A Preliminary Study of *c-kit* and Spermatogonial Stem Cells Differentiation

ZHANG, Lei

A Thesis Submitted in Partial Fulfilment
of the Requirement for the Degree of
Doctor of Philosophy
in
Obstetrics and Gynaecology

The Chinese University of Hong Kong
August 2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Table of content

Abstract

In adult mouse and human testes, spermatogenesis starts at spermatogonial stem cells (SSCs, also termed as undifferentiated spermatogonia) and ends at mature sperm. It has been proved that *c-kit* is crucial for proliferation, migration, survival and maturation of spermatogenic cells in embryonic and postnatal gonads. Expression of *c-kit* in the spermatogenic cells is periodic as they differentiate towards functional spermatozoa. However, expression profiles of *c-kit* mRNA and protein during SSCs differentiation and the upstream regulatory factors are unclear.

The aim of this study is to reveal and compare *c-kit* expression profile in the mouse SSCs before and after commitment of differentiation and find out the key factors to regulate this transition.

In this experiment, a SSCs cell line (c18-4, originated from the undifferentiated spermatogonia), a type B spermatogonia cell line (CRL-2053, originated from the differentiating spermatogonia) and testes from different aged mice (5 dpp, 10 dpp and 60 dpp) have been studied simultaneously. Transcription of *c-kit* was examined by Northern-blot with probes that hit against exons 10~12 or exons 18~20. Rapid amplification of cDNA ends (RACE) was applied to analyze the 5′ and 3′ end of *c-kit* transcripts. Real-time RT-PCR with *c-kit* transcript-specific primers was performed to assay the relative quantity of each *c-kit* transcripts. Immunofulorescence and Western blot were performed using antibodies specifically bound to the extracellular or intracellular domains of Kit, the protein product of *c-kit*. Having obtained the expression profiles of *c-kit* mRNA and protein in the two cell lines representing the

undifferentiated and the differentiating spermatogonia, we further confirmed the dynamics of *c-kit* mRNA in the testes from different aged mice and in the pluripotential embryonal carcinoma cell line (P19) stimulated by a known differentiation inducer RA (retinoic acid) *in vitro*.

Preliminary results showed that the full-length *c-kit* transcript was expressed in all type of cells and testes studied. A shorter transcript with a truncated 3′ end untranslated region (3′ UTR) was found in the differentiating spermatogonia but not in the undifferentiated spermatogonia. A SSCs specific transcript (starting from intron 9 and covering exons 10~21) was also discovered. Several truncated forms of *c-kit* transcripts encoding the intracellular domain of Kit were detectable in both the undifferentiated spermatogonia and the differentiating spermatogonia cell lines. Though the Kit protein containing the extracellular domain is not expressed before differentiation, one Kit protein that only has intracellular domain (50 kDa) exists in the SSCs as revealed by immunofluorescence and Western blot. The dynamics of the expression profiles of *c-kit* in the testis and the spermatogonial stem cell lines are different from each other after RA stimulation. The microenvironment around the germ cells in the testis may play an important role during RA induction. 2 μM of RA can induce embryonic carcinoma cell line (P19) to differentiate towards germ line cells. The wave-like changes of the quantitative expression pattern of *c-kit* (increase at first and decrease afterwards) during the induction process of P19 is similar to that of the *in vivo* male germ cell development process.

In conclusion, this is an on going study of the key gene *c-kit* in the SSCs differentiation. There are dynamic transcription and translation changes of *c-kit* gene before and after SSCs differentiation. RA is an important upstream regulatory factor for SSCs differentiation and *c-kit* expression. These changes may be either causes or consequences of differentiation.

中文摘要

在成年小鼠及人類睾丸中，精子發生起源于精原幹細胞（SSCs，為未分化精原細胞），終止於成熟精子。*c-kit* 基因對於胚胎期和出生後的生精細胞增殖、遷移、存活及成熟具有至關重要的作用。*c-kit* 基因的表達在生精細胞發育為功能成熟的精子這一過程中具有週期性。目前 *c-kit* mRNA 和蛋白表達譜在 SSCs 分化前、後的區別以及它們的上游調控因數尚不清楚。

本研究的目的是展現並比較 SSCs 分化前、後 *c-kit* 基因表達譜的區別，並探尋引起這種變遷的重要調控因數。

該實驗以一株 SSCs 細胞系（c18-4, 代表未分化精原細胞）， 一株 B 型精原細胞細胞系（CRL-2053，代表分化中精原細胞）和出生後 5 天、10 天、60 天小鼠睾丸為研究物件。用 Northern blot 的方法以針對全長 *c-kit* 轉錄體的 10~12 外顯子或 18-20 外顯子的探針檢測 *c-kit* 的轉錄。進一步以 RACE 分析不同轉錄體的 5′和 3′末端序列。以 Real-time PCR 分析不同 *c-kit* 轉錄體的表達水準。以針對 Kit 胞外區或胞內區的抗體行 western blot 和免疫螢光分析來明確 *c-kit* 蛋白表達情況。獲得 *c-kit* 在不同細胞系及不同年齡小鼠睾丸組織中的表達情況以後，我們以精原細胞誘導分化劑—維甲酸（RA）對不同細胞系、不同年齡小鼠睾丸組織和胚胎癌細胞細胞系（P19）進行體外誘導分化，並研究 *c-kit* mRNA 表達情況在該過程中的動態變化。

初步研究結果顯示，全長 *c-kit* 轉錄體在所有被研究的細胞系和睾丸組織中均有表達。在分化中精原細胞及睾丸組織中，我們檢測到一種包含截短 3′非轉錄區（3′ UTR）的新型 *c-kit* 轉錄體的存在，而該轉錄體在精原幹細胞中無表達。

此外，在我們還檢測到一種 SSCs 特異性轉錄體的存在，該轉錄體起源于 ***c-kit*** 基因 9 號內含子，包含 10~21 外顯子。僅由編碼 Kit 胞內區序列構成的截短型轉錄體在 SSCs 和分化中的精原細胞中均有表達。免疫螢光和 western blot 研究發現：雖然包含 Kit 胞外區的蛋白在 SSCs 中不表達，但是，一種可能僅包含部分 Kit 胞內區的蛋白（50 kDa）在 SSCs 中有表達。生精細胞微環境在 RA 誘導的 SSCs 分化和 *c-kit* 基因表達中可能發揮重要作用。2 μM RA 可以誘導多潛能的 P19 細胞定向分化為生精細胞，在這一誘導分化過程中，*c-kit* 基因的表達模式與生精細胞在體內發育的表達模式非常相似。

結論：該研究進一步探索了 *c-kit* 基因在 SSCs 分化過程中表達的變化和其重要調控因數。*c-kit* 基因的表達在精原幹細胞分化前、後發生動態變化；這種變化可能是 SSCs 分化的誘因，同時也可能是由於 SSCs 分化引起。RA 是 SSCs 分化和 *c-kit* 基因表達的重要調控因數。

Declaration

I hereby declare that this thesis represents my own work, except where due acknowledgement is made, and that it has not been previously included in a thesis, dissertation or report submitted to this University or to any other institution for a degree, diploma or other qualification.

Signature

ZHANG, Lei

Acknowledgements

I would like to express my sincere thanks to Prof. Han Yibing for her patient guidance and unlimited support throughout my course of study.

I would like to thank every member in Department of Obstetrics and Gynaecology, CUHK and my friends for all their support during these years.

I would also like to thank my family, especially my wife, for their encouragement during my study.

List of abbreviations

aa	amino acid
A_{al}	A aligned
ADH	alcohol dehydrogenase
AKT	thymoma viral proto-oncogene 1
Aldh1a2	aldehyde dehydrogenase family 1, subfamily A2
ALK3	(BMPR Ⅰ), bone morphogenetic protein receptor, type 1A
A_{pr}	A paired
A_s	A single
ATCC	American type culture collection
ATRA	all-trans retinoic acid
BAD	BCL2-associated agonist of cell death
Bcl-2	B-cell leukemia/lymphoma 2
Bcl6b	B-cell CLL/lymphoma 6, member B
bFGF	Basic fibroblast growth factor
BMP4	bone morphogenetic protein 4
BMP8b	bone morphogenetic protein 8b
Bp	base pair
CBP	sarcoplasmic calcium-binding protein
cDNA	complementary DNA
cdk2	cyclin-dependent kinase 2
c-Fos	FBJ osteosarcoma oncogene
c-JUN	Jun oncogene
c-kit	kit oncogene
CRABP	cellular retinoic acid binding protein
CSF1	colony stimulating factor 1
Cyp26b1	cytochrome enzyme P450
DAPI	4',6'-diamidino-2-phenylindole

DAZL	deleted in azoospermia-like
Dmc1	DMC1 dosage suppressor of mck1 homolog
DMEM/F12	dubecco modified eagle medium/F12
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNaseI	deoxyribonuclease I
dNTPs	deoxynucleoside triphosphate
dpc	days post coitum
dpp	days postpartum
E2F	E2F transcription factor
ERK	elk-related tyrosine kinase
Erm	ets related molecule
ESCs	embryonic stem cells
FBS	fetal bovine serum
FGF2	fibroblast growth factor 2
FSH	follicle stimulating hormone
FSHR	follicle stimulating hormone receptor
G-CSF	granulocyte colony stimulating factor
GDNF	glial cell line-derived neurotrophic factor
GRB2	growth factor receptor-bound protein 2
Kb	kilo base pair
kDa	kilo dalton
Kit	c-kit receptor
Kitl	Kit ligand
$Kitl_m$	membrane form of Kitl
$Kitl_s$	soluble Kit ligand
LASEC	laboratory animal service centre
LH	luteinizing hormone
Lhx1	LIM homeobox 1

mRNA	messenger RNA
MAPK	mitogen-activated protein kinases
Mvh	(Ddx4), DEAD (Asp-Glu-Ala-Asp) box polypeptide 4
Nanos2	nanos homolog 2
Nanos3	nanos homolog 3
Oct3/4	(Pou5f1), POU domain class 5 transcription factor 1
OD	optical density
ORF	open reading frame
p70S6K	rps6kb1, ribosomal protein S6 kinase
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PGCs	primordial germ cells
PI3K	phosphoinositide 3-kinase
PLCG	phospholipase C, gamma
PZLF	leukemia zinc-finger factor
RA	retinoic acid
RACE	rapid amplification of cDNA ends
RALDH	retinaldehyde dehydrogenase
RAR	retinoic acid receptor
RARE	retinoic acid responsive element
RAS	Rat Sarcoma
Rb	retinoblastoma protein
RBP	retinoic acid binding protein
RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
SCF	stem cell factor
SCP3	synaptonemal complex protein 3
SiRNA	small interfering RNA

Smad5	MAD homolog 5
Sohlh1	Spermatogenesis and oogenesis specific basic helix-loop-helix transcription factor
Sox3	SRY-box containing gene 3
SRC	rous sarcoma oncogene
Sry	sex-determining region Y
SSCs	spermatogonial stem cells
STAT3	signal transducer and activator of transcription 3
Stra8	stimulated by retinoic acid gene 8
Sycp3	synaptonemal complex protein 3
TAF4b	TATA-bingding protein associated factor 4b
TGF-β	transforming growth factor β
Tr-Kit	truncated form of Kit
UTR	untranslated region
VAD	vitamin A deficient

List of publications

Lei Zhang, Jiang jing Tang, Christopher John Haines, Yibing Han. *c-kit* expression profile and regulatory factors during spermatogonial stem cells differentiation. (in preparation)

Lei Zhang, Jiang jing Tang, Christopher John Haines, Huai Feng, Liang xue Lai, Xiao ming Teng, Yibing Han. *c-kit* and spermatogonial differentiation. (Already submitted)

Chan KY, Xiang P, Zhou L, Li K, Ng PC, Wang CC, Zhang L, Deng HY, Pong NH, Zhao H, Chan WY, Sung RY.Thrombopoietin protects against doxorubicin-induced cardiomyopathy, improves cardiac function, and reversely alters specific signalling networks. Eur J Heart Fail. 2011. 13(4): 366-376.

Ping Xiang, Hai Yan Deng, Karen Li, Guo-Ying Huang, Yuan Chen, Liu Tu, Pak Cheung Ng, Nga Hin Pong, Hailu Zhao, Lei Zhang, Rita Yn Tz Sung. Dexrazoxane protects against doxorubicin-induced cardiomyopathy: upregulation of Akt and Erk phosphorylation in a rat model. Cancer Chemother Pharmacol. 2009. 63(2): 343-349.

List of conference abstracts

Lei Zhang, Yibing Han. A primary study of *c-kit* and spermatogonial stem cells differentiation. Thirty Years of Advances in Reproductive Endocrinology and ART: A Celebratory Congress. Montreal, Canada, May 27-31, 2011.

Lei Zhang, Christopher John Haines and Yibing Han. EXPRESSION OF *C-KIT* MRNA AND KIT PROTEIN DIFFERS IN SPERMATOGONIAL STEM CELLS BEFORE AND AFTER DIFFERENTIATION IN MOUSE. 35th ASA Annual Meeting, Houston, USA, April 10-13, 2010.

Chapter 1

Introduction

1.1 Spermatogenesis

1.1.1 Definition and staging

Spermatogenesis is a highly regulated process of differentiation and complex morphologic alterations that leads to the formation of sperm in the seminiferous epithelium. It takes approximately 35 days in mice and 64 days in human. The entire process can be subdivided into two successive sections. The first section is termed spermatogenesis which comprises the cells from the diploid spermatogonia to the haploid secondary spermatocyts (meiosis). The second section is termed as spermiogenesis which comprises the the haploid spermatids and the matured spermatozoa. The following developmental stages are thereby passed through in the entire spermatogenesis and spermiogenesis from A-spermatogonium, B-spermatogonium, primary spermatocyte (spermatocyte order I), Secondary spermatocyte (spermatocyte order II), and Spermatid to Sperm cell (spermatozoon) (Figure 1.1, page 25).

1.1.2 Location of spermatogenesis

Spermatogenesis takes place within several structures of the male reproductive system. The initial stages occur within the testes and progress to the epididymis where the developing gametes mature and are stored until ejaculation. The germinal epithelium (seminiferous tubules) of the testes are the starting point for the process.

where stem cells adjacent to the inner tubule wall divide in a centripetal direction-beginning at the walls and proceeding into the lumen-to produce immature sperm. Final maturation of the spermatozoa occurs in the epididymis. The tails are formed during spermiogenesis in the testis, and some sperm obtained from the testis can slowly move. Modification of plasma membrane of spermatozoa occurs in the epdidymis that make the sperm move normally.

1.1.3 Structure of the germinal epithelium

The male germ epithelium consists of the Sertoli cells and the spermatogenic cells. The Sertoli cells form a single-layered lamina and extend from the basal lamina to the tubule lumen. With their labyrinthine cellular processes, they surround the individual types of germ cells. Spermatogenesis is accomplished in close contact with the Sertoli cells, which not only have supportive and nourishing functions, but also secrete hormones and phagocytize cell fragments. Above the basal lamina they are bound to each other through complicated occluding junctional complexes (tight junctions), so that 2 separated compartments are present in the epithelium: a basal one, in which the spermatogonia are lined up, and a luminal one, in which all the other stages of spermatogenesis are found. Through the occluding junctional complexes of the Sertoli cells a "blood/testicle" barrier is created in the tubule. This means that outside this barrier, in the tubular periphery, cells, substances and hormones from the blood have unhindered access.

On the other hand, the inner compartment of the tubule is protected by the "blood/testicle" barrier, which is selectively permeable and serves as an entry check.

This is of practical importance because haploid cells in the inner part of the tubule exhibit surface antigenic properties, different from all other body cells. They must thus be kept secluded from the immune system of the organism by the "blood/testicle" barrier (Figure 1.2, page 26).

1.1.4 Spermatogenic cycle

In adult men and mice, the production of spermatozoa is constant due to temporal cycling, termed as the "spermatogenic cycle", and spatial phasing, termed as the "spermatogenic wave" of germ cell differentiation (Aponte et al, 2005). The result of the temporal and spatial constraints is the cyclic appearance of specific cellular associations within a given cross section of the seminiferous tubule. These recurring cellular associations are termed the "stages of the cycle of the seminiferous epithelium", twelve of which are recognized in the mouse (Figure 1.3, page 27) (OAKBERG, 1956; Onishi et al, 1983) . In normal adult mice, all twelve stages are present at any given time and occur sequentially along the length of the seminiferous tubule. The distance between a given stage and the reappearance of that stage is termed a "spermatogenic wave" (Figure 1.4, page 28). The processes that give rise to the spermatogenic wave result in continuous or asynchronous release of spermatozoa from the seminiferous tubules.

1.2 Spermatogonial stem cells (SSCs) and the cohorts derived from them

In adult mouse and human testes, spermatogenesis is started from the spermatogonial stem cells (SSCs). SSCs also called A single (A_s) spermatogonia, are

located on the basal membrane of seminiferous tubules. The A_s spermatogonia can self-renew or produce Apaired (A_{pr}) spermatogonia. After successive divisions, A_{pr} spermatogonia differentiate and form chains of 4, 8 or 16 aligned spermatogonia (A_{al}) and migrate along the basal membrane. Based on morphological criteria, SSCs, A_{pr} and A_{al} spermatogonia are classically called undifferentiated spermatogonia. The undifferentiated spermatogonia comprise <1% of the entire testicular cells, have been shown experimentally to harbor eventually all stem cell activity (Shinohara et al, 2000; Ohbo et al, 2003). At present, there is no phenotypic, biochemical, or molecular characteristics that can distinguish the undifferentiated spermatogonia populations from one another. A_{al} spermatogonia differentiate into more committed A1 spermatogonia that will divide and differentiate into A2, A3, A4, intermediate and B spermatogonia, which will undergo meiosis after a final mitosis (Barroca et al, 2009) (Figure 1.5, page 29). Theoretically, a total of 4096 mature spermatozoa could be produced from a single SSC in the testis. However, the overall efficiency of spermatogenesis is estimated to be around 10 to 25% in adult rat because of apoptosis. As a result, a considerably fewer spermatozoa than potential are actually generated (Tegelenbosch and de Rooij DG, 1993).

Surrounded by Sertoli cells, SSCs live in the microenvironment (niche) formed by the Sertoli cells. Within the niche, growth factors and extracellular signals regulate the fate decisions of SSCs either to self-renew or to form daughter cells that will begin the complex differentiation process of spermatogenesis, resulting in mature spermatozoa. The timing of sequential steps in spermatogenesis is tightly regulated by

the germ cell, and the Sertoli cells support the differentiation process(Brinster，2007). In a normal seminiferous epithelium, the ratio between self-renewal and differentiation of spermatogonial stem cells should be about 1.0. More self-renewal than differentiation would reduce the seminiferous epithelium to only stem cells and a tumor might form. If differentiation prevails, the stem cells would deplete themselves and only the Sertoli cells would be remained in the seminiferous epithelium (de Rooij DG，2001).

1.3 Self-renewal and differentiation control factors for the SSCs

1.3.1 Self-renewal control factors

Very little information is known about spermatogonial renewal mechanisms. Extracellularly secreted factors should play essential roles in the stem cell-niche interactions. Glial cell line-derived neurotrophic factor (GDNF), one of the transforming growth factor (TGF-β) superfamily produced by Sertoli cells in the mammalian testis, is one of the possible control factors (Meng et al，2000；Tadokoro et al，2002). GDNF is identified as a critical factor *in vivo* for the replication of spermatogonia (Meng et al，2000). *In vitro* studies using serum-free culture medium demonstrates that GDNF is the primary growth factor supporting mouse SSCs self-renewal (Kubota et al，2004). GDNF binding and signaling occur through GDNF-family receptor α1 (GFRα1) and the Ret receptor in the spermatogonia. In the presence of GDNF, SSCs grow on the feeder cells in the shapes of clumps. If GDNF is removed, the clump cells begin to grow in chains resembling the initial stages of stem cell differentiation, as seen *in vivo* (de Rooij DG，1998；Hamra et al，2005；

Ryu et al, 2005). Thus, GDNF appears to be a primary regulator for SSCs self-renewal in mice and rat (Kubota et al, 2004; Ryu et al, 2005). It is also probably a conserved self-renewal signal for all mammalian SSCs (Kubota et al, 2004; Hamra et al, 2005; Ryu et al, 2005).

Functional transplantation assay and long-term culture system of SSCs make it possible to examine possible intracellular signals that influence self-renewal and differentiation *in vitro* in a rigorous manner which is usually not available for most of the other types of adult stem cells (Oatley et al, 2006). Studies have demonstrated that Oct3/4 (also called pou5f1) and SRY-box-containing gene 2 (Sox2), which regulate Nanog, are expressed in the SSCs. However, Nanog, the key determinant of embryonic stem cells (ESCs) self-renewal and pluripotency, is not expressed in the SSCs (Oatley et al, 2006). Therefore, it seems that the signaling mechanisms regulating self-renewal in SSCs and ESCs are different. These studies also demonstrate that the expression of three transcription factors including B cell CLL/lymphoma 6 member B (Bcl6b), Ets-related molecule (Erm), and LIM homeobox 1 (Lhx1), is highly regulated by GDNF *in vitro* (Oatley et al, 2007). Functional transplantation assays have confirmed the importance of Bcl6b. Bcl6b is a BTB/POZ family transcription factor, and experiments with mouse or rat SSCs indicate that Bcl6b is important for maintenance of the undifferentiated state of SSCs, but not for more differentiated germ cells (Oatley et al, 2006; Schmidt et al, 2009). Erm expression in Sertoli cells is believed to affect niche function. Mice with targeted disruption of Erm have a loss of maintenance of spermatogonial stem cell

self-renewal without a block in normal spermatogenic differentiation and thus have progressive germ cell depletion and a Sertoli-cell-only syndrome (Chen et al，2005). Other studies have also indicated that several genes, not necessarily being regulated by GDNF, including the promyelocyte leukemia zinc-finger factor (Plzf), TATA-binding protein-associated factor 4b (TAF4b), and Nanos 2 (NOS2) may play roles in SSCs self-renewal (Falender et al，2005；Wong et al，2005；Oatley et al，2006；Suzuki et al，2007).

Other extracellular factors including CSF1 (colony stimulating factor 1, also known as granulocyte-colony stimulating factor, G-CSF) (Kokkinaki et al，2009；Oatley et al，2009) and bFGF (basic fibroblast growth factor, also known as FGF2) (Kanatsu-Shinohara et al，2003；Kubota et al，2004) may also be candidates of SSCs self-renewal factors. They promote cell proliferation of SSCs. All of these control factors may work together synergetically or antergically.

1.3.2 SSCs differentiation control factors

The culture and transplantation systems for germ cell differentiation are absent. This complicated the examination of regulatory mechanisms of this process. Moreover, the intricate three-dimensional structural organization of spermatogenesis has compounded the problem. In addition to surrounding the stem cell to provide a regulatory niche, the Sertoli cell extends about 90 μm from the basement membrane to the lumen of the seminiferous tubule, contacting and surrounding germ cells in many stages of differentiation. Furthermore, an individual germ cell may associate

with more than one Sertoli cell (de Rooij DG，1998). Therefore, knowledge on SSCs differentiation regulatory factors is limited.

STAT3 (signal transducer and activator of transcription 3) may promotes the differentiation of SSCs. STAT3 has been identified as a central regulator of mouse ESCs cell pluripotency and self-renewal. In the germ-line of neonatal mouse testes expression of STAT3 is localized to gonocytes and the undifferentiated spermatogonial population, which contains SSCs. Transient impairment of STAT3, signaling enhances SSCs self-renewal *in vitro* without affecting general spermatogonial proliferation, indicating an alteration in the balance of SSCs fate decisions that inhibited differentiation (Oatley et al，2010). In males, spermatogenesis and oogenesis specific basic helix-loop-helix (bHLH) transcription factor (Sohlh1) is preferentially expressed in prespermatogonia and type A spermatogonia. Loss of Sohlh1 down-regulates expression of *c-kit* and causes infertility by disrupting spermatogonial differentiation into spermatocytes (Ballow et al，2006). Sohlh2, another member of the same group as Sohlh1 might coordinate with Sohlh1 to promote the differentiation of *c-kit*$^{+}$ germ cells *in vivo* (Toyoda et al，2009).

The transition of undifferentiated spermatogonia into differentiating spermatogonia coincides with the gain of *c-kit* expression. *c-kit* receptor (Kit) expression and interaction with Kit ligand (Kitl, also named stem cell factor (SCF)) are considered to be crucial for proliferation, migration, survival and maturation of germ cells in embryonic and postnatal gonads (Yoshinaga et al，1991；Dym et al，1995；Packer et al，1995；Orth et al，1997；Ohta et al，2000；Yan et al，2000；

Guerif et al, 2002; Prabhu et al, 2006; Runyan et al, 2006; Gu et al, 2009). Activation of Kit signaling is required for differentiation of spermatogonia into spermatocytes, as revealed by the inductive expression of early meiotic markers (such as Dmc1 and Scp3) (Pellegrini et al, 2008). Down-regulation of Kit signaling after meiosis is also important. Mice carrying a mutation rendering a constitutively active Kit kinase have an interrupted transition from the round into the elongating spermatids (Schnabel et al, 2005).

1.4 *c-kit*, SSCs differentiation and spermatogenesis

1.4.1 Transcription and translation of *c-kit* and its ligand in spermatogenic cells

c-kit gene is allelic to the W locus on mouse chromosome 5 (Chabot et al, 1988). The 21-exon gene encodes a 5150 bp transcript which is translated into a product of 145 kDa protein with 979 amino acid residues which is called Kit (Yarden et al, 1987). *c-kit* mRNA and protein synthesis are regulated separately possibly by circulating hormones as the undifferentiated spermatogonia contains only *c-kit* mRNA but not protein (Prabhu et al, 2006). Kit belongs to a family of growth factor receptors with intrinsic tyrosine kinase activity that transduces growth regulatory signals across the plasma membrane. Kit has three main functional regions: the extracellular domain, the transmembrane region and the intracellular domain. The extracellular domain consists of five immunoglobin-like repeats with about 520 amino acids which are required for ligand binding and dimerization (Blechman et al, 1995). The transmembrane region is a 23 amino acid hydrophobic domain, which anchors the receptor to the cell membrane. The 433 amino acid intracellular domain consists of

three domains, with a proximal kinase region for ATP binding, a 70-100 amino acid non-conserved insert and a distal phosphotransferase kinase region (Blechman et al, 1993). Tyrosine (Tyr) residues in the intracellular juxtamembrane domain serve as docking sites for signal transduction molecules that undergo activation (Roskoski Jr, 2005). Binding to the ligand of Kit (Kitl) induces a rapid and complete receptor dimerization that involves activation by autophosphorylation of the catalytic tyrosine kinase and generates signal transduction (Lev et al, 1992). Detailed structure of *c-kit* mRNA and protein are shown on Figure 1.6 (page 30).

c-kit has function not only during spermatogonial proliferation, but may throughout all stages of male germ cell development before and after birth. Northern blot analysis of germ cells at different developmental stages has shown the presence of two alternative mRNA of *c-kit*, 3.2 and 2.3 kb in length respectively, in the haploid cells of the mouse testis (Sorrentino et al , 1991). The two alternative spermatid-specific *c-kit* transcripts originate from the 16^{th} intron of the mouse *c-kit*, and contain all the downstream exons (Rossi et al, 1992) (Figure 1.6, page 30). These alternative *c-kit* mRNA encode for a truncated and soluble form of the Kit of ~30 kDa, called Tr-Kit, with an ORF (Open Reading Frame) that starts in the intron 16 and encodes for 12 hydrophobic amino acids followed by the last 190 carboxy terminal residues of the Kit (Rossi et al, 1992; Sette et al, 1997). Tr-Kit derived from alternative promoter in the intron 16 of *c-kit* and encodes part of the non-conserved insert from the C-terminal tail region and the distal phosphotransferase kinase region, and it lacks the entire extracellular and the transmembrane domain (Albanesi et al,

1996). This intronic promoter of the *c-kit* is only active in the late stages of spermatogenesis, suggesting a role for this truncated protein either during spermatid differentiation or for the function of mature sperm (Sette et al，2000). Tr-Kit is found in the residual sperm cytoplasm. There is evidence for its serving a function in the activation of oocyte at fertilization in mice (Sette et al，1997). Tr-Kit expression also appears to correlate with sperm DNA integrity (Muciaccia et al，2010).

