatggttg
aatcctttgcagaaaaagtcaaggcagacttgtgattcttgatgttaatgaaggtgtcacacattgtctagggaaagttgt
gcctggcaatcttccctgtttcggatgcctaccccagtgacatttcatctcacatccaccttaatgaaatgcataattaaca
tattctaattgtgctgagcgcttcgatgattttgacagaactcactttccatgtctaatttgatgctaatttaatctgttcctctc
ccgctgacatgacctttcttgattgagtttcactgaggaaattttttgatggattgatttatcagcattgcagagacaacaa
aggctcagggagcacgacagaaaataggaagggaacggccctcgaagccagcacgctctgaacttgccggctgg
aatgtcctctttccagagacggcttgtctgacctactgatggaatttgagaaatgtgacatctgtgcgtgggtttgatc

>embl|AZ755961|AZ755961 yeast_INO80-like_protein

gatctcatgatgacaagcatcttggttttagtgattgtggttgcatacagctttgcattgagtgattttgtcgcttcaattaaa
tacagttccccctcaccccctttcttctctccccaggcaaggagagcctcatcactgacagtggaaagctgtatgcccttg
atgtcctgctgactcggctcaagtctcaagggcataggtccttatctactcccagatgaccaggatgatagacctact
ggaggtaggagagggaatcttggggacctaggataccttgacatttaatttgtaggaatcagatgtccatgactaatgtt
ggtcctgcttaatgggagccccaaaactcccttttgtctttttttctttttcctttttattttttgagacagagtct

>embl|AZ755982|AZ755982 paired_box_gene_5

gatcctgccaattcacccattaccacagcctcctgtgagatgagcgctcctctcacggtgagttaacgccggattagat
tcggagatttcaaggaaattccaagtggaacgaatcgcgtttcctttctgtttgctgcagagccaggctggaggaaggt
gacaggcttcggcagcctcggctcgggtgccgctcggggtctctgcgggcttcaggtcagggagttgctcccgggc
taggagccgtccccatgcctggtccaaacaacccgctcccccccatcccaacagcctgggatc
>embl|AZ756058|AZ756058 AT_motif_binding_factor_1
gatcagttacattcaccttaagaatattgtaaagaattcctgaaaacctttctccatgaacactttggcattccatcttgtttc
aatttcttggggtcaagaataatgatactgaccaggctgctttcctgcctcttccaaattggaattcccatctcgatctgga
acagctgctgggaggaacatgctgcttggcatcaccatctcaggggtggtcacctgaggagcaagagcaaggaga
gtcaaatggagaggaaagcgacacaaaataggaaatgtttaaaattatctaataagcaaaattagccttcaccagtgta
aagctagaggtgtctcacttttcatggctcactgaagattcatgtttattagacaaagctcacatcagtcaaaagtgtttct
acggttacattctggacataattagcagtgcacacttgaagaataatcttctgattaatattttccttatgcagaattttga
ctttgatc
>embl|AZ755799|AZ755799 PR_domain_containing_12
gatcagcgctggggacagacctcgggctggcaggtgttggaggtgccaaagaagcaccggcagcagctggccag
gctctgcgctcgctgtgtgaccccgggcgcgcccgagtccctctctgctcccaggccctgaggtgccagggaggg
cggcgcccgggctcagacaaaccccgatgtcatctctctgggcggggaggagccacggcggaccccaactcccc
gagggatcccctgtagaccatctgcggcagggctggaaacagcgaggcgcacaagggcagcctccatgtcttaca
aggttcccctctaagctacctgggaaaatacccggcgagcgggccggccgaagggggggg
>embl|AZ755927|AZ755927 peptidase_(mitochondrial_processing)_alpha
gatcacaggcctggactccaggttcatcatgagcatggacgtcagctgcgtcttggctcgcccccagctccaccgag
tccacggttacgcccattaaaatcaactcctctgcggtgacttctaccatttctggaacctgcctcgggtcggcgctggc
naaggatgcacaggaggccagtgtcctcgtagctgaggtaac
>embl|AZ755668|AZ755668 carboxypeptidase_A5
tgcatccagcctgagtcaccaacccagacacacactcttgctggccccagggaggcccagggagcgcgccccttct
cggtttggtggtcctcacctgggatggggaccctggttagaaggggatgtgccggcagagctagggggtgaaagg
cagcccccaggctggaatcagcccatctaacagtggccttgtttgttatgtctgccaggagctcctgcacctgcttcact
tgatgaatgacaaacagcaacagatgcgcagagcggggggaggttgaaggctttgtctgccttcgcggcccacgca
cagaaagcggttcccctgccaacgaccaggtggctttgcccaaatatggagaaacactttaggccattagttctggttt
gagctaagcagaataaactaatcttcccctggccgaatgggaattggctgcctgcatctctgacgttcacagtcaaggt
cctctgagtccgcctgtttcacaagtacctgctaatggattcattagctatttattttccttctctaagaatgtaatttgtagc
aaggggaaatc
>embl|AZ755749|AZ755749
9546bp_at_5'_side_ATPase,_Class_V,_type_10B|144913bp_at_3'_side_hypotheti
cal_protein
gatctaatcgtcaaaattctgggaaaattgatggttctggtctatgaggcatggcccaggcagttgctcaaacttcccag
ctccctccttcccagtttctgcacaggcacacagtcgacaaacactgacacagagccgagagaaggattggtgccct
agaagacatgttaatgttggcccagttctaattacaactctaacagccagagccatgcctggtagaaaaaccaaacca
aaccagaggcacattaatttctgccccagctggctcctatcatgggctggctcacatcagcgtgttcgaaggggccct
gcctgaacctctgatgccgaaggaatgcaccagctcagactctctcatttttaattagaagttcccattaaaaccactac
agcaatgagacacacaccaggcagcacactacaggacaattagcacactcctcaggtggctaaaatggcctggccc
tggcctgggggcaacataattgcaccagaacatacagcctaccaggaggttcccattttaagtggaaagaaataatta
gtcacacaatctattgggcagtttga
>embl|AZ755614|AZ755614 189998bp_at_5'_side_transforming_acidic_coiled-
coil_containing_protein_1_shor...|62764bp_at_3'_side_hypothetical_protein

148

gatcccagagcccagcgagagtttctctaaacatctgtttgagtcaaagctcagggcccaccccaagcatcagcata
ggccccggccctccgcagcccgcagcagcatgaagcagggattccctgagctgcatatcactgcaggttgcgtgaa
agagtatgtccaggtgggtggggccccgaaaagcagggtctggggctgagctctttgtaagtgagacccaggagc
cccagtttaattagaattcctgcctaggatgaaggctcgcggctgccagcccaggccttgagccattggtcacggctc
ataaaggccacacacaagtcacagagcccaggcgaggccagcatcccttgtagacggagcctcagaagtctctatg
atc

APPENDIX III. DETERMINATION OF THE

CELL CYCLE LENGTH OF B16F10 CELLS

USING DOUBLE THYMIDINE BLOCK

AIII.1 OBSERVATION AND RESULT

The murine melanoma B16F10 cell line was used for Pax3 functional

study in melanoma pathology in the lab. We observed that B16F10 cells grew

much faster than many other cell lines we used, and Pax3 expression was related

to cell cycle progression (see Chapter 4), but the cell cycle of B16F10 was not

studied in detail prior to our study. In general, normal mammalian cells replicate

over a 24-hour time span, but cells in different tissue have their specific turnover

rate (Norbury and Nurse 1992). We therefore applied double thymidine block to

B16F10 population to learn more about the B16F10 cell cycle.

Thymidine block is regularly used for synchronizing cells. Treatment with

high concentration of thymidine will arrest cells at their S phase so that cells will

progress together during cell cycle and the timing of one cell cycle can be studied.

Double thymidine block was applied to ensure the B16F10 cells to arrest at early

S phase (see Figure AIII-1), and we collected cells every 2 hours up to 14 hours

after the cells were released from the double thymidine block. By plotting

percentages of each cell cycle phase over a 14-hour span (see Figure AIII-2), it

was obvious that one cell cycle of B16F10 cells is about 14 hours.

AIII.2 METHOD

B16F10 cells were seeded in a 6-well plate with the growth medium.