Kitl is produced in the Sertoli cells and a cytokine essential for haematopoiesis, melanogenesis and development of germ cells. Kitl has been identified as an analogue of the murine steel (Sl) locus and is located on chromosome 12 in humans. Two isoforms of Kitl are generated from the same gene by alternative splicing - a soluble ($Kitl_s$) and a transmembrane ($Kitl_m$) form (Anderson et al，1990; Flanagan et al，1991). The soluble form arises after proteolytic cleavage of a membrane-bound precursor (Toksoz et al，1992). In the spermatogonia proliferating stage, the membrane isoform is a predominant one; whereas in the Spermatogonia quiescent stage, the soluble form dominates.

1.4.2 Kit/Kitl-dependent mechanisms during spermatogenesis

Four pathways are known to be activated in response to Kit/Kitl activation in the spermatogonia (Figure 1.7, page 31). The PI3K pathway results in cell survival (via AKT and BAD regulation), adhesion (via c-JUN and c-FOS activation) and proliferation (via AKT and p70S6K). The PI3K/AKT pathway appears to be critical exclusively in postnatal stage spermatogenesis. Mice with a mutant form of Kit incapable of PI3K recruiting are sterile caused by reduced proliferation and

subsequent apoptosis in the spermatogonia (Blume-Jensen et al，2000). Cyclin might be one of the targets of Kit/Kitl pathway in the testis. Through PI3K pathway, Kit/Kitl facilitate the up-regulation and nuclear accumulation of cyclin D3, thus inducing spermatogonia to proliferate (Feng et al，2000；Dolci et al，2001). The SRC pathway involves the association of SRC family proteins with the intracellular juxtamembrane domain of Kit and affects cell migration and AKT phosphorylation in mice PGCs (Farini et al，2007). Tr-Kit activated PLCG through the PLCG pathway, mediates the resumption of meiosis of the fertilized eggs (Sette et al，2002). The MAPK cascade is activated by RAS with the binding of Kit and GRB2 (growth factor receptor-bound protein 2). MAPK directly mediates gene transcription in PGCs and proliferation in spermatogonia (Dolci et al，2001；Farini et al，2007；Mithraprabhu and Loveland，2009).

1.4.3 Roles of *c-kit* in embryonic and neonatal spermatogenic cells

1.4.3.1 *c-kit* and spermatogenic cell proliferation and restoration

In mouse, at around 7.2 days post coitum (dpc), somatic signals earmark a small cohort of proximal epiblast cells as potential germ cell precursors (Ginsburg et al，1990). This group of cells moves into the extraembryonic tissue at the base of the allantois where a second round of selection occurs, which results in a group of about 45 cells specified to be germ cell precursors or PGCs. After specification, the germ cells become transcriptional silent at 9.5 dpc and are subject to an extensive reprogramming of their genomes by histone modifications and alterations in the state of DNA methylation (Seki et al，2005). *c-kit* mRNA is first detected in the PGCs at

6.5-7 dpc and persists during their subsequent proliferation and migration to the genital ridge (7.5-13.5 dpc) (Orr-Urtreger et al，1990；Manova and Bachvarova，1991). In the mean time, the somatic cells along the migratory pathway and genital ridges synthesize Kitl (Matsui et al，1990；De Felici M et al，1996；Runyan et al，2006；Gu et al，2009). In the absence of either Kit or Kitl, mice are sterile and with a reduced number of PGCs (Gu et al，2009). Kitl secreted by the somatic cells seems to be an attractant for germ cells migration and are required for their adhesion, proliferation, migration and survival prior to 10 dpc after which down-regulation of Kitl is associated with switching on the intrinsic apoptotic pathway in ectopic germ cells (Godin et al，1991；Runyan et al，2006；Farini et al，2007). We wonder if Kit/Kitl pathway may facilitate SSCs survival by suppressing apoptosis as that in the ES cells (Bashamboo et al，2006).

Male PGCs arrest in G0/G1 of the mitotic cycle around 13.5 dpc and resume mitosis around 3 dpp during which Kit expression is markedly reduced in mice (Orr-Urtreger et al，1990). At around 3 dpp, expression of Kit is still low when the male PGCs actively proliferate again. The transition from *c-kit* independent type to ***c-kit*** dependent type occurs at about 5 dpp when the competence to enter meiosis is reached (Tajima et al，1994；Ohta et al，2000, 2003). Expression of Kit (3 dpp) is before the expression of Kitl (6-8 dpp) and their expression is closely coordinated (Rossi et al，1991；Tajima et al，1991). Unlike the chemo-attractant function Kitl takes during spermatogenic cell migration in the prenatal stage, expression of Kitl in the neonatal stage seems to be a triggering factor for spermatogenesis. Kit/Kitl expression

during spermatogenesis was specifically demonstrated on Figure 1.8 (page 32).

Expression of Kit in the SSCs is contradictory. In the early studies, Kit expression in the adult testis is detected by immunohistochemical analysis and in situ hybridization in the differentiating type A (A1–A4), intermediate, and type B spermatogonia, as well as preleptotene spermatocytes and interstitial Leydig cells, but not in undifferentiated spermatogonia and Sertoli cells (Yoshinaga et al，1991；Schrans-Stassen et al，1999). Hence, activation of the Kit/Kitl signaling pathway is not required for SSCs self-renewal (Kubota et al，2009；Morimoto et al，2009).

Studies that are more recent demonstrate that both Kit^- and Kit^+ cells showed comparable levels of stem cell activity after germ cell transplantation (Barroca et al，2009；Morimoto et al，2009；Trefil et al，2010). As SSCs can change their phenotype according to their microenvironment, Kit^+ cells might be an intermediate state during SSCs self-renewal (Morimoto et al，2009). Izadyar et al. further characterize the Kit^+ SSCs and find that the $POU5F1^+/Kit^+$ subset of mouse SSCs generates cell lines that express pluripotent ES markers and can differentiate into multiple lineages. However, *in vivo* testes regeneration assay shows that only the $POU5F1^+/Kit^-$ SSCs will regenerate the spermatogenesis of the recipient tests (Izadyar et al，2008). Kit seems not to affect SSCs self-renewal directly; it may affect the size of SSCs pool by playing a role during the phenotypic transition of SSCs (Figure 1.9, page 33).

1.4.3.2 Role of *c-kit* in onset of meiosis

Synthesis of *c-kit* mRNA and protein in postnatal mouse testes is concordant with the first appearance of differentiating spermatogonia, which persists at relatively

lower levels in meiotic pachytene spermatocytes (Prabhu et al，2006). The presence of Kit has been routinely used as a marker to identify differentiating spermatogonia (Shinohara et al，1999, 2000).

It has been demonstrated that the timing of meiosis entry is indirectly controlled by the Sertoli cells through the activation of Kit/Kitl system when they are induced by RA (Lufkin et al，1993；Packer et al，1995；Vincent et al，1998；Yan et al，2000；Pellegrini et al，2008). *In vitro* experiment has proved that addition of RA to the Kit expressing spermatogonia induces the onset of spermatogenesis but not the Kit negative spermatogonia (Pellegrini et al，2008). Kit/Kitl activation causes a transient activation of ERK1/2 and PI3K-dependent AKT kinase. These events are followed by a rapid nuclear redistribution of cyclin D3 and accumulation of cyclin E and promotes cell cycle progression via the PI3K/p70 S6 kinase pathway (Feng et al，2000；Dolci et al，2001). Hyperphosphorylation of retinoblastoma protein Rb by cyclin E/cdk2 is followed by the release of Rb-associated transcription factor E2F, which elicits timely induction of other genes required for S-phase progression (Rossi et al，2003). Silencing *c-kit* expression by siRNA in the spermatogonia induces cell cycle arrest also proves the role of Kit on meiosis entrance (Sikarwar and Reddy，2008). Transcriptome analysis of the Spermatogonia treated with Kitl indicates that Kitl stimulates their entrance of meiotic program by up-regulating the G1/S transition inhibitors and G2/M promoters and by down-regulating the G1/S promoters (Rossi et al，2008). Microgravity also promotes spermatogonia expression of *c-kit* and stimulated by retinoic acid gene 8 (stra8) and induces the last round of DNA

replication (preleptotene stage) in the *c-kit*$^{+}$ spermatogonia and activates the PI3K pathway (Pellegrini et al，2010).

1.4.3.3 *c-kit* and mature sperm

It has been shown that the *c-kit* gene can be translated into two kinds of proteins, Kit and Tr-Kit respectively, during spermatogenesis in human. The Tr-Kit is found to be expressed in the post-meiotic haploid germ cells and maintained in the motile spermatozoa. Cytometric analysis of several human sperm samples has showed variable degrees of the Tr-Kit-specific immunolabeling, and a significant inverse correlation of the Tr-Kit to the markers of sperm damage, i.e. DNA fragmentation, as revealed by terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate-nick end labeling (TUNEL) analysis and the intense clusterin positivity. Therefore, the maintaining of Tr-Kit in the haploid spermatogenic cells appears to correlate with the next stage spermatozoa DNA integrity (Muciaccia et al., 2010). Whether Kit is present in these haploid germ cells is under debate. Muciaccia et al. find that Kit and its coding mRNA are not detected in the spermatozoa (Muciaccia et al., 2010). Feng et al. show that the mature human spermatozoa expresses Kit and its presence appears to be correlated with sperm capacitation and the acrosomal reaction. The percentage of sperm undergoing acrosomal reaction declines and the percentage of head-to-head agglutination increases after blocking Kit with its antibodies (Feng et al., 2005).

1.5 Mechanisms/factors controlling the activation of Kit/Kitl pathway

The mechanisms/factors controlling *c-kit* expressing during spermatogenesis are not very clear. However, several upstream regulating factors have been discovered in the studies of PGCs, oogenesis and other organisms.

1.5.1 Vitamin A and its derivatives

Retinoic acid (RA), an active metabolite of vitamin A, is a vital signaling molecule for normal fetal development, pattern formation, cell proliferation and differentiation, and apoptosis (Livera et al，2002). RA is synthesized by the mesonephroi to which the gonads are attached (Bowles et al，2006). Metabolism of vitamin A derivatives are indicated in Figure 1.10 (page 34). RA is also a PGCs inducer (Vogel，2003). ES cells will differentiate to PGCs when culture with 2 μM RA for 5 days (Eguizabal et al，2009). RA has been successfully used to induce ES cells into functional spermatids expressing stra8 and Mvh (Nayernia et al，2006). RA is considered to be crucial for germ cells to enter meiosis in both male and female (Bowles et al，2006；Koubova et al，2006). *Aldh1a2* encodes the alcohol dehydrogenase that turns retinol into retinal (functional RA) and *Cyp26b1* encodes a family of cytochrome P450 enzyme that can convert RA into inactive forms. Thus, *Aldh1a2* and *Cyp26b1* acts as a 'source' and 'sink' of RA, whose finely control defining when and where RA mediated signaling will occur (Li et al，2009). Expression of *Cyp26b1* is detected in the peritubular myoepithelial cells in the testis around 11.5 dpc when *Sry* begins to express. In the *Cyp26b1*-null male mice, *Stra8* expression is up-regulated amd an early entrance of meiosis in 14.5 dpc is observed. However, the spermatogonia with an early entry of meiosis go apoptosis instead of

normal spermatogenesis with a lethal morphogenetic phenotype (*Spina bifida*) (Baltus et al，2006；Bowles et al，2006；Koubova et al，2006；MacLean et al，2007). Expression of *Aldh1a2* in the gonad is initiated at 11.5 dpc in the mesonephros and is maintained until 14.5 dpc in mice. Postnatally, *Aldh1a2* transcript are detected at 1 dpp, and its expression increases significantly until 20 dpp when protein is detected in the pachytene spermatocytes. In the adults, expression of *Aldh1a2* is restricted to the round spermatids (Wu et al，2008). Disruption of *Aldh1a2* rescues the *Cyp26b1*-null mice lethality (Niederreither et al，2002).

Testes of adult vitamin A-deficient mice/rat (VAD; deprived of dietary vitamin A) exhibit seminiferous tubules containing only Sertoli cells, type A Spermatogonia and few preleptotene spermatocytes (Morales and Griswold，1987). The type A spermatogonia are almost all arrested before differentiation at A1 Spermatogonia, with a reduced *c-kit* expression and no stra8 expression. Administration of vitamin A to these animals results in synchronized spermatogenesis emerging from type A spermatogonia and enhanced expression of *c-kit* (van and de Rooij DG，1991).

RA may regulate proliferation and differentiation of spermatogonia mainly through RARα mediated signal pathway. During post-natal development, RARα and RXRβ are confined to Sertoli cells, whereas RARγ is expressed in spermatogonia followed by a colocalization of RARβ, RXRα, and RXRγ to the step 7-8 spermatids (Vernet et al，2006). RARα knockout models ($RAR\alpha^{-/-}$) showed germ cell apoptosis and seminiferous epithelium dysfunctions related to the disruption of Sertoli cells cyclical gene expression, which preceded testis degeneration. Deletion of RARβ or

RARγ, on the contrary, does not cause primary testis defects (Lufkin et al，1993；Vernet et al，2006). As RARα is only expressed in Sertoli cells, it is theoretically possible that RA acts on Sertoli cells, which then send a secondary signal to germ cells to induce meiosis (Vernet et al，2006). RA binding with RARs recognizes and binds an RARE sequence which induces stra8 expression (Chiba et al，1997). RA pathways affect on spermatogonia self-renewal and differentiation is shown on Figure 1.11 (page 35).

RA acts to initiates meiosis both in male and in female. In male, exogenous RA can induce XY A_{al} staged germ cells in a cultured mouse fetal testis to enter meiotic prophase (Snyder et al，2010). It is not yet known whether the action of RA in inducing differentiation and *c-kit* expression is direct, or indirect via Kitl in the Sertoli cell. It is accepted that RA control the timing of meiosis indirectly by juxtacrine of Sertoli cells (Pellegrini et al，2008). Some studies show that RA directly act on spermatogenic cells by stimulating Stra8 and *c-kit* gene expression, whereas some studies show exogenous RA could not stimulate *c-kit* expression in spermatogenic cells but cause apoptosis of the A_{al} Spermatogonia (Wang and Culty，2007；Zhou et al，2008；Snyder et al，2011). Besides, RA also up-regulates Kitl levels in the Sertoli cells, resulting in increased levels of the early meiotic cell markers. This activation is independent of germ cell viability and occurs through the phosphatidylinositol 3-kinases (PI3K) and MAP kinase (MAPK) pathways (Pellegrini et al，2008).

1.5.2 BMP4/ALK3/SMAD5 signaling pathway

BMP4, one of the TGFβ-BMP superfamily growth factor, is produced by Sertoli

cells very early in the postnatal life and is down regulated during peri-pubertal. BMP4 treatment of the PGGs, SSCs and spermatogonia increases Kit levels and causes a mitogenic response to Kitl (Pesce et al，2002；Pellegrini et al，2003；Carlomagno et al，2010). BMP4 expression was significantly up regulated in the testes of VAD mice and was down regulated in freshly isolated germ cells treated by retinol. This reflects a direct requirement for retinoid by germ cells for the resumption of spermatogenesis in VAD animals via mechanisms that involve the suppression of BMP4 expression (Baleato et al，2005). Receptors of BMP4 (ALK3 and BMPIIR) are specifically expressed in mitotic spermatogonia during the first week after birth. BMP4 action is mediated by a rapid nuclear translocation of *Smad4* and *Smad5*, where the Smad4/Smad5 complexes are able to recruit the transactivating factor CBP and to bind Smad-responsive DNA sequences.

Another member of the TGFβ-BMP superfamily growth factor is BMP8b that stimulate both PGCs and spermatogonia to proliferate. BMP8b$^{-/-}$ mice show impairment of PGC commitment, defects of spermatogonia proliferation and spermatocyte apoptosis (Zhao et al，1996；Ying et al，2000).

1.5.3 FSH

FSH, LH and the testis androgen are involved in the process of orchestrated control of spermatogenesis. FSH is not essential for spermatogenesis but is required for quantitatively normal sperm production in both mice and human (Kumar et al，1997；Dierich et al，1998；Kumar et al，1999；Abel et al，2000；Vaskivuo et al，2002). FSH works directly on Sertoli cells via their receptors, the FSH receptor

(FSHR). Kitl is expressed by Sertoli cells under FSH stimulation. Therefore, the Sertoli cells of the genetic mutant mice lacking FSH receptor will produce less Kitl. These mutant mice exhibit weight loss of testis, epididymis, and seminal vesicle as well as low levels of testosterone. A significant increase of *c-kit*$^{+}$ spermatogonia and a significant decrease of the elongated spermatids are observed in these mice. The increase in the percentage of *c-kit*$^{+}$ cells and decrease in the testosterone values of FSH receptor in the mutant mice may be due to the reduced levels of Kitl available for intercellular communication in the absence of FSH receptor signaling (Krishnamurthy et al，2000). In the Sertoli cells, FSH can regulate transcriptional function of the RARα, thus controls the cell proliferation and differentiation (Santos and Kim，2010). Taken together, FSH might determine the expression of *c-kit* in the spermatogonia via Sertoli cell factors including Kitl and RARα.

1.6 *c-kit* negative mitotic arrest and testis tumor

Plzf, Nanos, Bcl6b, Oct3/4, Neurogenin3 and *Sox3* are markers of the undifferentiated spermatogonia. The DNA sequence-specific transcriptional repressor, *Plzf*, is considering involved in stem cell maintenance. Loss of *Plzf* function shifts the balance between spermatogonial stem cell self-renewal and differentiation toward differentiation at the cost of self-renewal and leads to an increase of post-meiosis apoptotic cells (Buaas et al，2004；Costoya et al，2004). It is shown that *Plzf* directly represses the transcription of Kit (Filipponi et al，2007). *Nanos* encodes for a zinc-finger RNA-binding protein and shows a translational repression activity requiring the interaction with the ubiquitously expressed protein Pumilio. The

Nanos-Pumilio protein complex binds to the nanos-responsive element (NRE) in the 3′ UTR of target mRNAs and represses their translation (Sonoda and Wharton, 1999). It has been indicated that *Nanos3* is required to prevent PGCs from undergoing apoptosis during migration (Suzuki et al, 2008). Over-expression of *Nanos3* causes an increase of the G1 stage undifferentiated spermatogonia. RA significantly decreases the expression of *Nanos3* in the undifferentiated spermatogonia (Asaoka-Taguchi et al, 1999; Lolicato et al, 2008). Therefore, *Nanos3* is important for maintaining the undifferentiated stage of spermatogonia. *Nanos2* suppresses meiosis by preventing stra8 expression. *Nanos2*$^{-/-}$ male PGCs go into apoptosis at 16.5 dpp and completely lost before birth (Tsuda et al, 2003). *Oct3/4* is the stem cell and germ line specific marker encoding DNA binding domain POU and it is also called *Pou5f1* (Scholer et al, 1991). When RA binds to RARs, expression of *Oct3/4* is inhibited. In the germ cell-specific nulls of Pou5f1, XX and XY germ cells undergo apoptosis between 9.5-10.5 dpc, before colonization of the gonad (Kehler et al, 2004).

Germ cells in the testis enter mitotic arrest in G0 until near birth (Durcova-Hills and Capel, 2008). At this stage, low level expression of meiosis-associated genes, such as *Sycp3* and *Dmc1*, indicates that all these germ cells are capable to enter meiosis (Di et al, 2000). The reason for them to enter the mitotic arrest is hypothesized that developing testicular tissue produces a meiosis-inhibiting factor (McLaren, 1984; Francavilla and Zamboni, 1985; Buehr et al, 1993). *Cyp26b1* seems to be one of the meiosis-inhibiting factors by distinguishing the inducer RA. The switching on/off of meiosis entrance is *Cyp26b1* which regulates the amount of RA in

the prenatal gonads (Trautmann et al，2008). Germ cells fail to enter such arrest forms teratomas in the Dnd1 mutant male mice (Cook et al，2011). The cut-off of *c-kit* expression in these pro-spermatogonia is important for prevent cancer formation and the afterward normal spermatogenesis as mutations of *c-kit* (result in self-activation of Kit) often occurs at this post-migration stage which will cause bilateral testicular germ cell tumors (Mol et al，2003；Biermann et al，2007).

A summary of *c-kit* upstream/downstream signals during spermatogenesis and testes tumors is shown on Figure 1.12 (page 36).

1.7 Hypothesis and aim

Based on the early experiments about *c-kit* during spermatogenesis, differences of *c-kit* mRNA and protein expression profiles in the spermatogonial stem cells before and after differentiation, as well as their upstream control factors are unclear. Our hypothesis is that the expression profiles of *c-kit* in SSCs are dynamically changed before and after differentiation; the change is controlled by some core regulatory factors, such as RA.

The aim of this study is to further explore the expression profiles of *c-kit* during SSCs differentiation and find out the core *c-kit* expression regulatory factors.

To accomplish this aim, the first problem we meet is that there is lack of effective methods to obtain and maintain a pure population of SSCs and differentiating spermatogonia *in vitro*. Therefore, in the first part of our study, we will use a previously established mouse SSCs cell line (c18-4), a mouse differentiating spermatogonia cell line (CRL-2053) and different age mouse testes to elucidate the

expression profiles of *c-kit* in SSCs before and after differentiation. The second problem is that *c-kit* involved SSCs differentiation controlling factors net works are not very clear. Therefore, after the expression profile of *c-kit* has been established, we will use an *in vitro* SSCs differentiation induction model to validate the obtained expression profile and at the mean time, set *c-kit* as the core clue to further study the interaction between *c-kit* and other important SSCs differentiation regulatory genes.

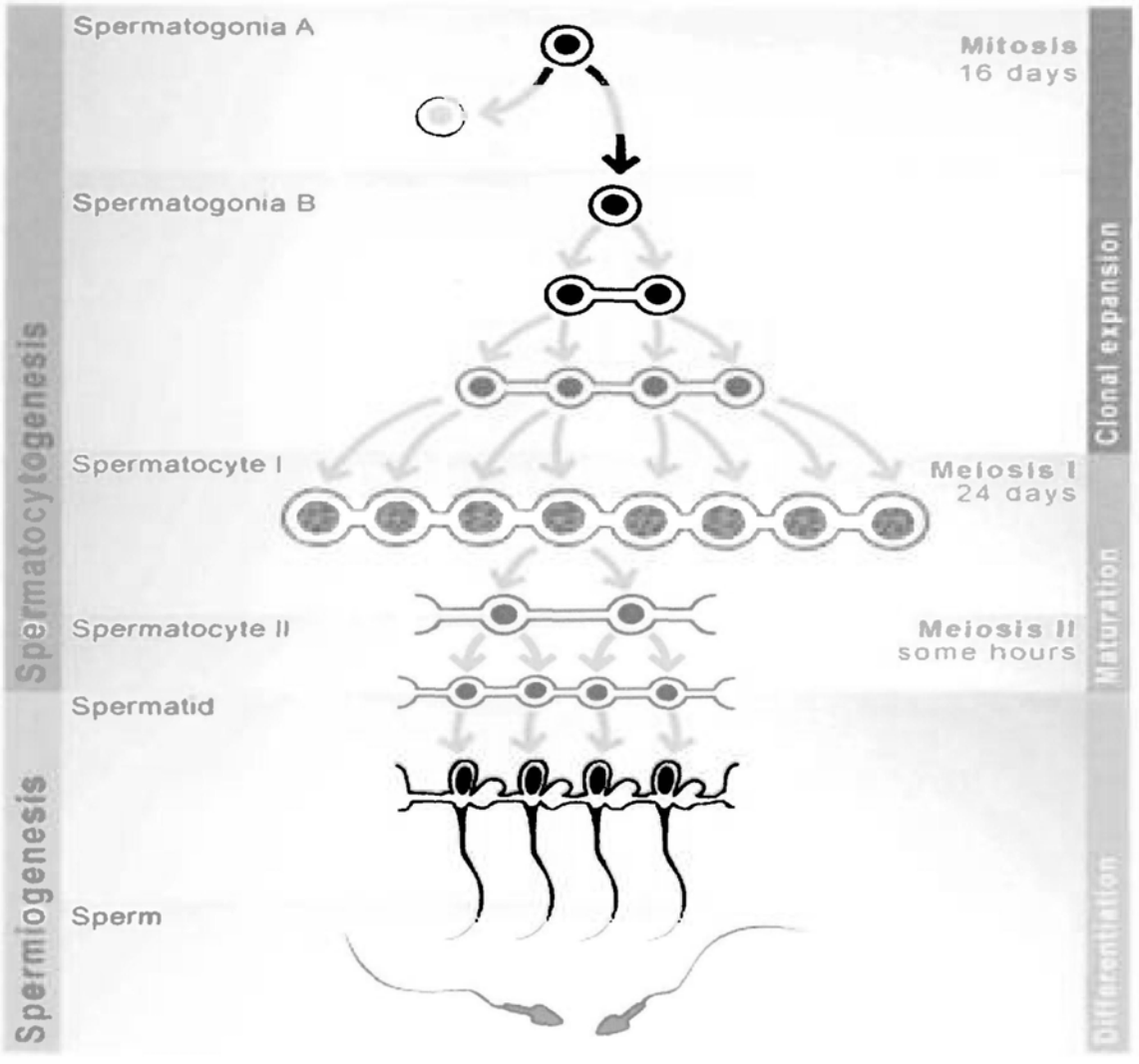


Figure 1.1 The human spermatogenesis generations.

The stem cell population of the germinal cells lies on the basal lamina of the convoluted seminiferous tubules. These are Type A spermatogonia. These cells undergo mitosis: one of the daughter cells renews the stock of type A spermatogonia, the other becomes a type B spermatogonia. The type B spermatogonia keep dividing and their daughter cells migrate towards the lumen. In roughly 64 days, the functional spermatozoa locating at the outer surface of the epithelium are generated .

Adapted from http://www.embryology.ch/anglais/cgametogen/spermato03.html.

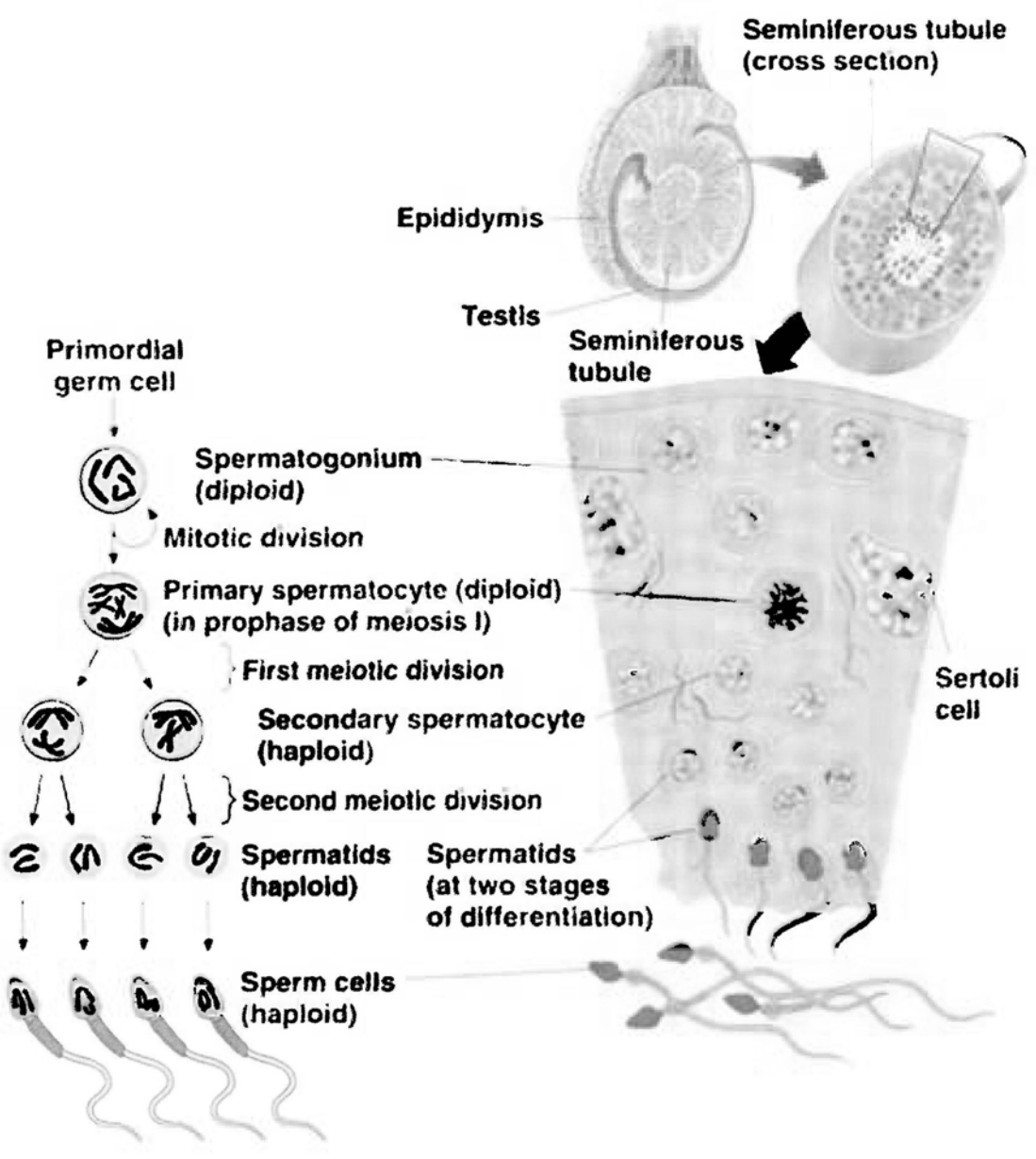


Figure 1.2 Schematic diagram of spermatogenesis. Spermatogonia in the seminiferous tubules of the testis give rise to diploid primary spermatocytes, which undergo meiosis (via secondary spermatocytes) to form haploid spermatids. Spermatids develop into mature spermatozoa. The germ cells are nurtured by growth factors secreted from the Sertoli cells.

Adapted from

http://iceteazegeg.wordpress.com/2009/02/25/gametogenesis/spermatogenesis/

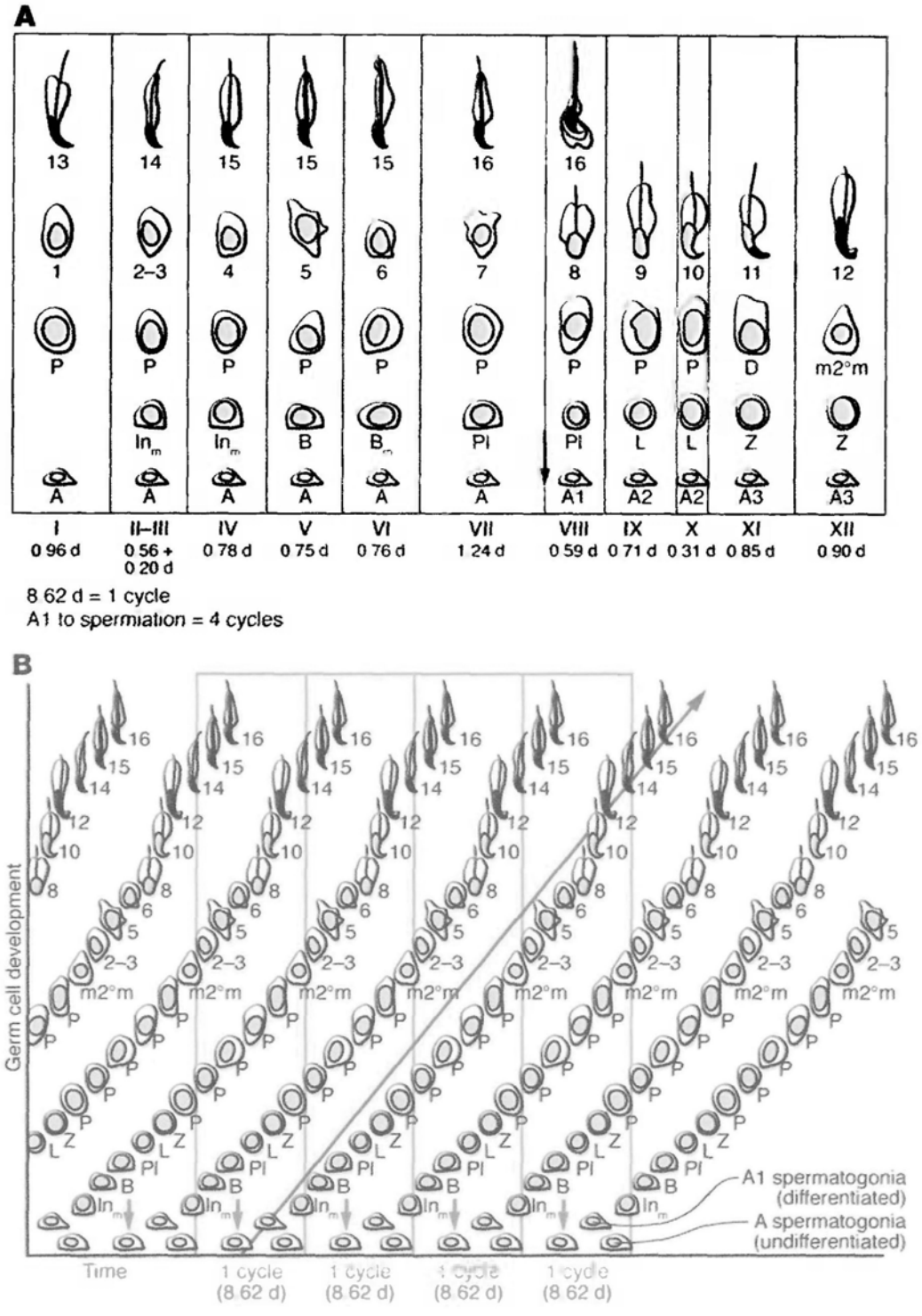


Figure 1.3 The cycle of the seminiferous epithelium. (A) Standard depiction of the cycle of the seminiferous epithelium for the mouse testis. The variable distances between the stages of the cycle are proportional to the duration of each of these cellular associations. The red arrow indicates the time in the cycle when vitamin A is required for the commitment to meiosis. (B) Depiction of how the cycle is generated. Spermatogonia undergo mitotic expansion. Stimulated by vitamin A (in the form of

RA) (red arrows), they initiate meiosis and ultimately become spermatozoa. The time required for this process from the time of the onset of meiosis to the formation of spermatozoa is particular to the species and the germ cells themselves (blue arrow). The periodic initiation of the differentiation process by vitamin A generates the cellular associations that define the cycle in A. Inm, intermediate (mitosis); B, B spermatogonia; Pl, preleptotene spermatocytes; L, leptotene spermatocytes; Z, zygotene spermatocytes; P, pachytene spermatocytes; D, diplotene spermatocytes; m2°m, secondary spermatocytes. Round and elongating spermatids are labeled as steps 2-3, 8, 12, 16. Adapted from (Hogarth and Griswold，2010).

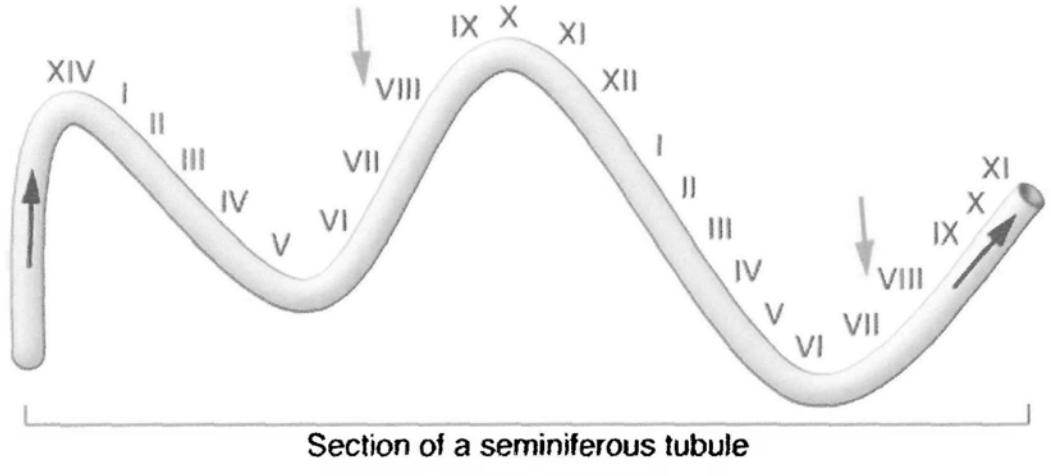


Figure 1.4 Depiction of the spermatogenic wave. A single seminiferous tubule is depicted, and the stages of the cycle (cellular associations) along the tubule are shown. The spermatogenic wave describes the process in space, while the cycle of the seminiferous epithelium refers to the process in time. The point of meiotic initiation (red arrows) moves along the tubule in the direction of the black arrows. The net result of the wave is the asynchronous (and therefore continual) release of spermatozoa. Adapted from (Hogarth and Griswold，2010).

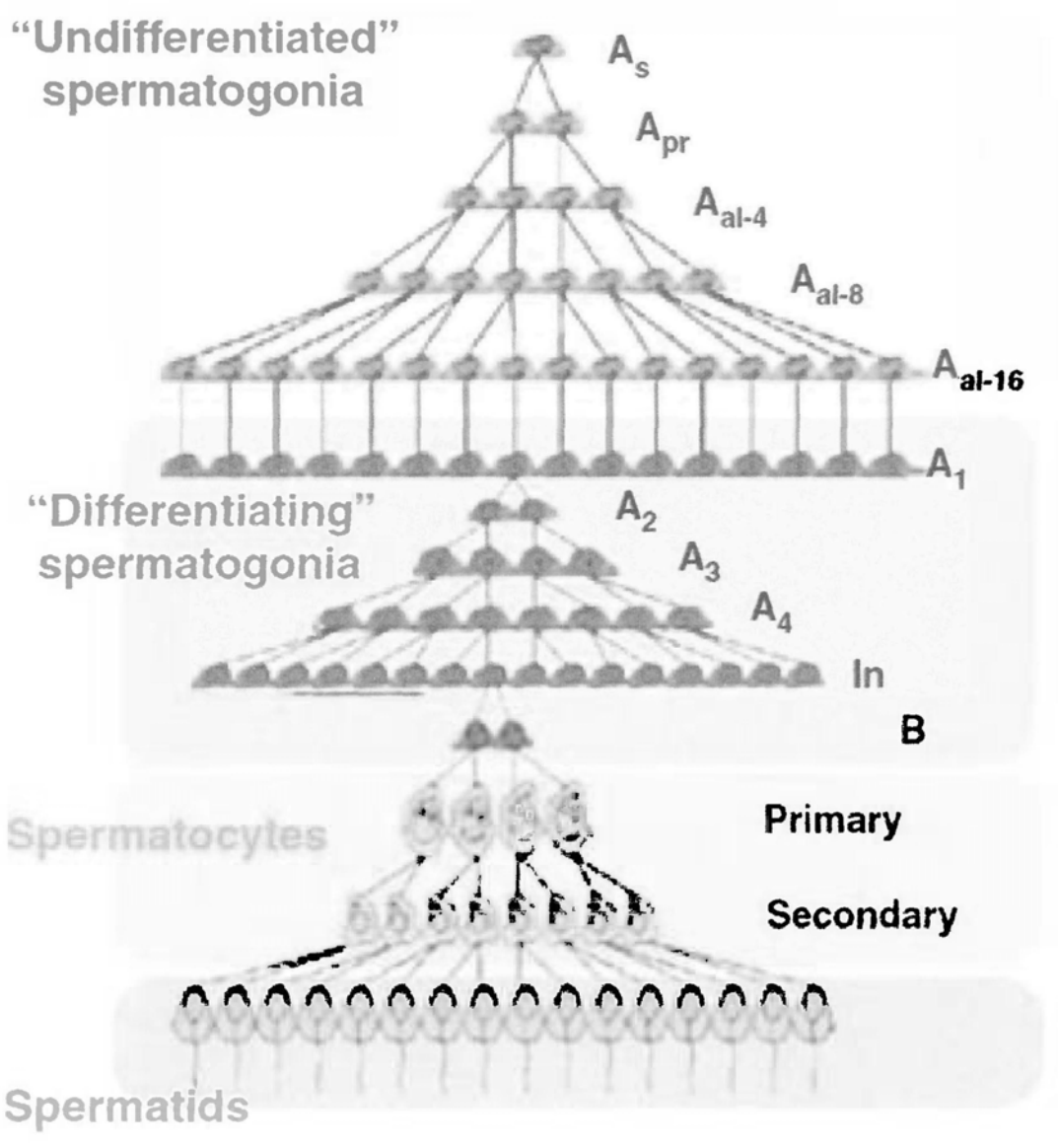


Figure 1.5 The organized cell populations corresponding topographically to the stages of spermatogenesis. The A_s spermatogonia is considered to be the most primitive cell. In mice, a A_s spermatogonium undergoes 10 successive mitotic divisions and theoretically produces 1024 primary spermatocytes are believed to enter meiosis in a synchronized manner. Adapted from Develop. Growth Differ. (2010) 52, 311–317

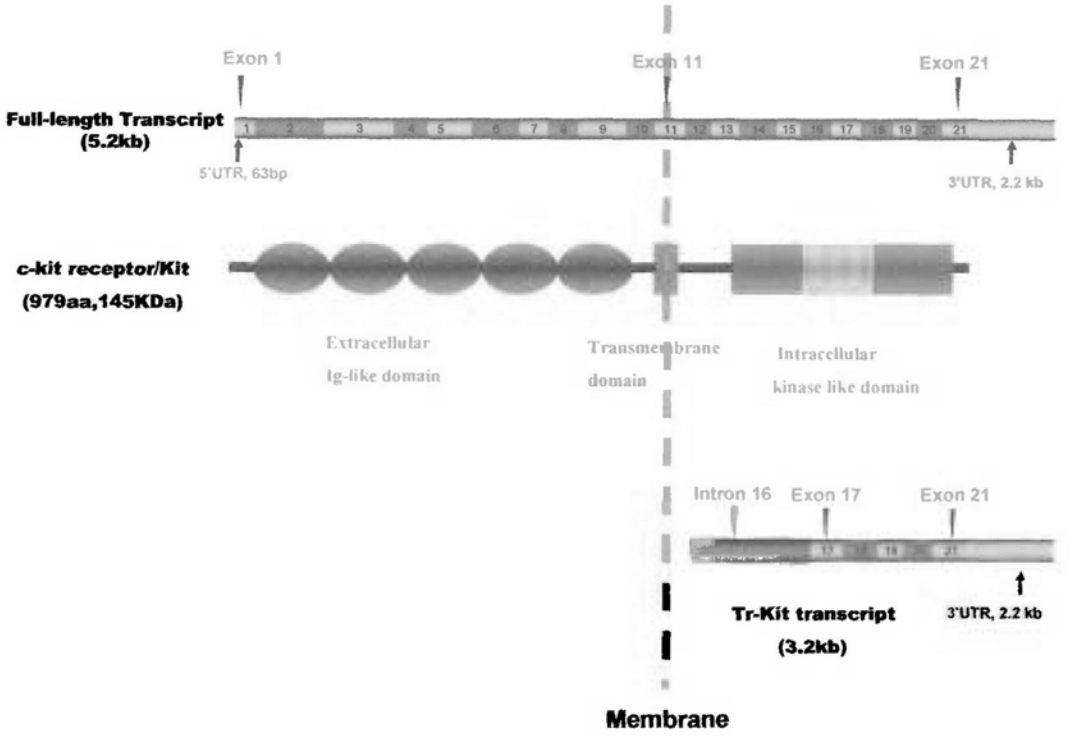


Figure 1.6 *c-kit* mRNA and protein structure. The full-length *c-kit* mRNA is 5150 bp in length, consists of 21 exons and encodes a 145 KDa protein with 979 amino acids. Exons 1-10 encode the extracelluar components with 5 Ig-like domains and the signal sequences. Exon 11 encodes the transmembrane segment. Exons 12-21 encode the intercellular segments including the juxtamembrane segment, proximal kinase domain, kinase insert domain and distal kinase domain.

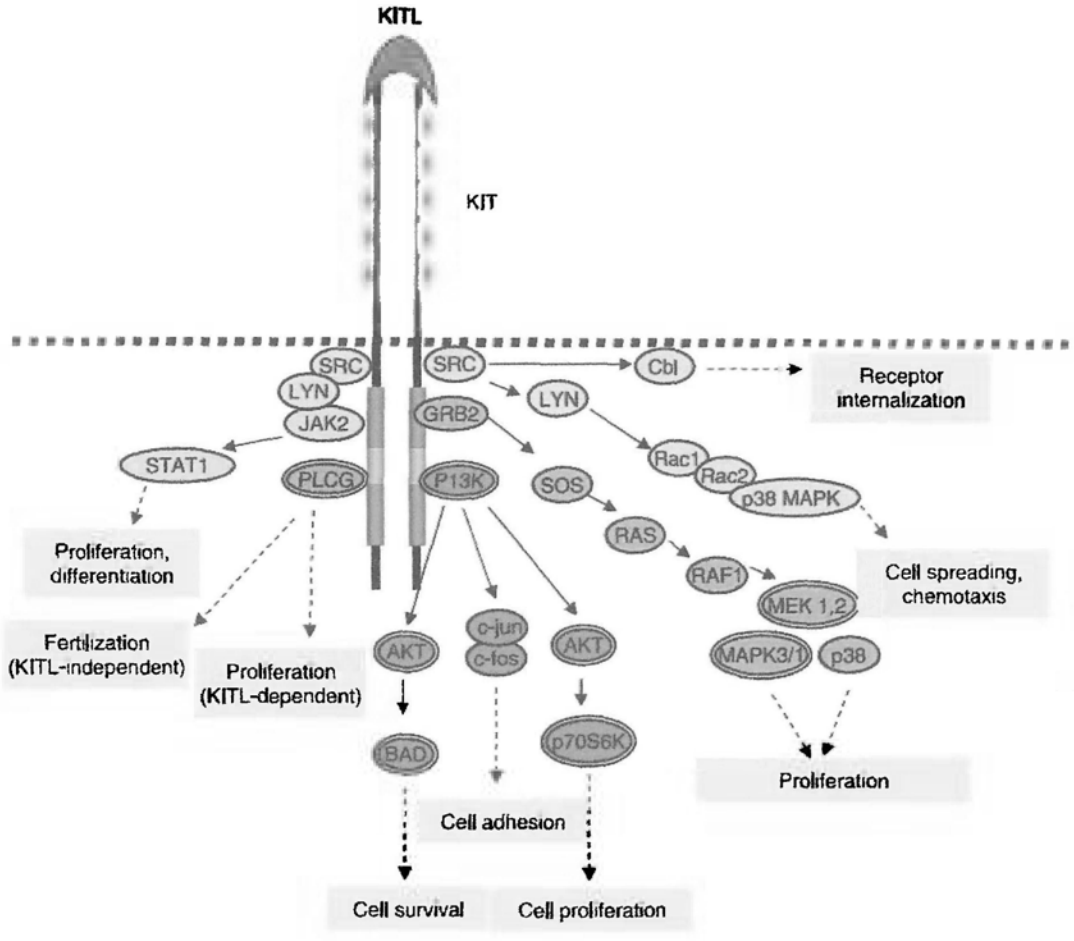


Figure 1.7 Signalling pathways and downstream cellular functions of activated Kit/kitl. Four pathways are known to be activated in response to Kitl stimulation. The PI3K pathway results in cell survival, adhesion and proliferation (blue circles). The SRC kinase pathway is required for receptor internalization, chemotaxis and proliferation (purple circles). The MAP kinase pathway (green) mediates gene transcription and proliferation; JAK/STAT (yellow) signalling is essential for proliferation and differentiation. The PLCG pathway (orange) mediates Kitl-independent sperm–egg fertilization and Kitl-dependent proliferation in haematopoietic cells. Double-line circles represent the pathways that are active during spermatogenesis Adapted from (Mithraprabhu and Loveland，2009).

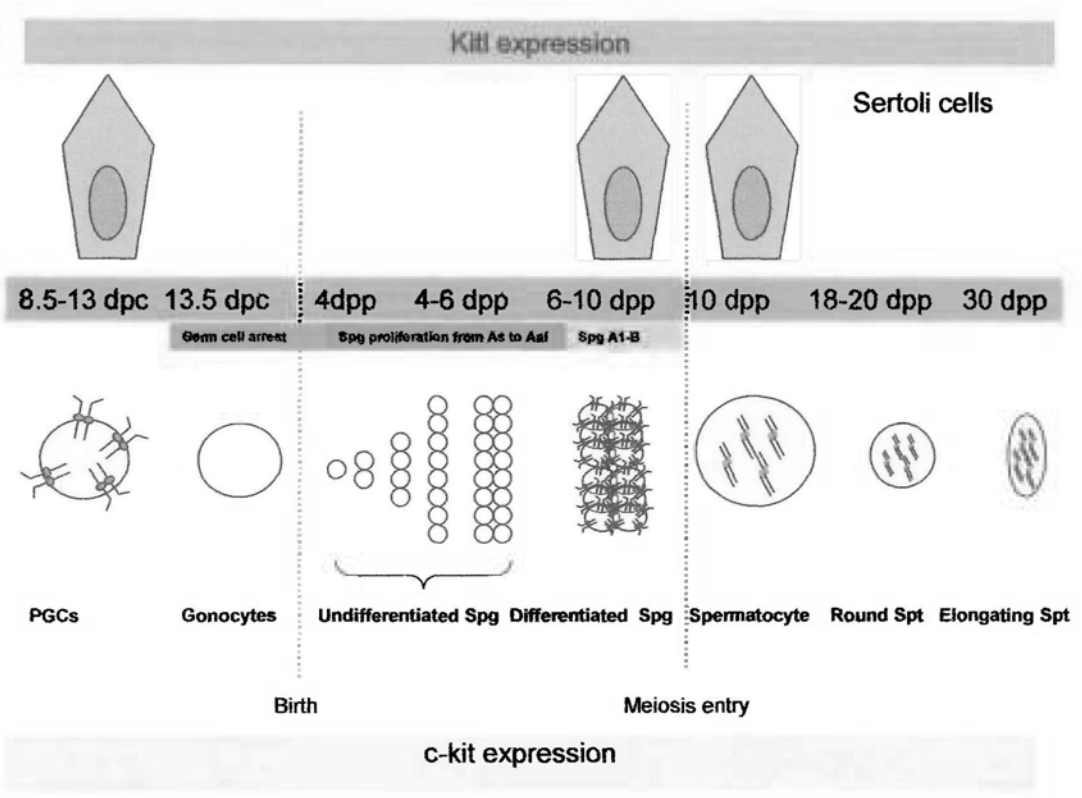


Figure 1.8 Expression of Kit during spermatogenesis in mice. PGCs from 8.5-13 dpc express *c-kit*. From 13.5 dpc to 3 dpp, the *c-kit*$^-$ gonocytes are arrested. The A_s to A_{al} staged spermatogonia, which are undifferentiated, do not express *c-kit*. Expression of *c-kit* is restarted at around 7 dpp from the A1-B stage Spermatogonia. Truncated Kit keeps expressed in the spermatocytes and later stage germ cells (spermatocytes, spermatids and spermatozoa). Sertoli cells express Kitl from 8.5 dpc to 13.5 dpc and 6-10 dpp only. Entry of meiosis occurs around 10 dpp. Spg represents spermatogonia; Spt represents spermatids.