Thymidine was added after 24 hours at a final concentration of 2mM. Upon 16

hours of thymidine treatment, the old growth medium with thymidine was

removed and cells were washed 3 times with PBS. The fresh growth medium was

then applied onto cells, followed with an 8-hr incubation. These cells were

blocked with second portion of thymidine. The cells were then washed 3 times

with PBS and released in fresh growth medium after 16 hours. Cells were finally

harvested at 0hr, 2hrs, 4hrs, 6hrs, 8hrs, 10hrs, 12hrs and 14hrs upon release, and

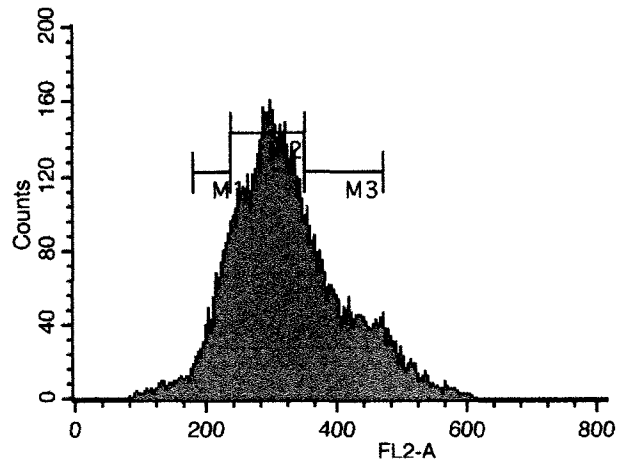prepared for cell cycle analysis using FACS that was described in Chapter 2.

Figure AIII-1. DNA content graph of B16F10 population treated with double thymidine block. B16F10 were harvested immediately after double thymidine block and were analyzed using FACS with PI staining. Markers M1, M2 and M2 indicate the regions of G0/G1, S and G2/M phase of cell cycle with % cells of 10.73%, 57.59% and 25.60%, respectively.
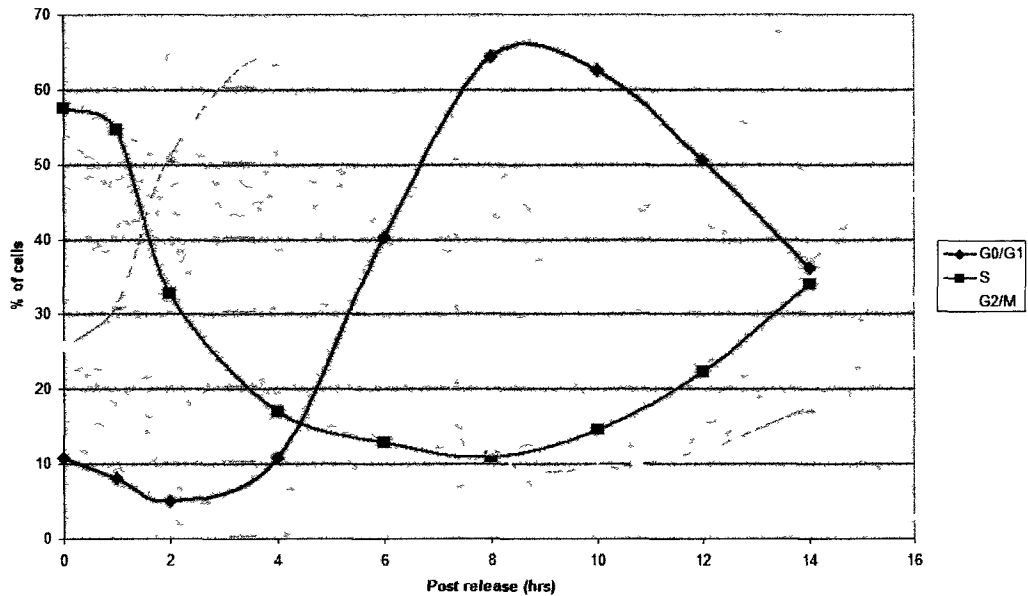

Figure AIII-2. Cell cycle progression of synchronized murine melanoma B16F10 cells. After double thymidine block, B16F10 cells were progressed from S phase (0hr) to G2/M (4hrs) and G0/G1 (~9hrs), and returned to S phase around 14 hours. G0/G1, S and G2/M cells are indicated in blue, pink and yellow, respectively.