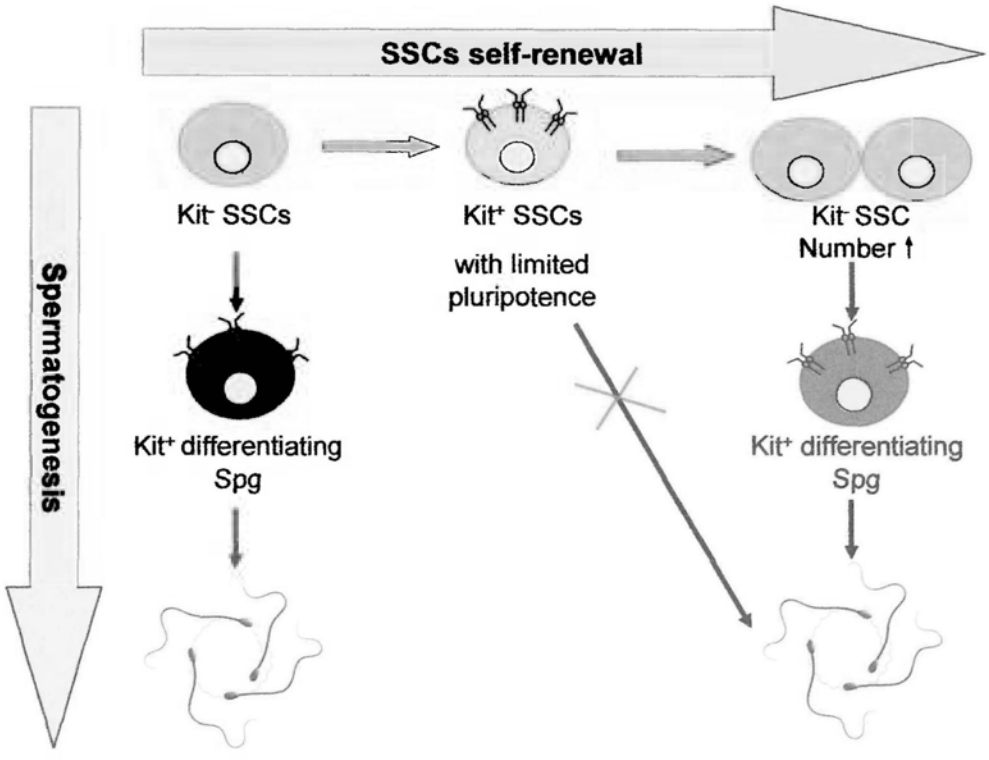


Figure 1.9 New discoveries on *c-kit* and SSCs self-renewal/differentiation. There are 2 kinds of SSCs including the Kit^+ and Kit^- SSCs in the early stage of postnatal spermatogenesis in mice. The Kit^- SSCs is the performer of normal spermatogenesis. In order to complete self-renewal, Kit^- SSCs change their phenotype to Kit^+ SSCs according to their microenvironment. The Kit^+ SSCs have limited pluripotence and cannot regenerate spermatogenesis until transformed into Kit^- SSCs.

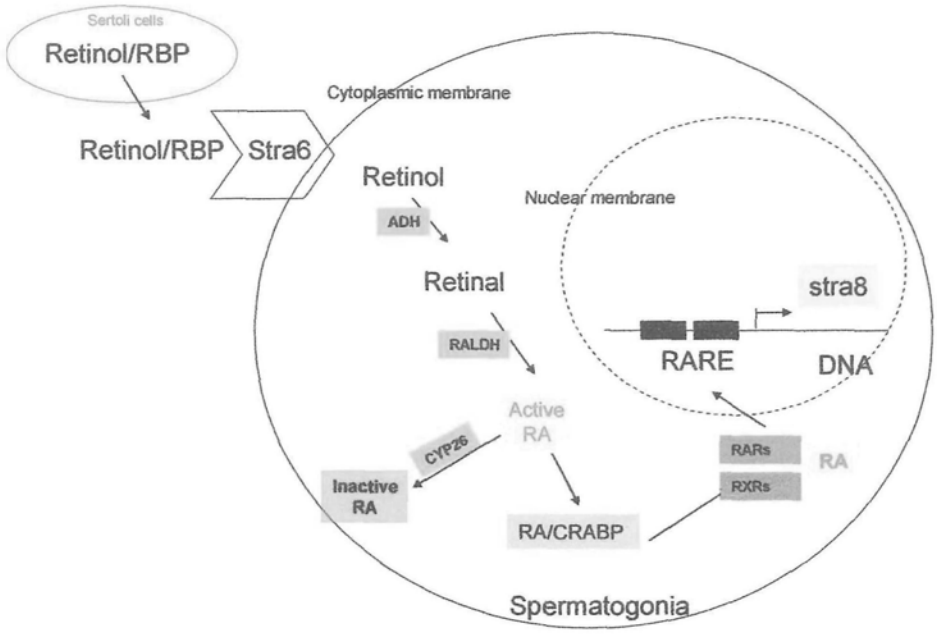


Figure 1.10 RA and its derivatives in the Spermatogonia. Retinol is delivered to germ cells on the retinol-binding protein (RBP) and is internalized via the membrane receptor STRA6. Retinol is catalyzed by alcohol dehydrogenase to retinal. Retinal is irreversibly oxidezed by a family of retinaldehyde dehydrogenases (RALDH1A1, RALDH1A2 and RALDH1A3) form RA. Inside the cell, RA is bound up by an excess of cellular retinoic acid-binding protein (CRABP). In the nucleus of the spermatogonia, RA binds to the RARs and the activated receptor can work on the RARE region and stimulate transcription of a number of genes including *Stra8*. A family of cytochrome P450 enzymes (Cyp26a1, Cyp26b1 and Cyp26c1) converts excess RA to 4-oxo and 4-hydroxy forms which are excreted from the cell (Niederreither and Dolle，2008). RA acts by binding to nuclear RA receptors (RARs) and heterodimerize with retinoid X receptors (RXRs) in the cytoplasm. After binding to RA, the bounded RAR-RXR dimmers move to the nucleus and bind to the RA-response elements (RAREs) and thereby control the expression of RA-response genes (Chambon，1996；Mark et al，2006).

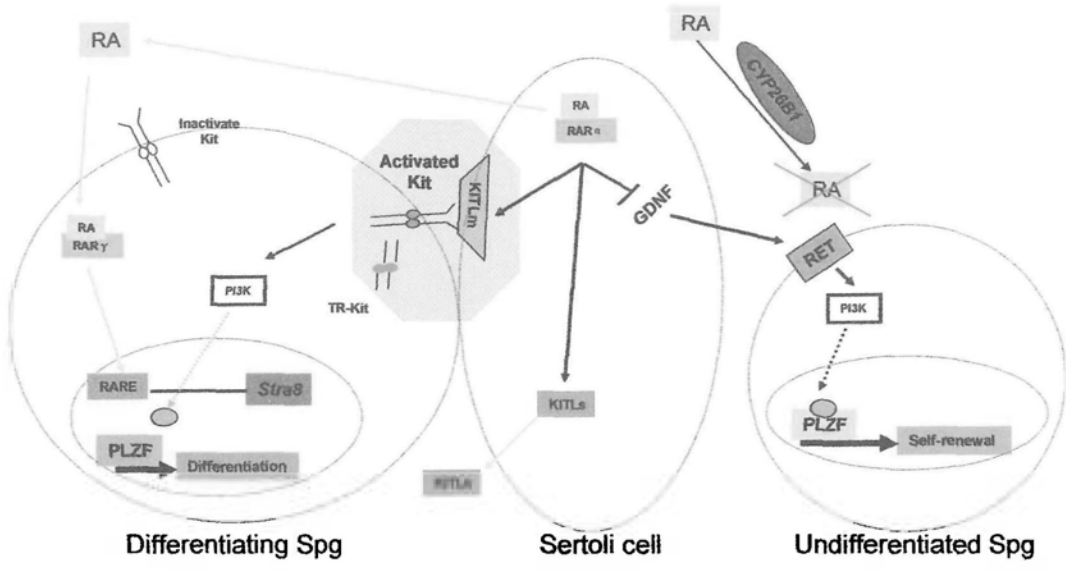


Figure 1.11 RA pathways affect on spermatogonia self-renewal and differentiation. Both Kit and Kitl consist of the membrane and soluble forms represented by Kitm, Tr-Kit, Kitlm and Kitls respectively. By binding to RARα, RA stimulates the Sertoli cells to synthesize Kitl, which will bind to Kit and will activate the Kit/Kitl pathway and block the GDNF pathway. Soluble RA may also bind to the RARγ in the the differentiating spermatogenic cells and act on RARE to stimulate meiosis genes like *Stra8*. In contrast, the RA is distinguished by Cyp26b1 before reaching the undifferentiated spermatogonia. Whether RA directly activates Kit/Kitl cascade or not is still questionable. In the differentiating Spermatogonia, the phosphorylation state of the PLZF associated corepressors is altered by activation of Kit/Kitl/PI3K pathway followed by the inactivation of the corepressors. As a result, meiosis is triggered. In the undifferentiated spermatogonia, GDNF/Ret/PI3K pathway works to allow PLZF interacts with its corepressors to exert its repressing activity. As a result, meiosis is prohibited.

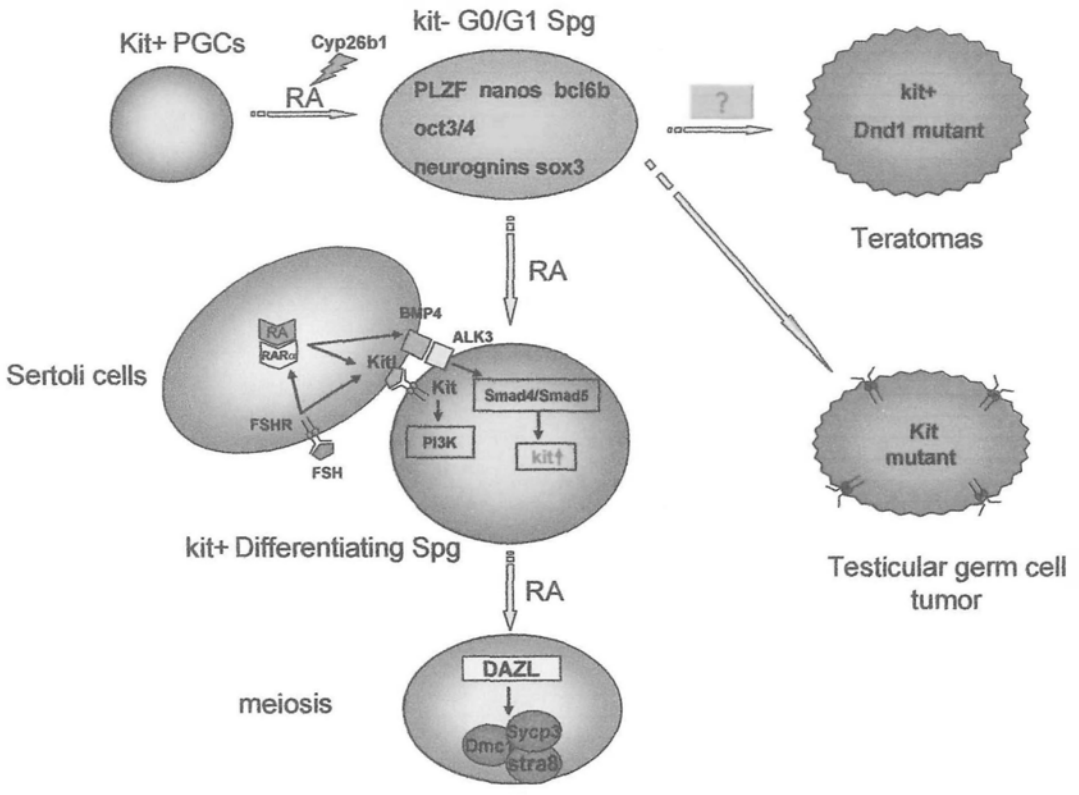


Figure 1.12 Summary of Spermatogenesis and *c-kit* upstream/downstream signals. The undifferentiated spermatogonia highly express *Plzf, Nanos, Bcl6b, Oct3/4 (pou5f1), Neurogenin3* and *Sox3*. The *c-kit*$^{+}$ PGCs gradually lose the expression of *c-kit* and go into arrest when the induction signal RA is distinguished by Cyp26b1. With RA stimulation, the arrested spermatogonia can resume normal spermatogenesis. Gain of *c-kit* expression accidentally during the arrest stage will lead to formation of testis teratomas. RA, FSH and BMP4 congenerously activate the Kit/Kitl signal pathway in the spermatogonia which will activate expression of *Stra8, Dmc1* and *Sycp3* and finally lead to meiosis.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Cell lines

Mouse SSCs cell line (c18-4) was previous established by Marie-Claude Hofmann et al. c18-4 was immortalized type A spermatogonia using the simian virus large T-antigen gene (LTAg) under the control of an ecdysone-inducible promoter. The cell line exhibited typical morphological features of spermatogonia at the light microscopic level. Moreover, the cells expressed detectable levels of germ cell markers such as *Dazl*, germ line stem cells markers such as *Oct-4, GFRα-1,* and GDNF, and stem cell markers such as *Piwil2* and *Prame11*. The c18-4 cell line represents a good *in vitro* model for studying mouse germ line stem cell biology (Hofmann et al，2005).

CRL-2053 (ATCC) is a type B spermatogonia cell line that was immortalized by transfection with pSV3-neo (a plasmid containing coding sequences for the SV40 large T antigen and neomycin resistance). The cell line shows characteristics of a stage between type B spermatogonia and primary spermatocytes. The cells express two testis specific isoproteins, cytochrome c and lactate dehydrogenase C4 (Hofmann et al，1992).

CRL-2196 (GC-2spd(ts), ATCC) is a spermatocyte cell line which was established by stable co-transfection of freshly isolated spermatocytes with the SV40 large T antigen gene (pSV3neo) and a temperature sensitive mutant of the p53 tumor

suppressor gene (LTRp53cG9). Cells were selected with G-418, cultivated for 6 months and single cell cloned three times by limiting dilution. No clonal proliferation was observed in soft agar cultures, indicating that these cells were immortalized but not transformed. The cells have lost their differentiation potential, and are currently arrested at a premeiotic stage (Hofmann et al，1994, 1995).

CRL-1825 (P19) cell line was a line of pluripotent embryonal carcinoma able to grow continuously in serum-supplemented media (McBurney，1993). The differentiation of these cells can be controlled by nontoxic drugs. RA effectively induces the development of neurons, astroglia and microglia - cell types normally derived from the neuroectoderm. The concentration of RA determines the differentiated cell types formed (Edwards and McBurney，1983).

2.1.2 Animals

C57/BL6 mice at different age were purchased from laboratory animal service center (LASEC), CUHK. All procedures were approved by the Animal Research Ethics Committee of the university.

2.2 Methods

2.2.1 Cell culture

All the cells were cultured in the Dubecco Modified Eagle Medium/F12 (DMEM/F12, Invitrogen, Catlog No. 11330-032, USA) supplemented with 10% fetal bovine serum (FBS, Invitrogen, catlog No. 10099-141, USA) under 37℃, 5% carbon

dioxide (CO_2) and saturated humidity. A subcultivation ratio of 1:6 to 1:10 was applied. Media were renewed 1 to 2 times per week. The cells were frozen in the complete growth medium supplemented with 5% (v/v) DMSO and stored in the liquid nitrogen.

2.2.2 Mouse testes collection

5 days post partum (dpp), 10 dpp and 60 dpp C57/BL6 were sacrificed by cervical dislocation. Testes from both sides were harvested and pooled according to age. For RNA extraction, testes were washed twice with phosphate buffered saline (PBS) and then immersed in "RNA-later" stabilization reagent (Qiagen, USA). For protein extraction, testes were washed twice with PBS, transported in ice box and store in -80℃. Three batches of animals were used for each experiment.

2.2.3 *In vitro* tissue culture inducing with RA

In vitro tissue culture was carried out according to the methods described by previous study (Zhou et al，2008). Testes from 5 dpp, 10 dpp and 60 dpp mice were detunicated, cut into small pieces per testis, placed on Millicell CM filters (Millipore, Bedford, MA) floating on the surface of medium and covered with drops of medium (DMEM/F12+10% FBS). RA (Sigma-Aldrich Co., Saint Louis, MO, U.S.A) diluted in ethanol was added to the culture medium to make a final concentration of 0.7 μM or 2 μM. Tissues were harvested after 24 hours of RA treatment. Total RNA was isolated using the RNeasy mini kit (Qiagen, USA).

2.2.4 RA induction of the *in vitro* cultured germ cells

For germ cell induction assay, 2×10^6 c18-4 or CRL-2053 cells were pre-seeded into T25 cell culture flasks separately (2 flasks each group) overnight before the treatment in full medium (DMEM/F12+10% FBS). Induction media (DMEM/F12+10% FBS) with a final concentration of 2 μM RA dissolved in ethanol were used in the treatment (induction) group. No RA medium which containing the same amount of ethanol was set up as control group. After 24 hours of induction, removed the induction media and washed the cells with PBS twice, and collected the cells and stored at -80℃ until analysis. Three independent replications were carried out for each experiment.

2.2.5 *In vitro* induction of differentiation with RA on P19 cells

Methods for RA stock solution and *in vitro* cultured P19 cells with RA were described by previous studies (Jones-Villeneuve et al，1982). RA was prepared as a stock solution at 5 mM in 100% ethanol. The stock solution was diluted into the culture medium to obtain a desired working concentration. 7 different RA concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1 μM, 2μM, 4μM) were used. Media were changed every day and RA treated P19 cells were collected on 3 time-point which was 1 day, 3 days and 5 days after treatment respectively. Ethanol only treatment was set up as control. 3 independent replications were performed for each experiment.

2.2.6 Methods for RNA preparation, electrophoresis and Northern blot

Total RNA from cells and testes was isolated using the RNeasy mini kit (Qiagen, USA) following the manufacturer's instructions. RNA concentrations and purity were determined by absorbance at 260 and 280 nm with BioPhotometer (Eppendorf, USA). OD260/OD280≈1.8-2.0 and OD260/230>2 were considered to be acceptable. The procedures used for electrophoresis and Northern blot were performed as previously described (Sambrook and Russell, 2001). RNA size was estimated by comparing with 2 μg RNA Millennium size markers (Ambion, USA) by measuring the distance from each band to the loading well.

DNA fragments corresponding to exons 10-12 and exons 18-20 of the full-length *c-kit* transcript were obtained by PCR with *c-kit* specific primers using the 60 dpp mouse testis cDNA as template. Primers against exons 10-12: sense 5′ -TGGGGATCATTGTGATGGT-3′, anti-sense 5′-ATGGCAGCATCCGACTTAAT-3′; primers against exons 18-20: sense 5′ - CCTCTGGGAGCTCTTCTCCT-3′, anti-sense 5′- GCTGTCCGAGATCTGCTTCT-3′. Amplified DNA fragments were inserted into the Topo-TA vector (Invitrogen, USA). The plasmids were extracted by QIAprep spin miniprep kit (Qiagen, USA) and were sent to commercial company for sequencing.

RNA probes were prepared by MAXIscript kit (Ambion, USA) following the manufacturer's instructions. mRNA-complementary (antisense) transcripts are synthesized in a 20μl *in vitro* transcription system containing 1 μg DNA template, 2

μl 10×transcription buffer, 1 μl 10 mM ATP, 1 μl 10 mM CTP, 1 μl 10 mM GTP, 5 μl 800 Ci/mmol [α-32P]UTP at a concentration of 10 mCi/mL (Perkinelmer, USA) and 2 μl T3 enzyme mix. After purification with NucAway Spin columns (Ambion, USA), the RNA probes were hybridized with the blots with RNA samples in the ULTRAhyb ultrasensitive hybridization buffer (Ambion, USA) at 68 °C overnight. The same blot was stripped and re-probed with α^{32}P-labeled beta-actin RNA probe as internal control. Northern hybridization was performed twice with probes and membranes that were made independently. The sequence of RNA probes were showed in Table 2.1.

Table 2.1 RNA probe sequence of mouse *c-kit* gene

Probe name	Probe sequence
exons 10-12 probe	1 ATGGCAGCAT CCGACTTAAT CAAGCCATAT GCAGTGGCCT CAACGACCTT 51 CCCGAAGGCA CCAGCTCCCA ATGTCTTTCC AAAACTCAGC CTGTTTCTGG 101 GAAACTCCCA TTTGTGATCA TAAGGAAGTT GCGTCGGGTC TATGTAAACA 151 TAATTGTTTC CATTTATCTC CTCGACAACC TTCCATTGTA CTTCATACAT 201 GGGTTTCTGC AAATATTTGT AGGTGAGCAC CATCACAATG ATCCCCAT
exons 18-20 probe	1 GCTGTCCGAG ATCTGCTTCT CAATAAGTTG GACAACCTGC TTGAATGTTG 51 GCCTTTTCAA GGGGTCAGCG TCCCAGCAAG TCTTCATGAC GTCATACATT 101 TCGGCAGGCG CGTGCTCCGG GCTGACCATC CGGAAGCCTT CCTTGATCAT 151 CTTGTAGAAC TTGGAGTCGA CCGGCATCCC TGGGTAGGGG CTGCTTCCTA 201 AGGAGAAGAG CTCCCAGAGG

Exons 10-12 probe hybridizes to *c-kit* extracellular domain coding area. Exons 18-20 probe hybridizes to *c-kit* intracellular domain coding area.

2.2.7 Rapid amplification of cDNA ends (RACE), cloning and sequencing

The number and size of *c-kit* mRNA expressed in mice cell lines and testis were determined by Northern blot, the existence of these trancripts were further confirmed by RACE and sequencing. We used the BD-Smarter RACE protocol from BD Biosciences Clontech (PaloAlto, CA) in RACE analysis. The full-length cDNAs was made by joint action of the SMARTer II A Oligonucleotide and SMARTScribe Reverse Transcriptase (a variant of MMLV RT) in reverse transcription reactions. First strand cDNA synthesis was obtained from 1 μg total RNA. PCR amplification was done with specific primers hit exons 10-12 and exons 18-21 on the full-length *c-kit* transcript (Table 2.2 and Figure 2.1) in conjunction with universal primers which were provided by the kit. Advantage 2 PCR kit (BD Biosciences Clontech, USA) was used for the 5′ and 3′ PCR amplification. Nested PCR and touchdown PCR were used to make sure the specificity of the amplification. Conditions for primary and nested PCR were: 3 min at 94 °C; 5 cycles of 30 s at 94 °C, 4 min at 72 °C; 5 cycles of 30 s at 94 °C, 30 s at 70 °C, and 4 min at 72 °C; 20 cycles of 30 s at 94 °C, 30 s at 68 °C, and 4 min at 72 °C; final extension, 7 min at 72 °C. The products of the first round PCR were diluted 100 times with ddH_2O and nested PCR were performed with *c-kit* gene specific nested primers under the same PCR conditions of the first round. Electrophoresis of the PCR products, bands cutting and gel extraction (QIAquick gel extraction kit, QIAGEN) were performed. All of the clear RACE PCR product gel extractions were cloned to TA vector (Invitrogen, TOPO TA cloning kit

for sequencing) and sent to commercial company for sequencing. 5′ and 3′ RACE results were combined to obtain the full-length *c-kit* transcripts sequence information.

Table 2.2 RACE primers

Name	5′ or 3′	sequence(5′→3′)	No. of bases	exons hitting	position on NM_021099
e11 5′	5′	CAGCCTGTTTCTGGG AAACTCCCATTTG	28	exon 11	1825-1798
e12 5′	5′	GCAACTGTCATGGC AGCATCCGACTT	26	exon 12	1920-1895
e18 5′	5′	TGCTCTCTGGTGCCA TCCACTTCAC	25	exon 18	2552-2756
e20 5′A	5′	GGTCAGCGTCCCAG CAAGTCTTCAT	25	exon 20	2786-2762
e20 5′B	5′	AAGGGGTCAGCGTC CCAGCAAGTCT	25	exon 20	2790-2766
e20 5′C	5′	TGCTTGGTGCTGTCC GAGATCTGCT	25	exon 20	2856-2832
e21 5′	5′	GGGGTTGCAGTTTG CCAAGTTGGAG	25	exon 21	2887-2863
e10 3′	3′	AAATCCAGGCCCAC ACTCTGTTCACG	26	exon 10	1602-1627
e11 3′	3′	TGGGAGTTTCCCAG AAACAGGCTGAG	26	exon 11	1802-1827
e18 3′	3′	CCGTGAAGTGGATG GCACCAGAGAG	25	exon 18	2550-2574
e19 3′A	3′	AGGAAGCAGCCCCT ACCCAGGGATG	25	exon 19	2650-2674
e19 3′B	3′	GGGATGCCGGTCGA CTCCAAGTTCT	25	exon 19	2669-2693
e20 3′A	3′	TGACCCCTTGAAAA GGCCAACATTCA	26	exon 20	2782-2807
e20 3′B	3′	GCAGATCTCGGACA GCACCAAGCAC	25	exon 20	2833-2857

Requirement of a good gene specific primer for RACE: It should be 23-28 nt, has 50-70% GC and Tm > 70 °C, and does not complement to the 3′ of the Universal Primer Mix.

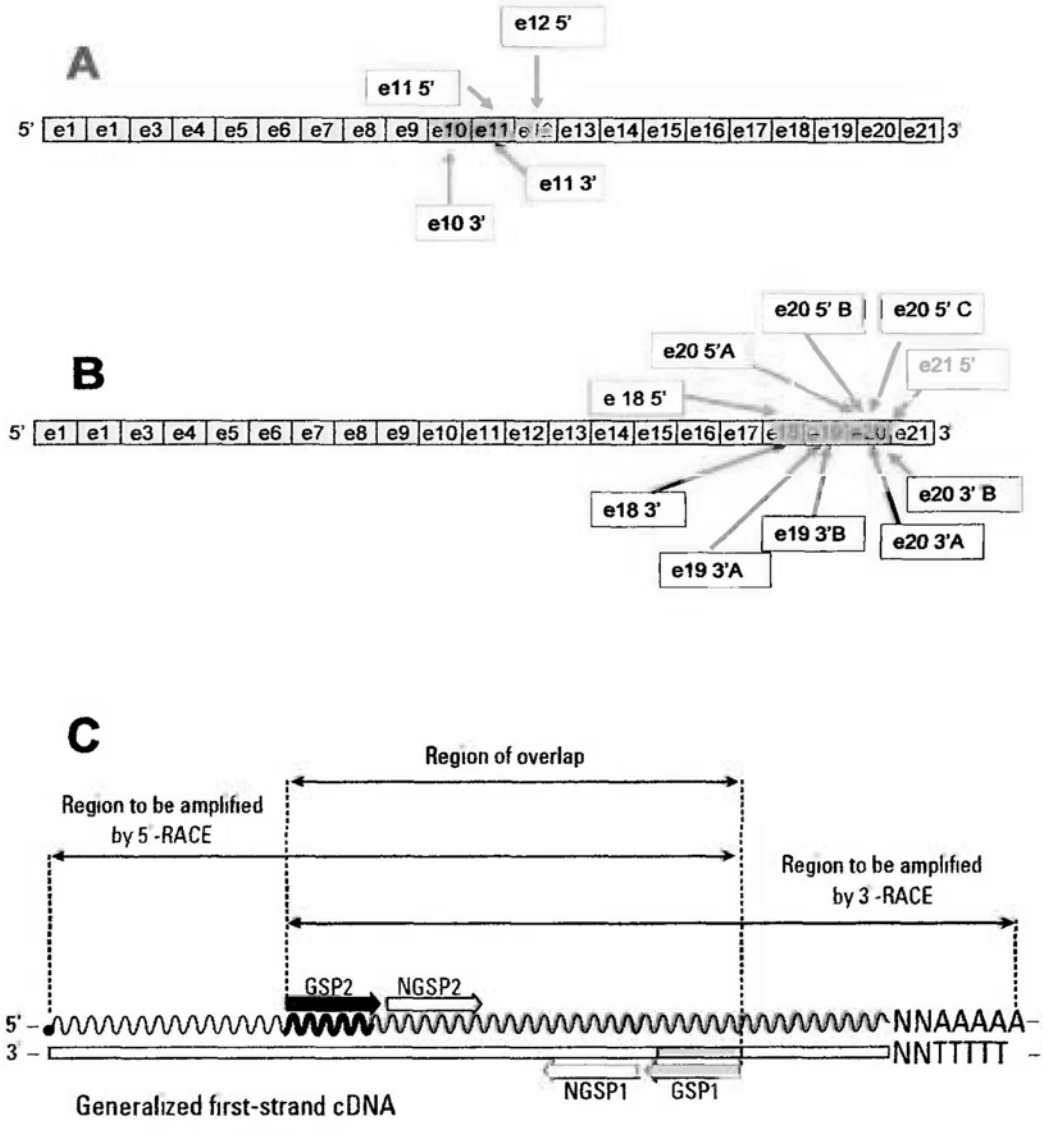


Figure 2.1 RACE primers binding position on *c-kit* full-length transcript and relationship of gene-specific primers to the cDNA templates. *c-kit* gene specific primers were designed corresponding to Northern hybridization probing area on the full-length transcripts. (A) Primary and nested primers hit exon 10-12. e12 5′: 5′ primary primer (GSP1); e11 5′: 5′ nested primer (NGSP1); e10 3′: 3′ primary primer (GSP2); e10 3′: 3′nested primers (NGSP2). (B) Primary and nested primers hit exon 18-20. Different combination of these primers were applied to get gene specific products, the combination pattern was the same as exon 10-12 primers. (C) The relationship between gene-specific primers and the cDNA templates. Gene specific primers designed produce overlapping RACE products.

2.2.8 Quantitative Real-Time RT-PCR

Total RNA (2 μg) was treated with DNase I (Sigma, Saint Louis, USA) for 15 min at room temperature and then reverse transcribed by High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA).

Real-time RT-PCR analysis of *c-kit* were performed with Taqman universal PCR master mix and Taqman gene expression assays on the ABI Prism 7900HT Real Time PCR System, according to the manufacturer's instructions (Applied Biosystems). The relative expression level of each target gene was calculated by the comparative CT method and was normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression. Three *c-kit* gene specific probes that hit different part of full-length transcript were used. Binding regions of these probes were exon 7-8 (AACGTTTACGTGAACACAAAACCAG), exon 20-21 (GCACCAAGCACATTTACTCCAACTT) and exon 21 (CTGATATGTTGTCCAACTGTTGACA).

Real-time RT-PCR analysis for other genes were performed with Power SYBR PCR master mix and gene specific primers on the ABI Prism 7900HT Real Time PCR System, according to the manufacturer's instructions (Applied Biosystems, USA). The relative expression level of each target gene was calculated by the comparative CT method and was normalized to GAPDH expression. The primers of the candidate genes are list on Table 2.3.

Each RT-PCR analysis was repeated 3 times and each reading was normalized by the reading of GAPDH.

Table 2.3 Gene specific primers of the candidate genes

Name	Sequence	Direction
BMP4-F	TTCCTGGTAACCGAATGCTGA	Forward
BMP-R	CCTGAATCTCGGCGACTTTTT	Reverse
Cyp26b1-F	GCAAGATCCTACTGGGCGAAC	Forward
Cyp26b1-R	TTGGGCAGGTAGCTCTCAAGT	Reverse
DAZL-F	GTCCTTACATGTACCATTCTGTGAC	Forward
DAZL-R	GACTCCAACAAAACAGCAGACAA	Reverse
EGR 3-F	AGCTGAACTGGGCTGTGTCT	Forward
EGR 3-R	AATGGGGAGTGGGTATGTGA	Reverse
Kitl-F	TCTGCGGGAATCCTGTGACT	Forward
Kitl-R	TGGAAGATTTGCCACCAGTTT	Reverse
PLZF-F	GCAAGAACAGCGTCAAGACA	Forward
PLZF-R	TGGGATCACGTGAAGCTATG	Reverse
RARα-F	TCCGAAGAGATAGTACCCAGC	Forward
RARα-R	AAAGCAAGGCTTGTAGATGCG	Reverse
Stra8-F	GTTTCCTGCGTGTTCCACAAG	Forward
Stra8-R	CACCCGAGGCTCAAGCTTC	Reverse

2.2.9 Western blot

Cells and testis tissues were lysed on ice in RIPA buffer (50 mM Tris-HCl, pH 7.4; 150 mM NaCl, 1% NP40, 0.25% Na-deoxycholate, 1 mM PMSF) containing 1% freshly added protease inhibitors (protease inhibitor cocktail, Sigma, USA). The lysates were incubate on ice for 30 min and centrifugated at 16,000 g, 4 ℃ for 10 min. Protein concentration was determined by BCA protein assay kit (PIERCE,

Rockford U.S.A.). Protein electrophoresis and gel bolting were performed with NuPAGE electrophoresis system (Invitrogen, USA) following the manufacturer's instructions. 20 μg of total protein lysates were denatured in NuPAGE LDS (Lithium dodecyl sulfate) Sample Buffer, separated on a NuPAGE Novex Bis-Tris Gel, and semi-dry blotted onto PVDF membranes. Verified the efficiency of the protein transfer by Ponceau red staining. Membranes were destained and blocked with 5% (wt/vol) non-fat dry milk (RT, 60 min). Membranes were probed for Kit protein at 4℃ overnight, using either 1 μg/ml of a monoclonal antibody (rat anti-mouse, NOVUS) directed against the extracellular domain of the Kit or a polyclonal antibody (rabbit anti-human, mouse, rat; NOVUS) directed against the amino acid near S715 of the human Kit (1 μg/ml). After incubation with the primary antibody, blots were labelled with species-matched HRP-conjugated secondary antibodies (Santa Cruz) at RT for 1 h, visualized by chemiluminescence staining (Thermo, USA) and documented on hyperfilm (Roche, Mannheim Germany).

Protein lysate from Kit expressed human megakaryoblast cell line (ATCC no. CRL-2021) was set as positive control and protein lysate from Kit negative mouse myoblast cell line (ATCC no. CRL-1772) was set as negative control. The same blot was stripped and re-probed with mouse beta-actin primary antibody (Santa Cruz) as internal control.

2.2.10 Immunofluorescence

c18-4 and CRL-2053 cells which grew on galss coverslips were wash twice with PBS and fixed for 20 min in 4% paraformaldehyde. The coverslips with the cells were washed 3 times with PBS and permeabilized wtih 0.1% Triton X-100 in PBS for 5 min. The coverslips with the cells were washed 3 times with PBS again and were incubated with 5% normal goat serum (Santa Cruz) in PBS for 30 min before being incubated with the primary antibody overnight at 4°C. The cells were then incubated with the secondary antibody and mounted with UltraCruz™ Mounting Medium with DAPI (SantaCruz, CA, USA). For negative controls, goat serum was used to replace the primary antibody for the overnight incubation. Antibodies used in this study included: the FITC monoclonal rat-anti-mouse Kit extracellular domain (1:200, 105805, BioLegend, USA), the monoclonal rat-anti-mouse kit extracellular domain (1:200, NBP1-43359, NOVUS, USA), the monoclonal rat-anti-mouse kit extracellular domain (1:100, KJ-14, Santa Cruz, USA); the polyclonal rabbit-anti-human/mouse/rat kit intracellular domain (1:200, NBP1-19865, NOVUS, USA), the polyclonal goat-anti-mouse kit C-terminus (1:100, M14, Santa Cruz, USA) and the polyclonal rabbit -anti-human/mouse kit C-terminus (1:100, C19, Santa Cruz, USA). The secondary antibodies used in this study included: the Alexa 488-conjugated goat-anti-rat IgG (1:500; Invitrogen Inc, Carlsbad, CA, USA); the Alexa 594-conjugated goat-anti-rabbit IgG (1:500; Invitrogen Inc, Carlsbad, CA, USA) and the Texas red-conjugated donkey-anti-goat IgG (1:100; Santa Cruz, USA). The cell nuclei were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) contained in the mounting medium.

The 5-μm-thick paraffin-embedded sections of testis were fixed on grease-free slides. The sections were deparaffinized in xylene and rehydrated through a graded ethanol series. Immerse the sections in sodium citrate buffer (10 mM sodium citrate, pH 6.0) at high power for two cycles of 5 min each in a microwave oven to retrive the epitopes. The sections were cooled down to room temperature in the retrieval buffer before continuing with the immunofluorescence procedure. Blocking was carried out using 10% normal goat serum in PBS for 30 min at room temperature, followed by an overnight incubation at 4 °C with the *c-kit* primary antibodies. The following procedures were same as that for cultured cells. Immunostaining of sections was repeated at least three times.

2.2.11 Statistical analysis

Statistical analysis was performed by unpaired two-tail student t test using SPSS (version 17.0). All experiments were performed at least three independent repeats and a P value less than 0.05 were considered statistically significant.

Chapter 3

Results

3.1 Transcription of *c-kit* before and after SSCs differentiation

We studied *c-kit* mRNA expression profile in both cell lines and testes by Northern blot with two RNA probes that hit extracellular (exons 10-12) and intracellular domain (exons 18-20) respectively. It was shown that, besides the full-length transcript, several shorter transcripts were transcribed. In cell line c18-4, a 1.5 kb and a 4~5 kb shorter transcript were discovered. In cell line CRL-2053, a 2.7 kb and a 4~5 kb shorter transcript were discovered. In 5 dpp, 10 dpp and 60 dpp testes, a 4~5 kb transcript was discovered (Figure 3.1). We analyzed all of the observed distinctive transcripts by RACE in the subsequent experiments.

Detailed sequences of *c-kit* transcripts expressed in c18-4, CRL-2053 and testes were assayed by RACE. In order to reveal all the possible transcripts, two set of *c-kit* gene specific primers were used. One set hits the region exons 10-12 of full-length transcript and the other set hits exons 18-21 (Figure 3.2). Several sets of primers were designed to ensure the specificity of the products obtained in RACE. It was shown that a shorter transcript with a length of 4.6 kb was expressed in the c18-4 but not in the CRL-2053. This transcript started from *c-kit* intron 9 and included all the downstream exons. Another shorter transcript with a size of 4 kb was expressed in the CRL-2053 but not in the c18-4. This transcript consisted 21 exons, but with a 1.2 kb shorter 3′ UTR than the full-length transcript. Several other truncated *c-kit*

transcripts were found in both the c18-4 and the CRL-2053 cell lines as shown in Figure 3.3. Multiple blast assay show that exons 17-21 is a highly conserved region in the *c-kit* transcripts. Detailed structures of these transcripts were listed on Figure 3.3 and a multiple blast of their sequences was showed on Appendix 1.

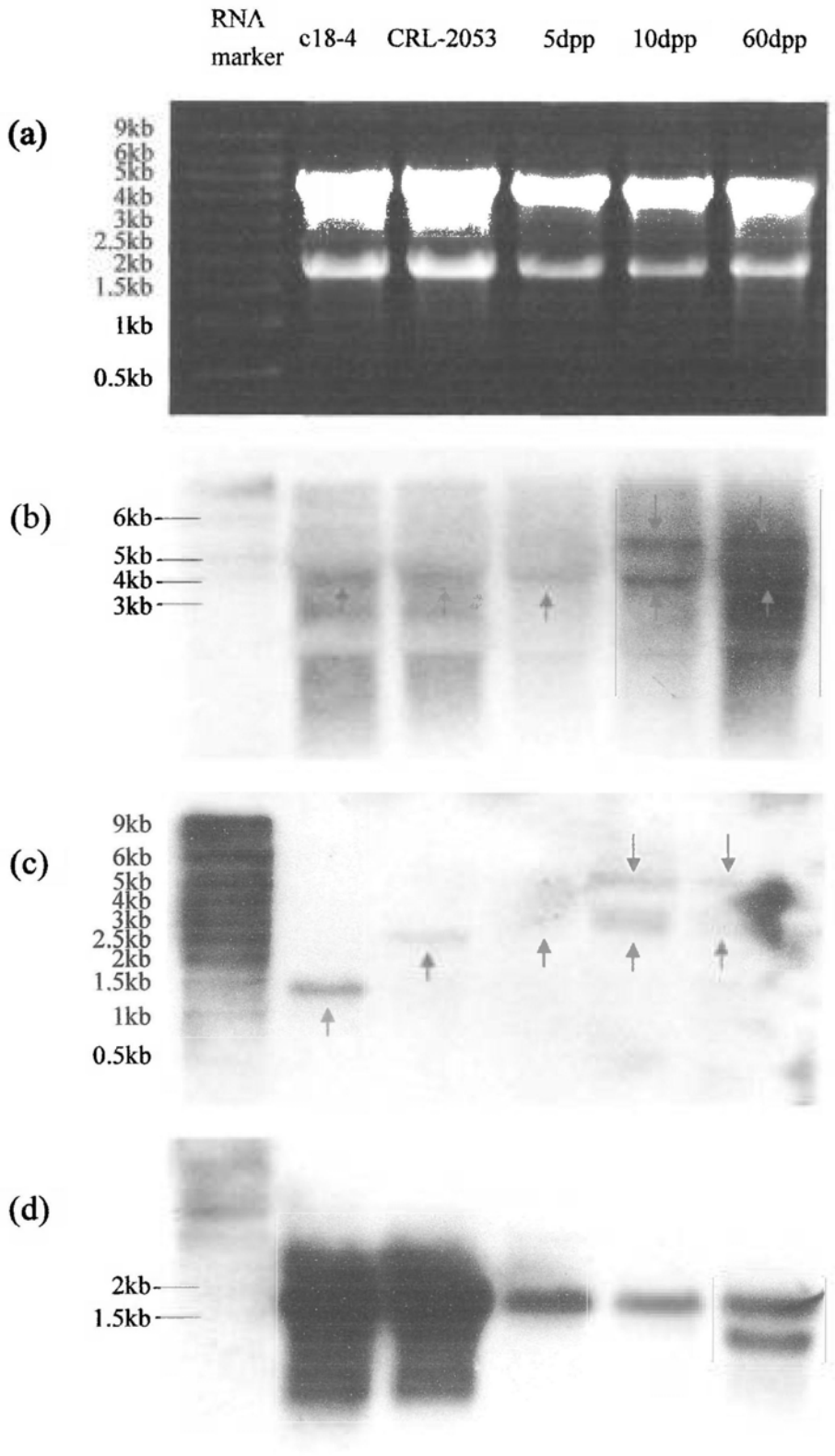


Figure 3.1 *c-kit* northern hybridization. (a) RNA electrophoresis. 10 μg total RNA from different samples were loaded. RNA sizes were marked with 2 μg (2μl) RNA Millennium size markers (Ambion). All RNA samples were in good qualities. (b) Northern hybridization with *c-kit* probe hit exons 10-12. Short transcripts (purple

arrowhead) with a size between 4 and 5 kb could be observed in c18-4, CRL-2053 and different aged testes. Long transcripts (blue arrowhead) with a size between 5 and 6 kb were seen in 10 dpp and 60 dpp testes. (c) Northern hybridization with *c-kit* probe hit exons 18-20. A 1.5 kb short transcript (red arrowhead) was observed in c18-4 and a 2.7 kb short transcript (green arrowhead) was observed in CRL-2053. Short transcripts (purple arrowhead) with a size between 4 and 5 kb could be observed in 5 dpp, 10 dpp and 60 dpp testes. Long transcripts (blue arrowhead) with a size between 5 and 6 kb were seen in 10 dpp and 60 dpp testes. (d) Northern hybridization of beta-actin (internal control).

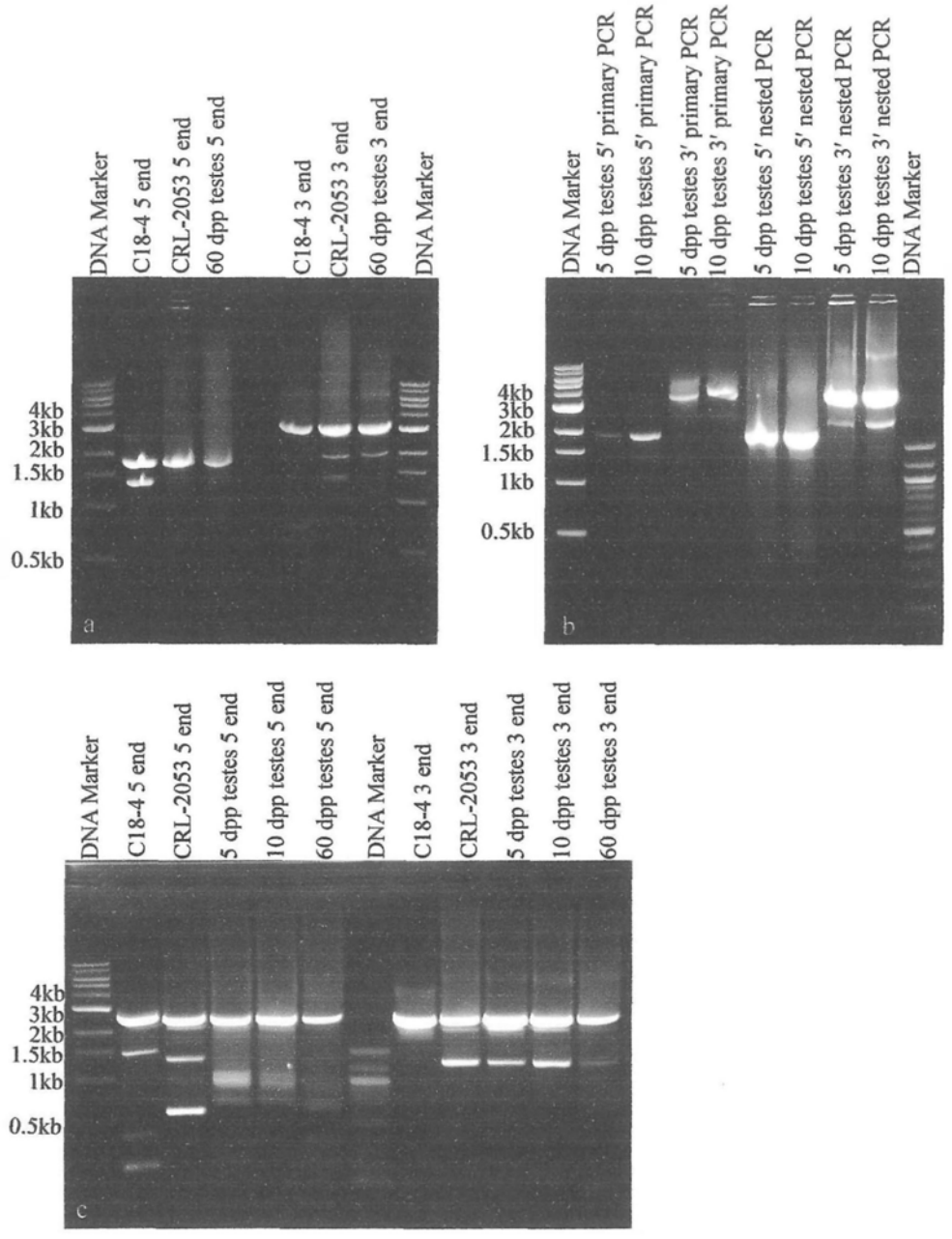


Figure 3.2 RACE PCR products gel electrophoresis. (a), (b) 5′ and 3′ RACE with primer sets hit exons 10-12. 2 bands could be observed in c18-4 5′ RACE, one band was about 1.8 kb and the other was about 1.4 kb. The 1.4 kb band could not be observed in CRL-2053 and 60 dpp testes, but may be very weak in 5 dpp and 10 dpp testes. For the 3′ RACE, a 3 kb bright band was observed in all of the samples. Another smaller band than the 3 kb one could be observed in CRL-2053 and testes but not in c18-4 cell lines. (c) 5′ and 3′ RACE with primer sets hit exons 18-21. A bright band around 2.6 kb was observed in 5′ of all the samples. Several short bands in different size could be seen in both c18-4 and CRL-2053. For the 3′ RACE, a

bright 2.5 kb band could be seen in all samples, a small band about 1.4 kb was observed in CRL-2053 and different age testes, but not in c18-4.

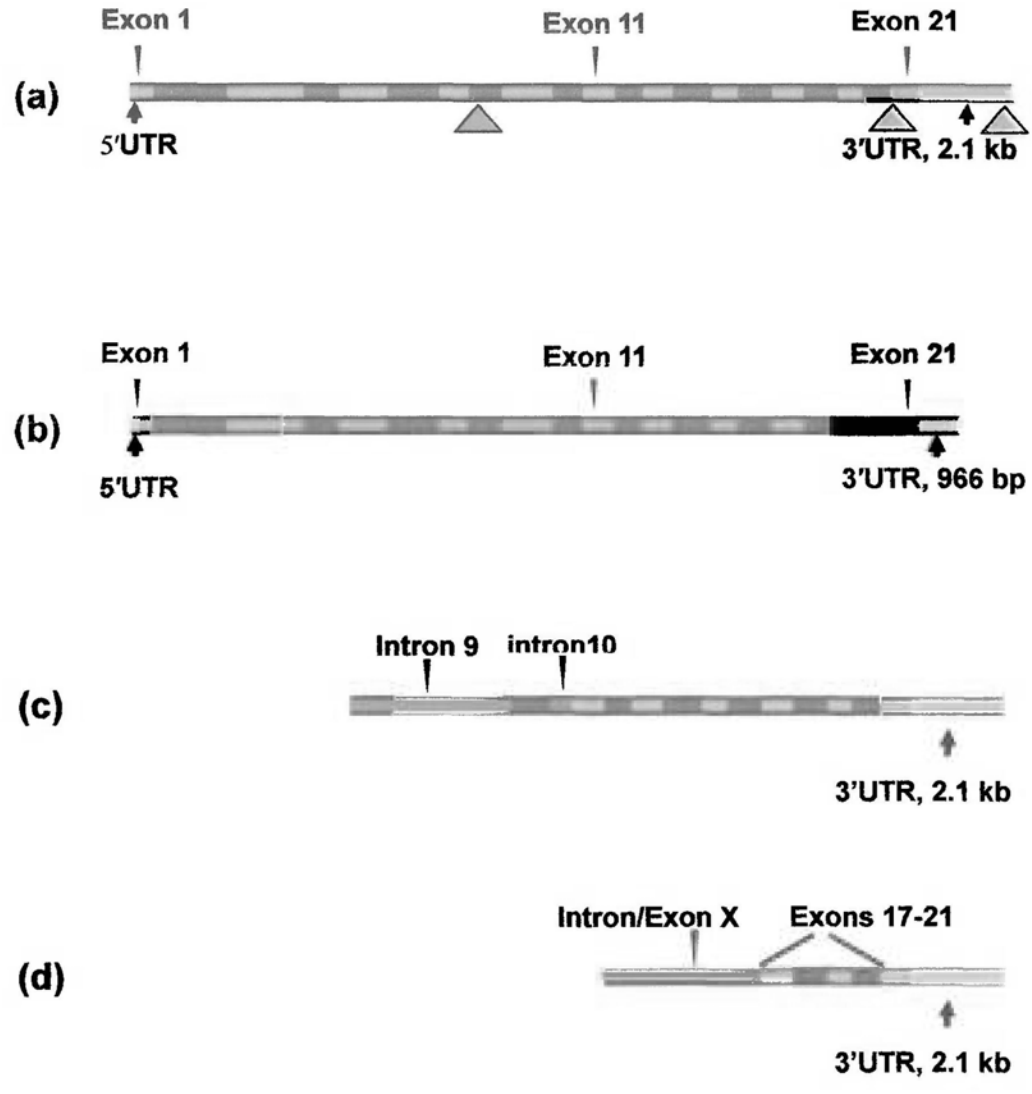


Figure 3.3 *c-kit* transcripts before and after differentiation of SSCs. (a) Full-length transcript. This transcript was composed of 21 exons and with a size about 5.2 kb. A 2.1 kb 3′ UTR region located at exon 21 of this transcript and it was expressed by SSCs before and after differentiation. Green triangles on this figure represented 3 groups of *c-kit* primers that hit exons 7-8, exons 20-21 and exon 21 respectively (b) Short 3′ UTR transcript. This transcript was composed of 21 exons and with a size about 4 kb. The 3′ UTR was 966 bp long. (c) SSCs specific transcript.

This transcript started from intron 9 of the full-length transcript, including exon 10, intron 10 and the following exons as the full-length transcript. It may code a new protein with a novel extracellular domain. (d) Truncated transcripts. It represented a group of transcripts that started from different intron or exon of *c-kit* gene, the Intron/Exon X means the beginning of these transcripts. We found some of these transcripts started from exon 13, exon 15 and exon 17 of the *c-kit* gene respectively. All these truncated transcripts we found containing a conserved domain from exon 17 to exon 21.

3.2 Quantitative analysis of different *c-kit* transcripts expression in the testes and the germ cell lines.

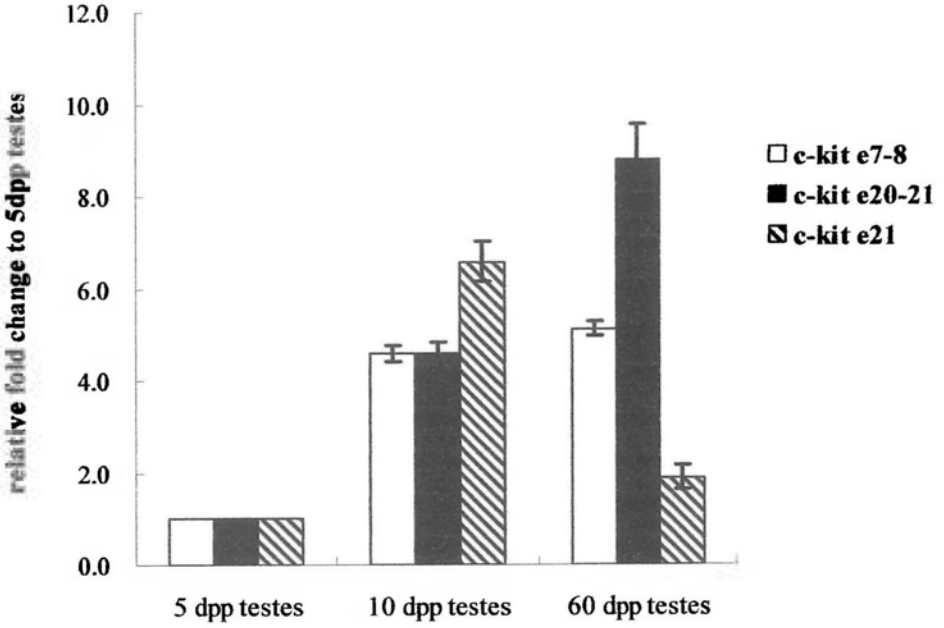
As most of the novel transcripts we found contained the same sequence as that in the full-length transcripts, it was hard to design specific markers to distinguish them. However, they could be grouped into different groups based on their structures. According to whether the extracellular domain coding sequence was involved or not, the *c-kit* transcripts could be grouped into extracellular domain coding sequence-containing transcripts (a and b on Figure 3.3) and intracellular domain coding sequence only transcripts (c and d on Figure 3.3). According to the size of 3′ UTR, these transcripts can be grouped into long 3′ UTR transcripts (a, c and d on Figure 3.3) and short 3′ UTR transcripts (b on Figure 3.3). Therefore, 3 groups of Taqman gene expression assay primers and probes were applied to quantitatively assay the expression of these transcripts. The first primer and probe mix hit exons 7-8 and could only detect the full-length transcripts. The second primer and probe mix hit exons 20-21 which could detect all of the *c-kit* transcripts. The third primers and probe mix hit the short 3′ UTR absent region on exon 21. Therefore, the third primers and probe mix could only detect the long 3′ UTR containing transcripts (Figure 3.3).

Expression of all kinds of *c-kit* transcripts increased significantly in the 10 dpp testes compared to that in the 5 dpp ones. From 10 dpp to 60 dpp, total *c-kit* transcripts (detected by *c-kit* e20-21) and the full-length *c-kit* transcript (detected by *c-kit* e7-8) increased moderately but long 3′UTR transcripts (detected by *c-kit* e21)

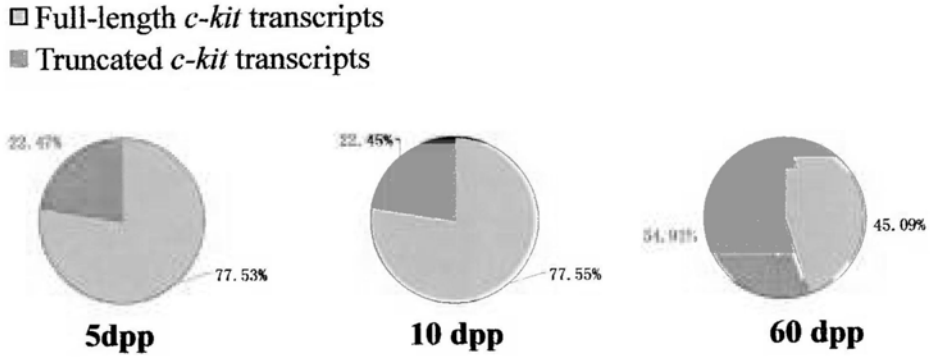
decreased. The ratio of the truncated *c-kit* transcripts were calculated with the formula (1-ratio$_{\text{full-length transcripts}}$)×100 %. From 5 dpp to 10 dpp, the ratio of full-length and truncated *c-kit* transcript did not changed, even though the total *c-kit* mRNA level increase significantly. From 10 dpp to 60 dpp, ratio of truncated *c-kit* transcripts increased and full-length *c-kit* transcripts decreased though absolute expression of full-length *c-kit* transcripts increased. This indicated that expression of truncated transcripts increased repidly than that of full-length *c-kit*. Expression level of long 3′ UTR *c-kit* transcripts only make up very minor part of total *c-kit* transcripts (Figure 3.4).

We compared the *c-kit* mRNA expression level in germ cell lines that represent different stages of germ cell development. The c18-4 is an undifferentiated spermatogonial stem cell line, the CRL-2053 is a differentiating type B spermatogonia and the CRL-2196 is a spermatocyte cell line. It was shown that the expression of *c-kit* mRNA was very low in the c18-4, which was consistent with previous studies. Ratio of the full-length transcripts in c18-4 was about 65%. The expression level of total *c-kit* mRNA increased significantly in the CRL-2053 comparing with that in the c18-4. The ratio of the full-length transcripts increased to 92.9% in the CRL-2053. The long 3′ UTR transcripts (GeneBank reference *c-kit* mRNA) was not detectable with probe e21 in all the three germ cell lines (Figure 3.5).

(a)


12.0
10.0
8.0
6.0
4.0
2.0
0.0
relative fold change to 5dpp testes
5 dpp testes
10 dpp testes
60 dpp testes
c-kit e7-8
c-kit e20-21
c-kit e21

(b)


Full-length c-kit transcripts
Truncated c-kit transcripts
22.47%
77.53%
22.45%
77.55%
54.91%
45.09%
5dpp
10 dpp
60 dpp

(c)

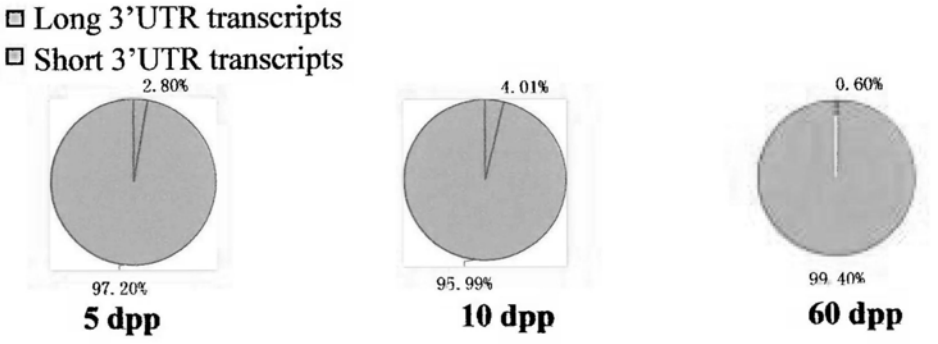
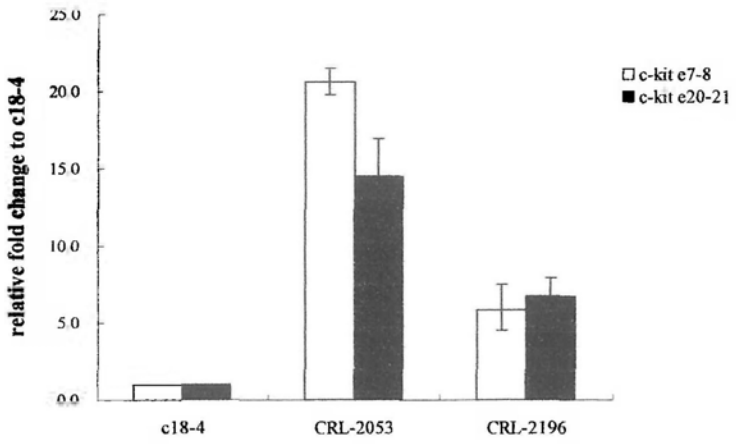

Long 3'UTR transcripts
Short 3'UTR transcripts
2.80%
97.20%
4.01%
95.99%
0.60%
99.40%
5 dpp
10 dpp
60 dpp

Figure 3.4 Profiles of *c-kit* transcripts in 5 dpp, 10 dpp and 60 dpp testes. (a) Relative *c-kit* mRNA expression level in the 5 dpp, 10 dpp and 60 dpp mouse testis. Error bars represent S.E.M. *c-kit* e7-8, real-time PCR using probe hit full-length *c-kit* transcripts on exons 7-8. *c-kit* e20-21, real-time PCR using probe hit full-length *c-kit* transcripts on exons 20-21. *c-kit* e21, real-time PCR using probe hit full-length *c-kit* transcripts on the end of exons 21. Differences between each two groups were significant ($P<0.05$). (b) The ratio of Full-length and truncated *c-kit* transcripts in 5 dpp, 10 dpp and 60 dpp mouse testes. The ratio of truncated *c-kit* transcripts were calculated by the formula $(1\text{-ratio}_{\text{full-length transcripts}})\times 100$ %. (c) The ratio of long 3′ UTR and short 3′ UTR *c-kit* transcripts in 5 dpp, 10 dpp and 60 dpp mouse testes. The ratio of short 3′ UTR *c-kit* transcripts were calculated by the formula $(1\text{-ratio}_{\text{long 3'UTR transcripts}})\times 100$ %.

(a)



(b)

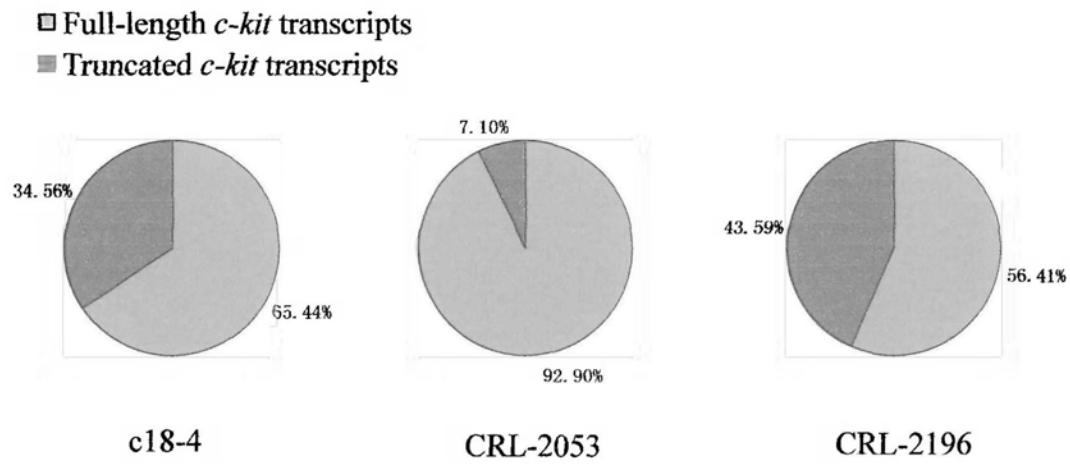


Figure 3.5 *c-kit* transcripts profile in germ cell lines. (a) Relative *c-kit* mRNA expression level in the c18-4, CRL-2053 and CRL-2196 germ cell lines. Error bars represent S.E.M. *c-kit* e7-8 represent full-length *c-kit* transcripts. *c-kit* e20-21 represent total *c-kit* mRNA. Differences between each two groups were significant ($P<0.05$). (b) The ratios of Full-length and truncated *c-kit* transcripts in c18-4, CRL-2053 and CRL-2196 germ cell lines. The ratio of truncated *c-kit* transcripts were calculated by the formula $(1\text{-ratio}_{\text{full-length transcripts}})\times 100\ \%$.

3.3 Translation of *c-kit* before and after SSCs differentiation

Two isoforms of Kit proteins had been found during mouse spermatogenesis. One was encoded by the full-length *c-kit* transcript and expressed in differentiating spermatogonia. An extracellular domain, a trans-membrane domain and an intracellular domain composed this isoform. Another isoform was called Tr-Kit encoded by truncated *c-kit* transcripts that was found only in the mouse spermatids. Tr-Kit was composed by part of the intracellular domain.

We found that the truncated *c-kit* transcripts are a group of different mRNA that initiated from different exons or introns of *c-kit* gene. The Tr-kit mRNA was not only expressed in the spermatids but also in the SSCs no matter they are before or after differentiation. We found a SSCs-specific transcript, which was only expressed before differentiation, and a short 3′ UTR transcript, which was only expressed after differentiation. Based on the sequences of these 2 transcripts, we predicted the corresponding proteins by commercial software NCBI ORF finder and Blastx. A 502 amino acid-long protein was predicted by the SSCs-specific transcript. The ORF of the SSCs-specific transcript started from intron 9 of *c-kit* gene and containing transmembrane and intracellular domain coding sequence as the full-length transcript. Therefore, after translation, a protein with a unique extracellular domain comparing with the full-length one might produced. The short 3′ UTR transcript encoded a protein that was the same as the full-length Kit. A multiple sequences alignment of ORF finder predicted *c-kit* proteins was showed in Appendix 2.

In order to discriminate the different types of Kit expressed before and after SSCs differentiation, we bought two groups of antibodies for Western blotting and immunofluorescence staining. The first group of Kit antibodies was monoclonal antibodies bound to the extracellular domain of the Kit. The other of group of antibodies was polyclonal antibodies bound to the intracellular domain of the Kit. According to our prediction, the anti-extracellular domain antibodies should only recognize the full-length Kit and the anti-intracellular domain antibodies should theoretically recognize both the full-length Kit and the truncated Kit. The results showed that full-length Kit was expressed by CRL-2053 cells and different age mouse testes from 5 dpp to 60 dpp but not by c18-4 cells. With the anti-intracellular antibody, a protein with a molecular weight around 50 KDa was detected in the c18-4, CRL-2053 and different age testes. Expression of this protein in testes increased as the mice grew up to 60 dpp (Figure 3.6).

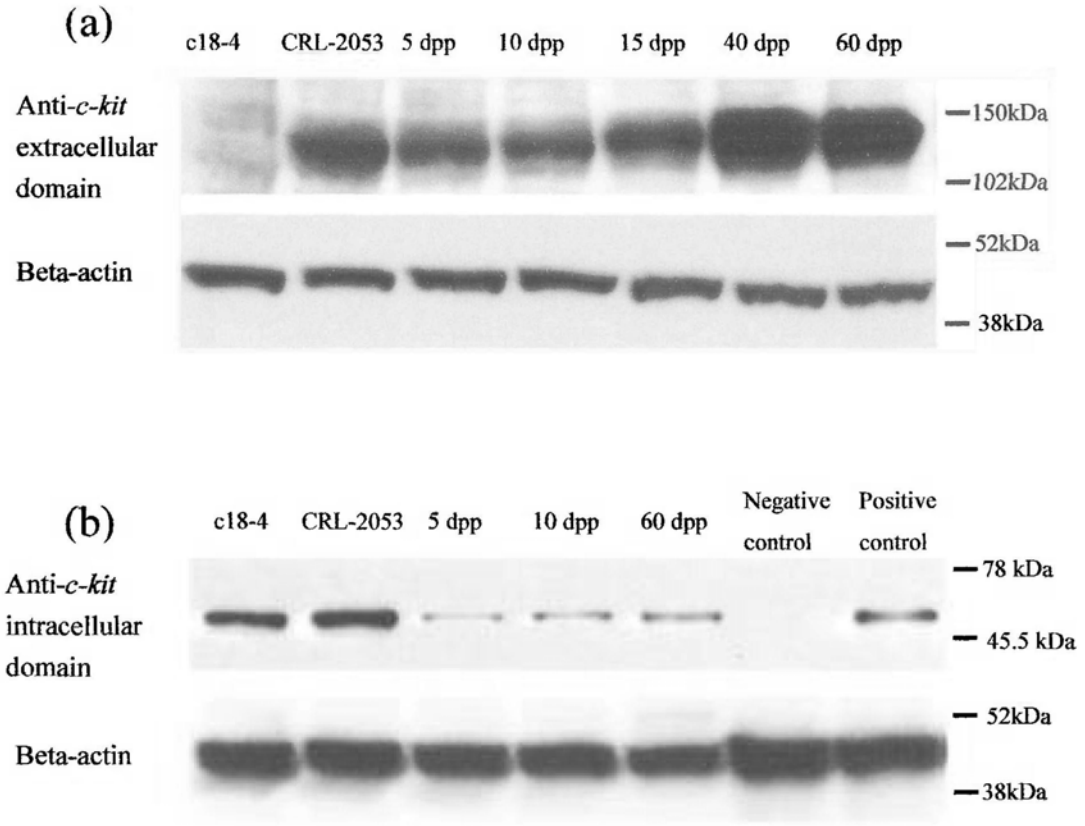


Figure 3.6 Western blot analysis of *c-kit* protein expression in germ cell lines and testes. (a) *c-kit* protein expression assayed by an anti-mouse Kit extracellular domain monoclonal antibody in germ cell lines and testes. (b) *c-kit* protein expression assayed by an anti-mouse Kit intracellular domain monoclonal antibody in germ cell lines and testes.

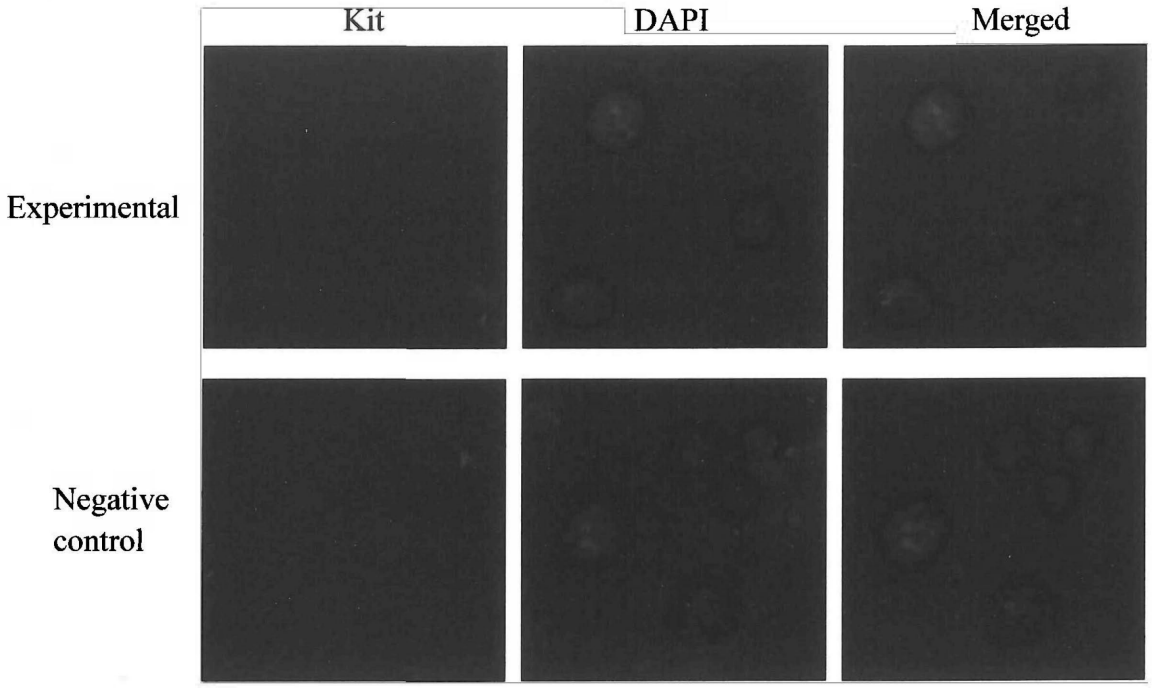
3.4 Distribution of Kit protein before and after SSCs differentiation

In order to identify the distribution of different Kit proteins, immunofluorescence staining of germ cell lines and paraffin fixed mouse testes sections were carried out.

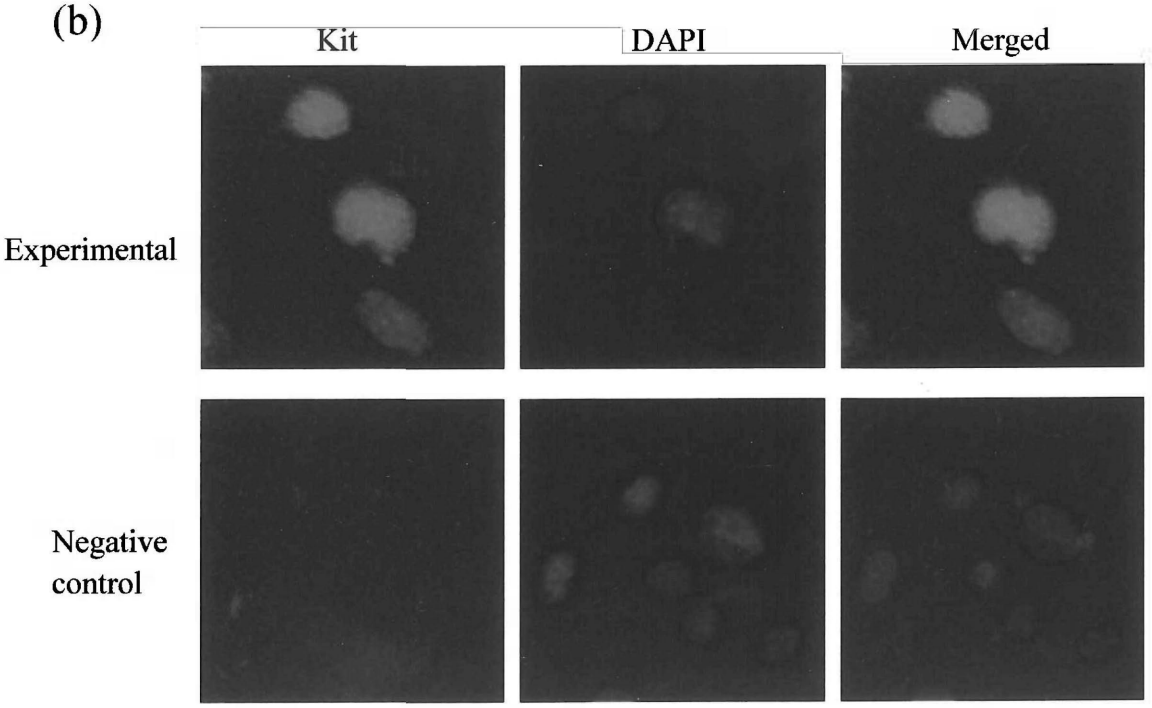
The full-length Kit was expressed in the CRL-2053 cells but not in the c18-4 cells. The full-length Kit was expressed on both the cell membrane and the cytoplasm. However, when staining with an anti-intracellular domain antibody, both the c18-4 and CRL-2053 cells were fluorescein stained. The truncated form of Kit kept expressing in the SSCs before and after their commitment of differentiation (Figure 3.7).

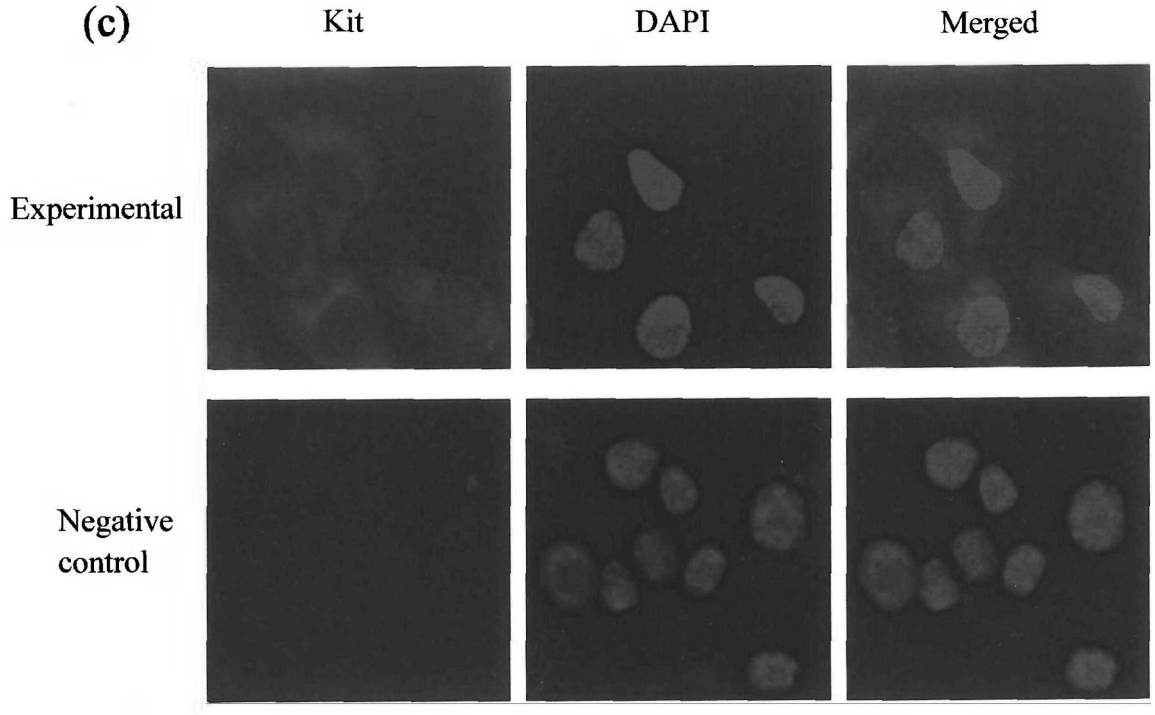
In the immunostaining of the testes sections, expression of full-length Kit was very low in the germ cells but high in the leydig cells of the 5 dpp testes. In the 10 dpp testes, some of the germ cells adjacent to the basement membrane of the seminiferous tubules developed into Kit positive. In the 60 dpp testes, some spermatogonia and spermatocytes were Kit positive. We also found that mature spermatozoa in some seminiferous tubules cross section were highly stained by the antibody. Tr-Kit was expressed by germ cells including spermatogonia, spermatocyte and spermatids, but not in mature spermatozoon. Leydig cells highly expressed the truncated Kit too (Figure 3.8).

(a)


Kit
DAPI
Merged
Experimental
Negative control

(b)


Kit
DAPI
Merged
Experimental
Negative control


(c)
Kit
DAPI
Merged
Experimental
Negative control

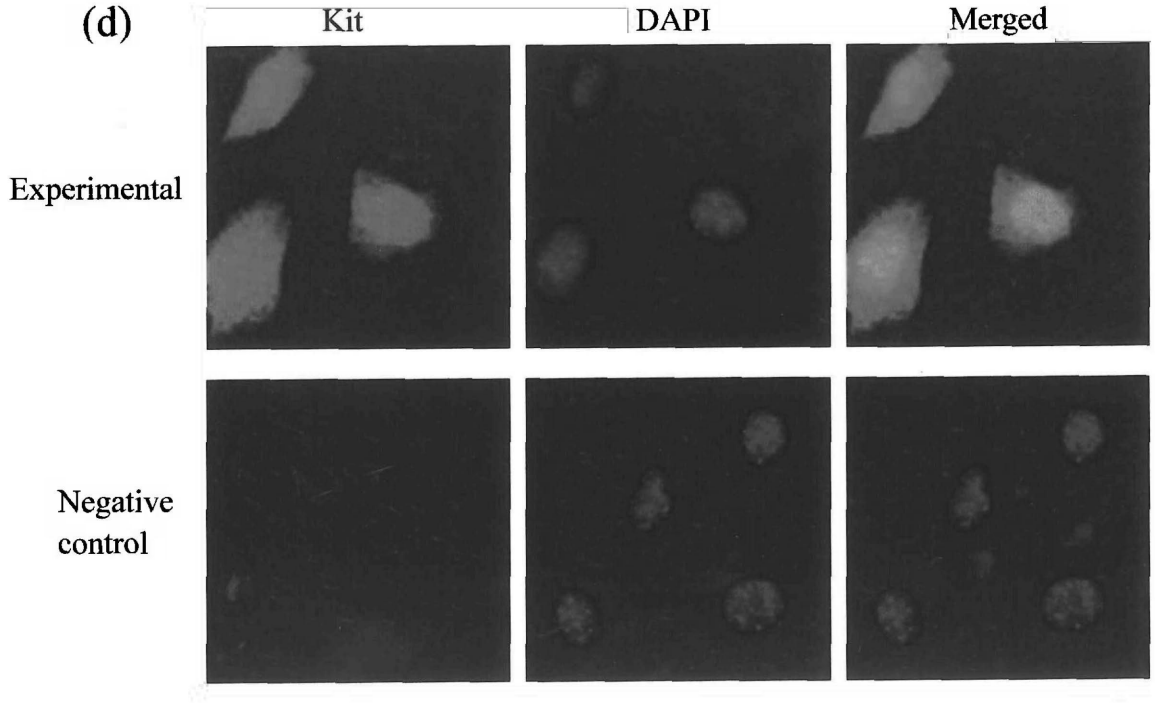
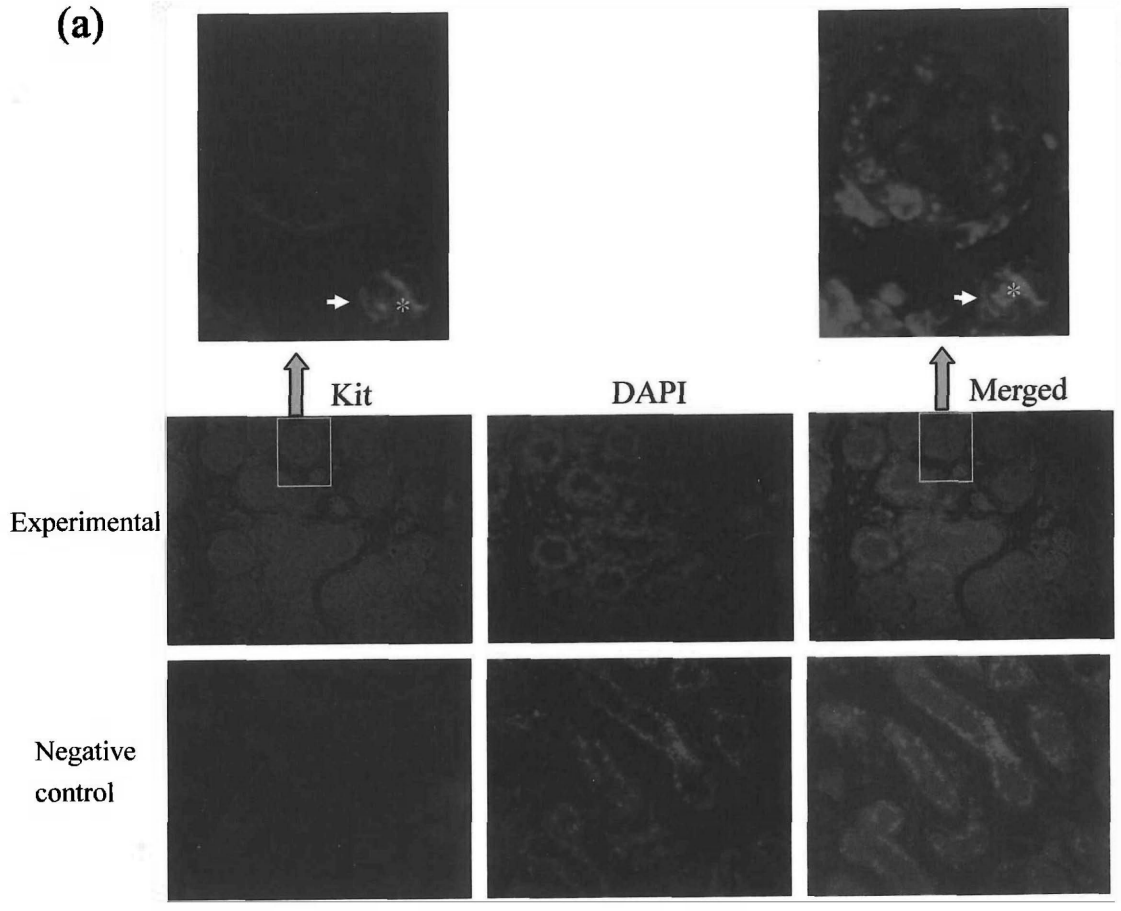
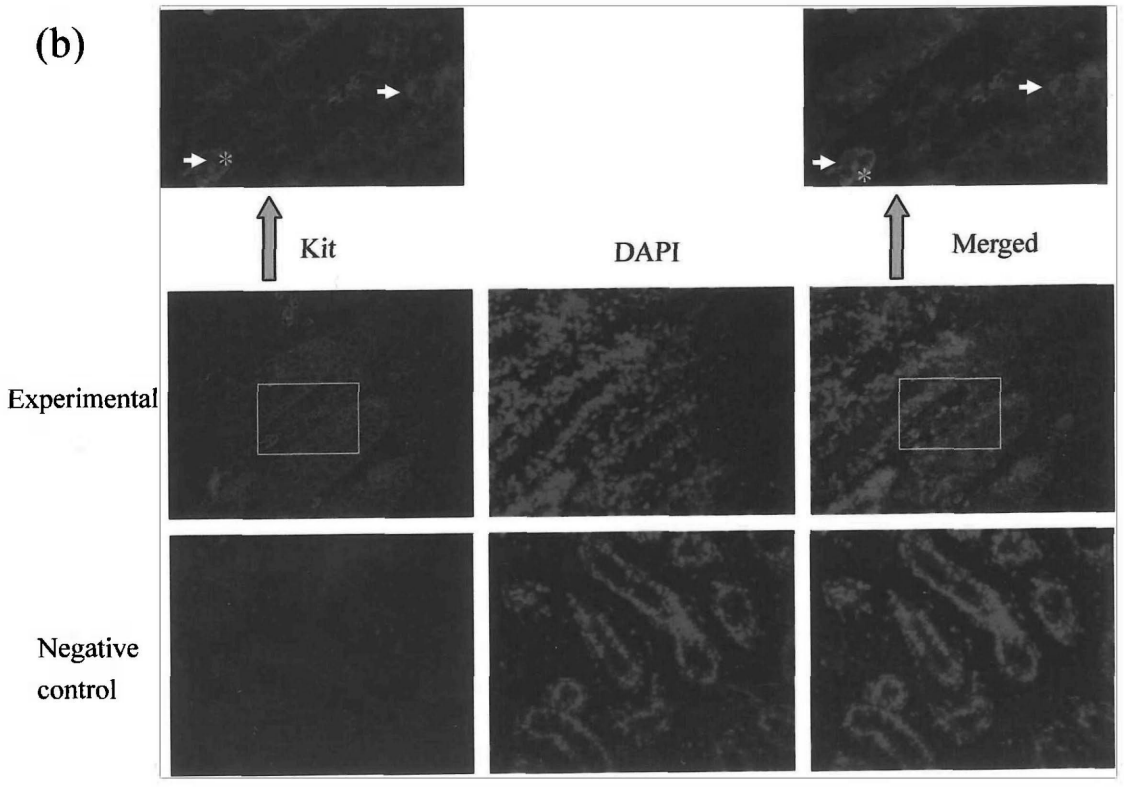

(d)
Kit
DAPI
Merged
Experimental
Negative control

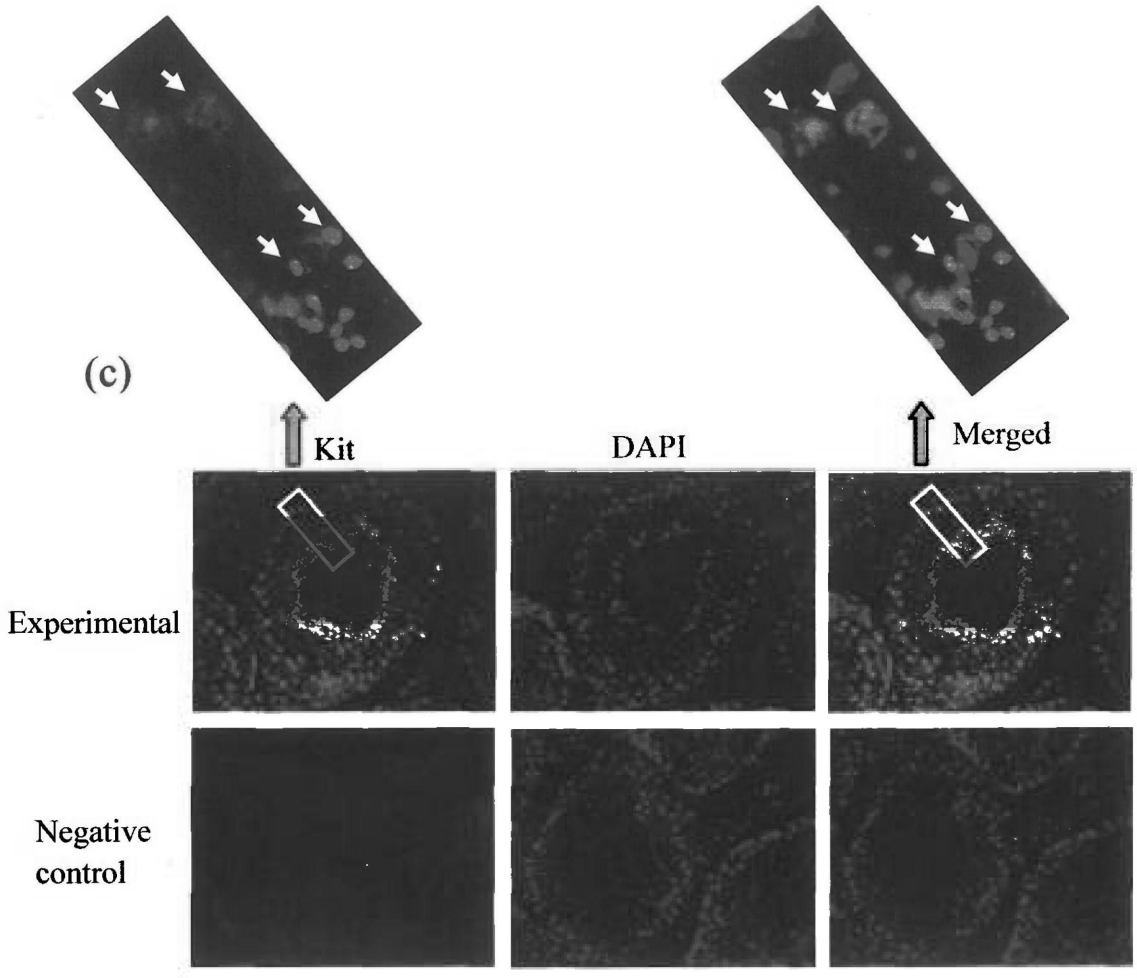
Figure 3.7 Immunofluorescence study of Kit expressed in c18-4 and CRL-2053. Images were obtained with ZEISS Axioplan 2 imaging system and Spot 4.7/SpotAdvanced software (Magnification ×400). Kit positive cells were stained with green or red fluorescein (arrow). The cell nuclei was counterstained with 4′,6-diamidino-2-phenylindole. Images were merged with Photoshop CS4. (a) FITC labeled anti-Kit extracellular domain in c18-4. (b) Alexa Fluo 594 labeled anti-Kit intracellular domain in c18-4. (c) FITC labeled anti-Kit extracellular domain in CRL-2053. (d) Alexa Fluo 594 labeled anti-Kit intracellular domain in CRL-2053.

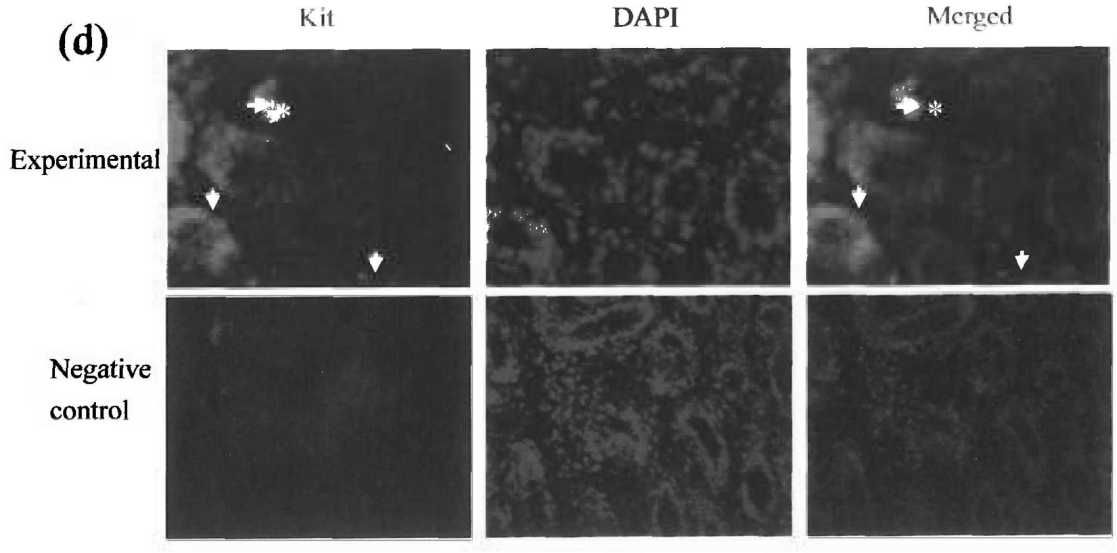
(a)


Kit
DAPI
Merged
Experimental
Negative control

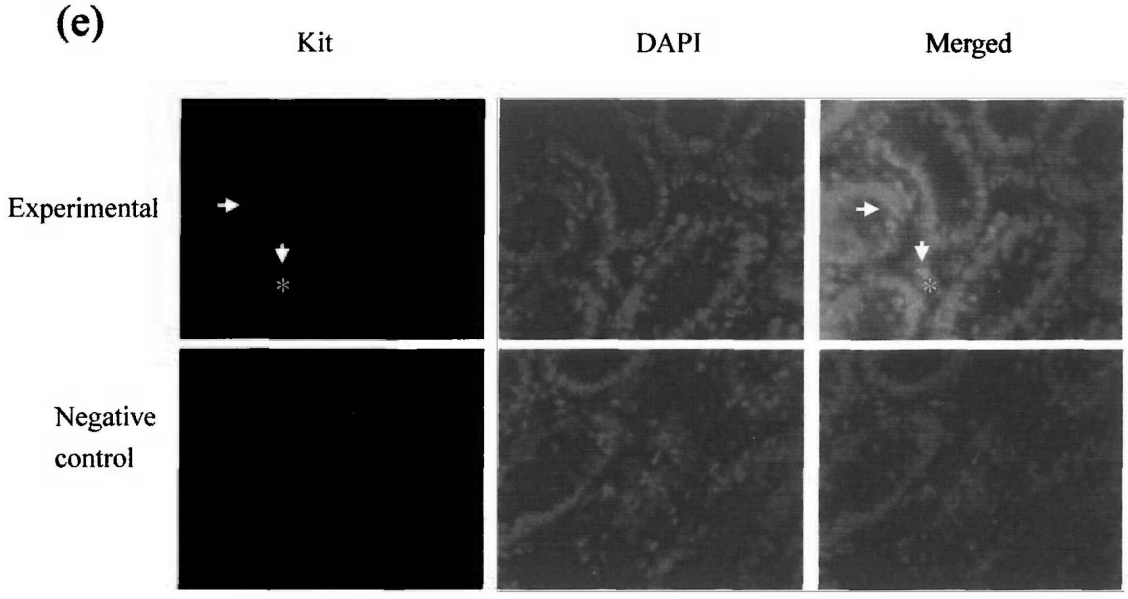
(b)


Kit
DAPI
Merged
Experimental
Negative control


(c)
Kit
DAPI
Merged
Experimental
Negative control


(d)
Kit
DAPI
Merged
Experimental
Negative control

(e)


Kit
DAPI
Merged
Experimental
Negative control

(f)

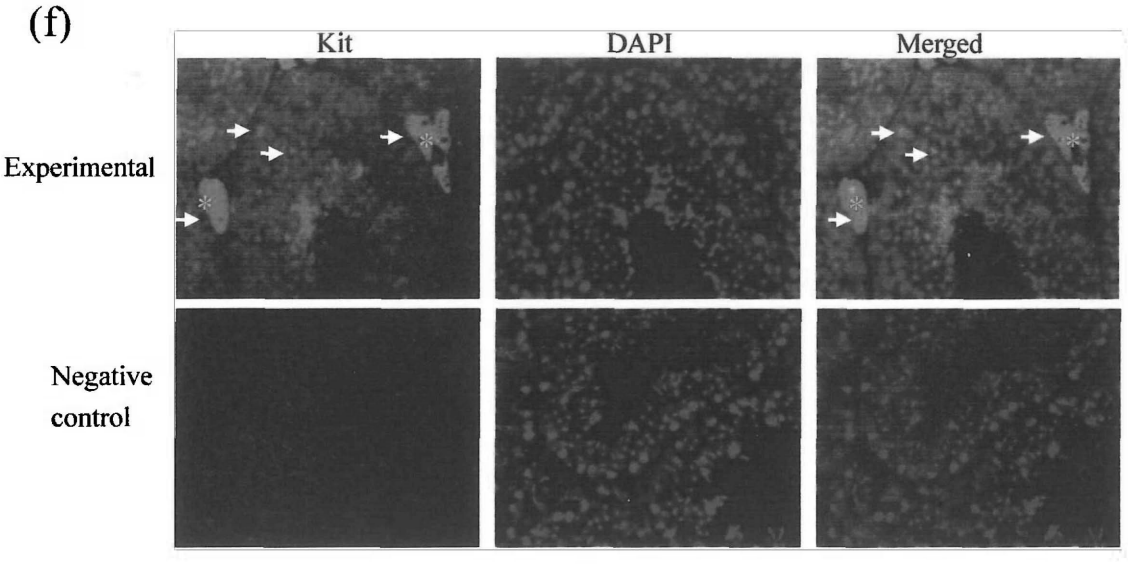

Kit
DAPI
Merged
Experimental
Negative control

Figure 3.8 Immunofluorescence study of Kit protein expression in 5 dpp, 10 dpp and 60 dpp mouse testes. Images were obtained with ZEISS Axioplan 2 imaging system and Spot 4.7/SpotAdvanced software (Magnification ×400). Kit positive cells were stained with green or red fluorescein (Arrow). Leydig cells (Asterisk) displayed a large prominent nucleus. The cell nuclei was counterstained with 4′,6-diamidino-2-phenylindole. Images were merged with Photoshop CS4. (a) FITC labeled anti-Kit extracellular domain in 5 dpp mouse testes. White arrow: Kit positive cells (b) FITC labeled anti-Kit extracellular domain in 10 dpp mouse testes. (c) FITC labeled anti-Kit extracellular domain in 60 dpp mouse testes. (d) Alexa Fluo 594 labeled anti-Kit intracellular domain in 5 dpp mouse testes. (e) Alexa Fluo 594 labeled anti-Kit intracellular domain in 10 dpp mouse testes. (f) Alexa Fluo 594 labeled anti-Kit intracellular domain in 60 dpp mouse testes.

3.5 Expression changes of *c-kit* and other germ cell differentation-related genes in the testes by RA stimulation

5 dpp, 10 dpp and 60 dpp mouse testes were treated with 0.7 μM or 2 μM RA *in vitro* for 24 h. Realtime PCR was applied to assay the expression of *c-kit* and SSCs differentiation related genes (Figure 3.9). The total *c-kit* mRNA level increased following RA treatment and showed a concentration related pattern in 5 dpp, 10 dpp and 60 dpp testes. Thus, RA might be one of the upstream control factors of *c-kit*. RARα and EGR 3 were not changed. Therefore, these two genes might play a parallel role with RA in SSCs differentiation. Cyp26b1 and Stra8 were significantly up-regulated. DAZL and Kitl increased modestly as *c-kit*. BMP4 was down-regulated after RA treatment which was consistent with the nature of being a ploripotent factor of SSCs.

The effects of RA on the expression of BMP4, Cyp26b1 and stra8 were more significant in the 5 dpp testes than that in the 10 dpp testes. On 60 dpp testes, the effects could nearly be neglect. As the ratio of the undifferentiated spermatogonia declined following the increasement of age form 5 dpp to 60 dpp, the results might indicate that RA mainly plays a role in early stage spermatogenesis.

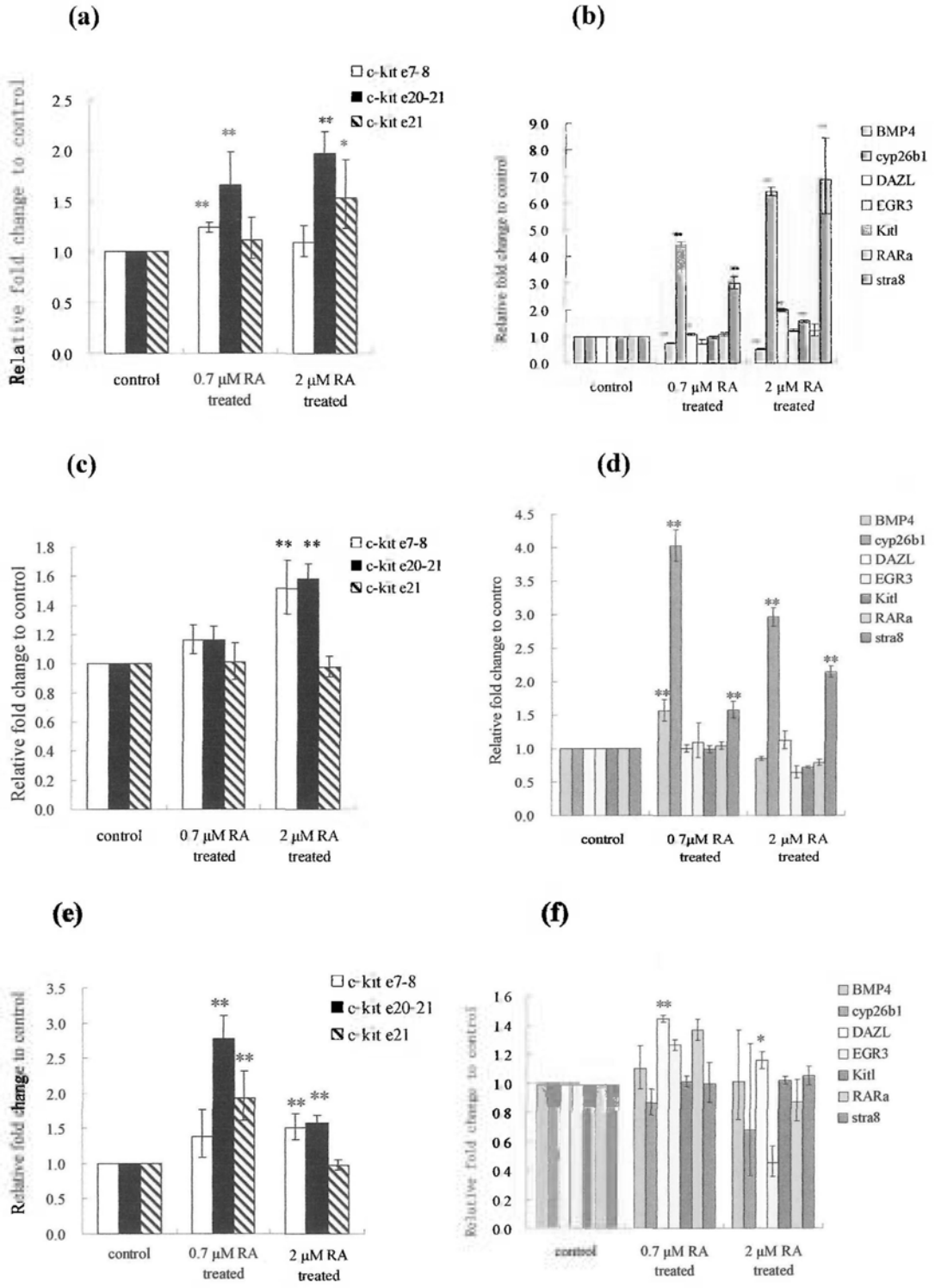


Figure 3.9 Expression of *c-kit* and other germ cell differentiation-related genes in the 5 dpp, 10 dpp and 60 dpp mouse testes stimulated by RA.

5 dpp, 10 dpp and 60 dpp testes were treated with 2 μM RA diluted in ethanol for 24 h *in vitro*. Testes treated with the same concentration of ethanol without RA were the

control group. Realtime PCR was carried out for quantitative determination of the expression of *c-kit* and its potential regulatory genes (BMP4, Cyp26b1, DAZL, EGR3, Kitl, RARα and Stra8). Values of the vertical axis represented the expression fold change comparing with the control group. The results were normalized to GAPDH values. Error bars represent the S.E.M. Values with ** represented a significance with a $P<0.01$ whereas values with * represented a significance with a $P<0.05$ comparing with the control group. (a) *c-kit* expression in the 5 dpp testes. (b) Expression of germ cell differentiation-related genes in the 5 dpp testes. (c) *c-kit* expression in the 10 dpp testes. (d) Expression of germ cell differentiation-related genes in the 10 dpp testes. (e) *c-kit* expression in the 60 dpp testes. (f) Expression of germ cell differentiation-related genes in the 60 dpp testes.

3.6 Expression dynamics of *c-kit* and some stem/germ cell-related genes in spermatogonia cell lines by RA stimulation

Testis is a mixture of germ cells, Sertoli cells and leydig cells. According to our results, Leydig cells were the biggest source of *c-kit* in testes. So, it was difficult to discriminate which type of cells contributed to the *c-kit* and other differentiation-related genes expression changes after RA stimulation in the experiment described in part 3.5. To make it clearer, we used germ cell lines as hosts and re-analyzed *c-kit* and other differentiation-related genes changes after RA stimulation. C18-4 and CRL-2053 cells were treated with 2 μM RA for 24 hours. Unlike in the testes, the expression of *c-kit* decreased in the c18-4 cells after RA stimulation. In CRL-2053 cells, the full-length *c-kit* transcripts decreased to about 50% but the total *c-kit* mRNA level increased to about 2-fold. Therefore, we inferred that the truncated forms of *c-kit* mRNA should increase significantly. The alteration

was consistent with the changes from type B spermatogonia to spermatocyte. Expression of Cyp26b1 and Stra8 increased significantly in the c18-4 and CRL-2053 cells ($P<0.01$). This indicated that RA was working in these 2 cell lines. Other germ cell differentiation-related gene showed different alteration after RA treatment. EGR3 decreased significantly in both cell lines. Expression of Pou5f1, a marker gene of SSCs, decreased significantly in the c18-4 cells but not in the CRL-2053 cells (Figure 3.10).

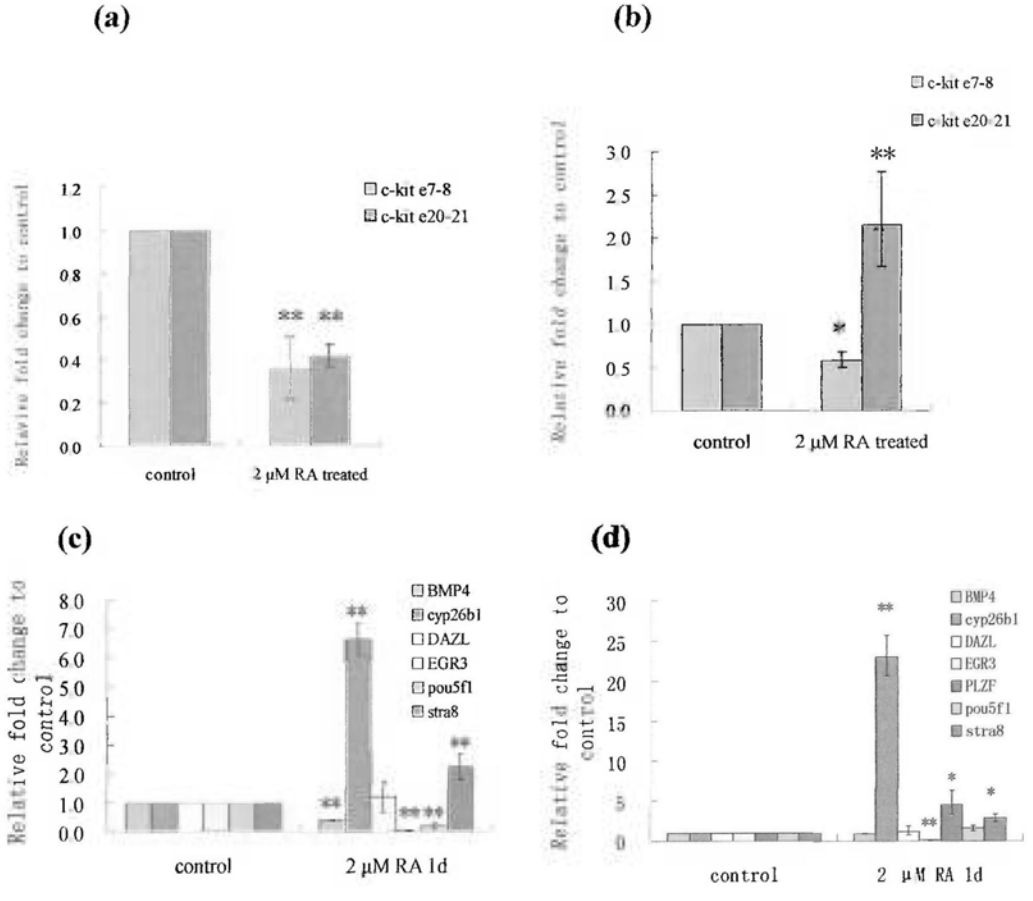


Figure 3.10 Expression of *c-kit* and other germ cell differentiation-related genes in the c18-4 and CRL-2053 cells stimulated by RA.

The c18-4 and CRL-2053 cells were treated with 2 μM RA diluted in ethanol for 24 h *in vitro*. Cells treated with the same concentration of ethanol without RA were the control group. Realtime PCR was carried out for quantitative determination of the expression of *c-kit* and its potential regulatory genes (BMP4, Cyp26b1, DAZL, EGR3, Pou5f1 and Stra8). Values of the vertical axis represented the expression fold change comparing with the control group. The results were normalized to GAPDH values. Error bars represent the S.E.M. Values with ** represented a significance with a $P<0.01$ whereas values with * represented a significance with a $P<0.05$ comparing with the control group. (a) *c-kit* expression in the c-18-4 cells. (b) *c-kit* expression in the CRL-2053 cells. (c) Expression of germ cell differentiation-related genes in the c18-4 cells. (d) Expression of germ cell differentiation-related genes in the CRL-2053 cells.

3.7 Morphological changes of RA induced pluripotential embryonic carcinoma cell lines (P19).

P19 cells were proved previously to have the ability to differentiate to neuron cells under induction of 500 nM RA. According to some unpublished data from our collaborators, P19 cells could be induced to germ cells with RA. In this study, P19 cells were cultured in the presence of different concentrations of RA *in vitro* for 5 days and samples were collected on 3 different time point (1 day, 3 day and 5 day after the start of induction). Cell morphology and some of the stem/germ cell marker genes were checked after RA treatment.

P19 cells were induced with 7 different RA concentrations from low concentration (10 nM), medium concentration (50 nM, 100 nM, 500 nM) to high concentration (1 μM, 2 μM, 4 μM). Dose-dependent apoptosis was showed when RA

was added for 2 days in the P19 cells. Morphological changes of P19 cells were observed after 5 days of RA treatment. The low RA concentration group showed indistinctive morphological changes comparing with the control group. Cells in the medium concentration group showed a neural like cell structure after 5 days of RA treatment. Most of the cells in the high RA concentration group gone apoptosis after 5 days of RA treatment. The survived cells in the high RA concentration group were further cultured with normal cell culture medium without RA for another 9 days. Some cell colonies formed and cells of these colonies showed a cell structure similar with the in vitro cultured male germ cell lines (Figure 3.11).

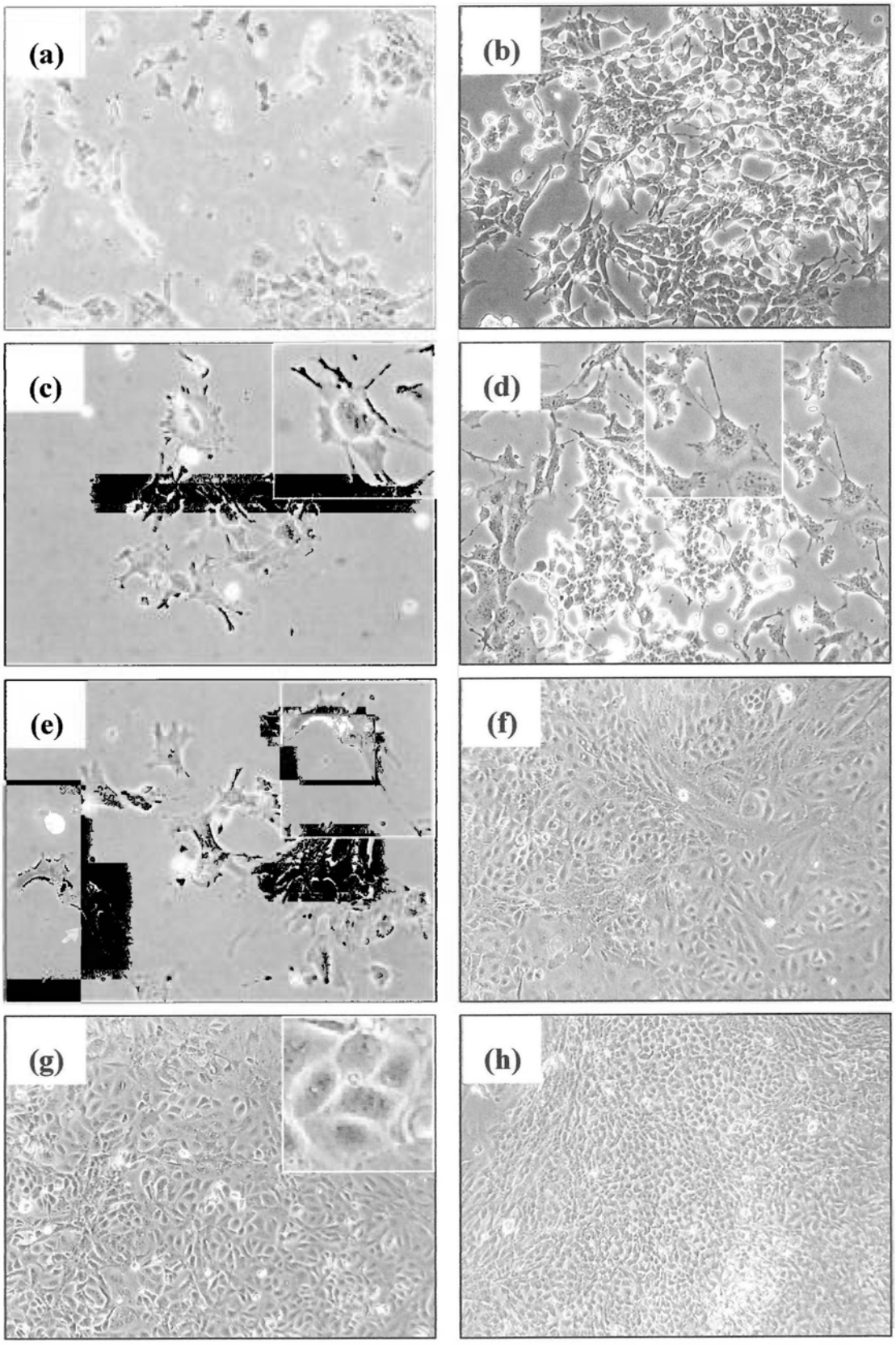
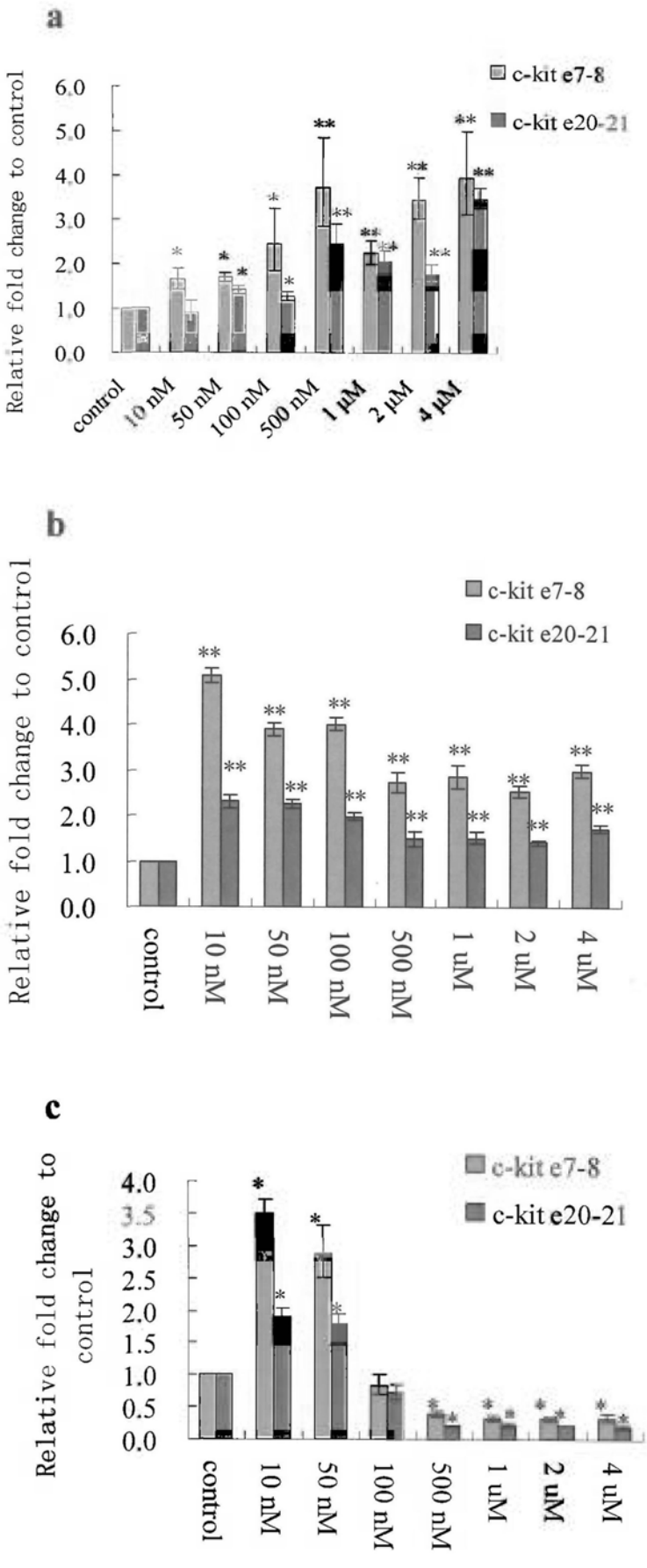

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 3.11 Morphological changes of RA induced pluripotential embryonic carcinoma cell lines (P19). (a) Control group cells. (b) Cells induced by 10 nM RA for 5 days. (c) Cells induced by 50 nM RA for 5 days. (d) Cells induced by 100 nM RA for 5 days. (e) Cells induced by 500 nM RA for 5 days. (f) Cells induced by 1 μM RA for 5 days followed by 9 days of culture in normal media. (g) Cells induced by 2 μM RA for 5 days followed by 9 days of culture in normal media. (h) Cells induced by 4 μM RA for 5 days followed by 9 days of culture in normal media. All photos were capatured with optical microscope (Leica) under a 200 magnification. Yellow arrows represent neural like cells; Red arrow represents cells that showed cell morphology similar with the in vitro cultured male germ cell lines.

3.8 Expression dynamics of *c-kit* and other stem/germ cell-related genes in P19 cells stimulated by RA

After 1 day of RA stimulation, transcription of both the long and the long+short isoforms of c-kit increased and the increase showed a dose dependent manner from 10 nM to 500 nM. It seemed that the long transcript respond to RA better than the short transcript as the long transcript increased more significantly than the long+short transcripts (Fig 3.12a). After 3 days of RA stimulation, expression of c-kit increased the most significantly in the 10 nM group. Expression of c-kit was not increased any more with more RA added (Fig 3.12b). After 5 days of RA stimulation, expression of c-kit dropped significantly in the >500 nM groups (Figure 3.12c). The responding time to RA stimulation of both isoforms of c-kit was around 1-3 days after which, a non-responding period is followed (Fig 3.12d and Fig 3.12e).


a
c-kit e7-8
c-kit e20-21
Relative fold change to control
6.0
5.0
4.0
3.0
2.0
1.0
0.0
control
10 nM
50 nM
100 nM
500 nM
1 μM
2 μM
4 μM
b
c-kit e7-8
c-kit e20-21
Relative fold change to control
control
10 nM
50 nM
100 nM
500 nM
1 uM
2 uM
4 uM
c
c-kit e7-8
c-kit e20-21
Relative fold change to control
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
control
10 nM
50 nM
100 nM
500 nM
1 uM
2 uM
4 uM

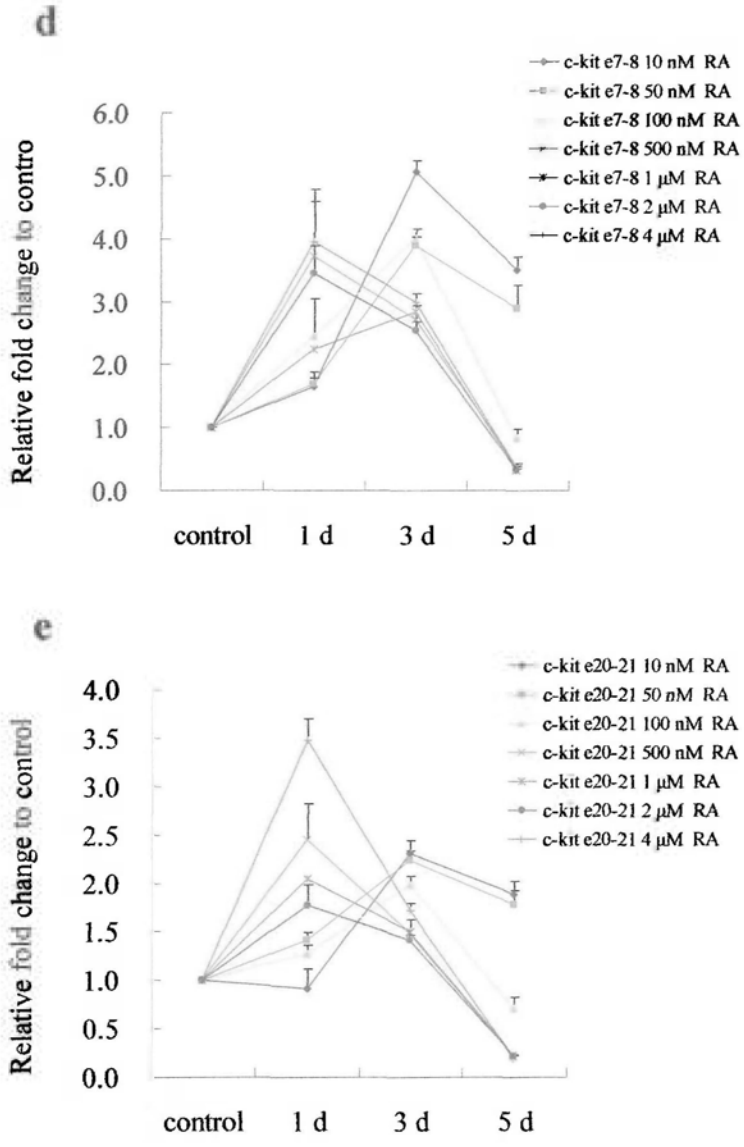


Figure 3.12 Expression dynamics of *c-kit* in P19 cells stimulated by RA.

P19 cells were cultured with different RA concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1 μM, 2μM and 4μM). Expression of *c-kit* was quantified by real-time PCR. Values of the vertical axis represented the expression fold change comparing with the control group. (a-c) Expression dynamics of the full-length *c-kit* transcript (amplified with primers hit c-kit e 7-8) and all *c-kit* transcripts (amplified with primers hit c-kit e20-21) after 1 day, 3 days and 5 days of RA induction respectively. (d) Expression dynamics of the full-length *c-kit* transcript during 5 days of RA induction. (e) Expression dynamics of all *c-kit* transcripts during 5 days of RA induction.

Dazl (germ cell-specific gene) and Plzf (SSCs marker gene) were not detectable before and 1 day after induction. Expression of Dazl and Plzf was gained after 2 more days of induction (day 3 and day 5) (Fig 3.13a). Expression of these 2 genes demonstrated that the P19 cells had lost their pluripotence and begun to differentiate into germ cell lineage cells. Expression of Dazl was almost stable from day 3 to day 5 except in the 100 nM RA induction group (Fig 3.13c). Expression of Plzf decresed significantly on day 5 as compared with that on day 3 (Fig 3.13b).

Expressions of the SSCs marker genes (Bmp4, Egr3 and Pou5f1) and RA responding genes (Cyp26b1 and Stra8) changed significantly. Expression of Bmp4 increased significantly after 1 day of induction and then decreased back to similar levels as in the control after 3-5 days of inductive culture (Fig 3.14a). Cyp26b1 was stimulated and showed a RA dose dependent pattery during the first 3 days of induction. However, this increase was slowed-down afterwards (Fig 3.14b). Egr3 responded to RA induction at 1 day of incubation and the response suddenly disappeared after 3 days and 5 days of incubation (Fig 3.14c). Changes of Pou5f1 were not significant in the whole process of induction (Fig 3.14d). Stra8 was stimulated on day 1 but decreased at day 3 and day 5 of induction (Fig 3.14e).

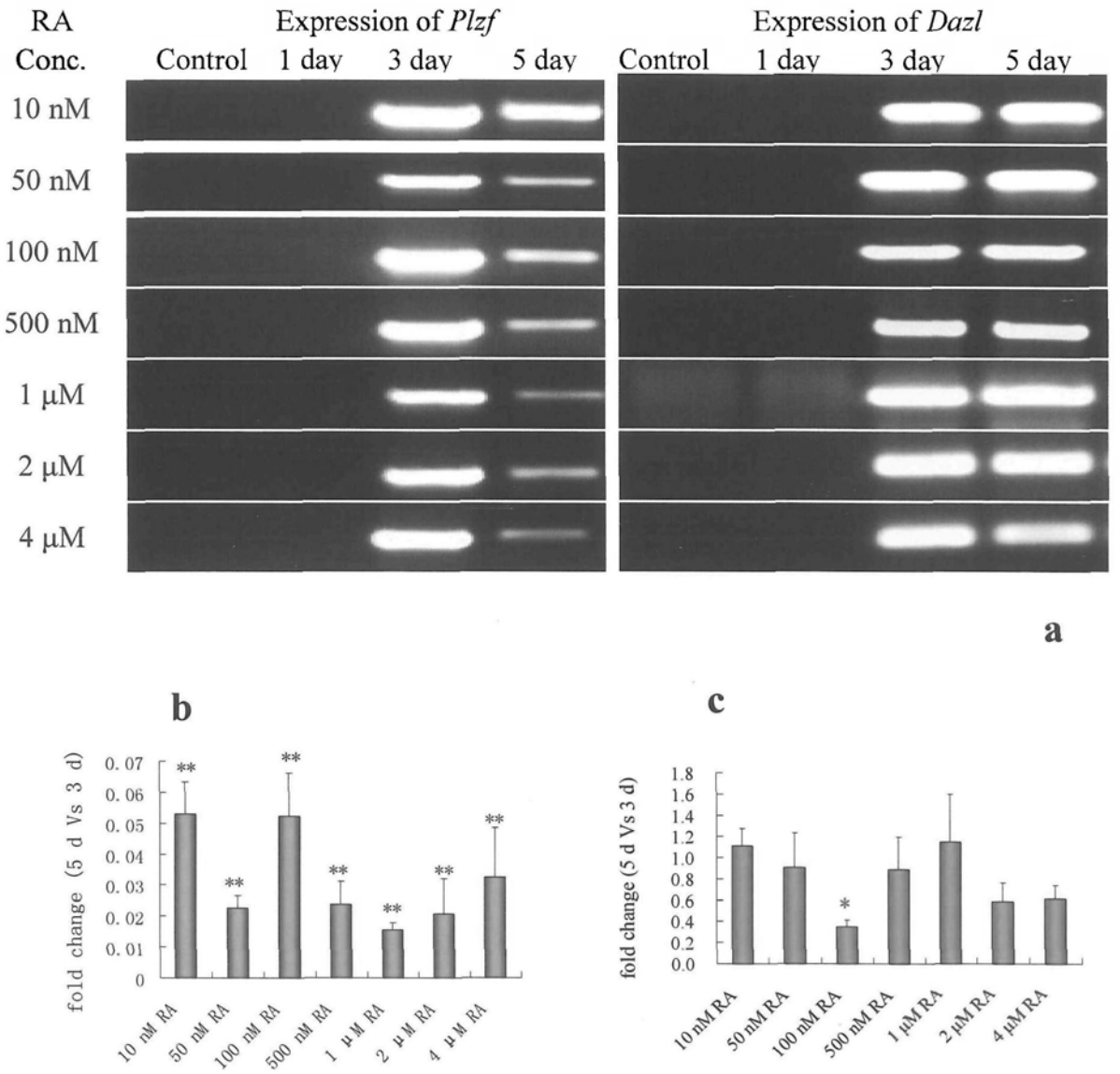
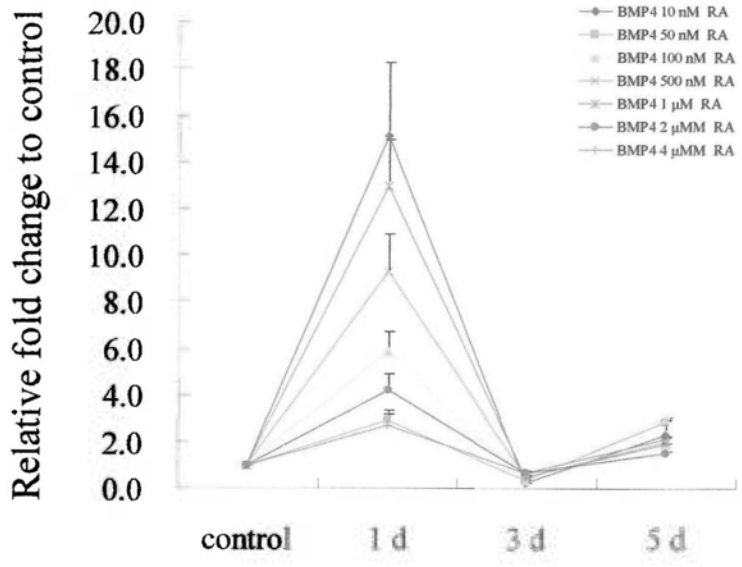
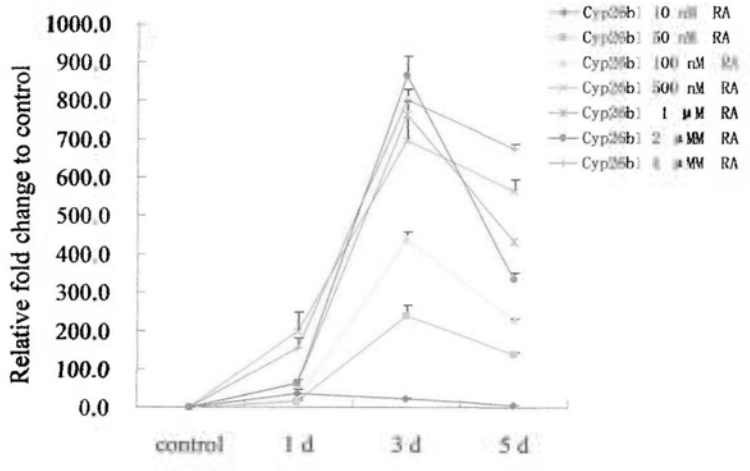


Figure 3.13 Expression dynamics of germ cell-related genes *Plzf* and *Dazl* in P19 cells stimulated by RA. P19 cells were cultured with different RA concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1 μM, 2μM and 4μM) for 1, 3 and 5 days respectively. Expression of *Plzf* and *Dazl* was quantified by real-time PCR. (a) Real-time PCR of *Plzf* and *Dazl* in P19 cells before and after RA induction. (b) Comparison of expression of *Plzf* between 5 days and 3 days after RA induction. (c) Comparison of expression of *Dazl*between 5 days and 3 days after RA induction.

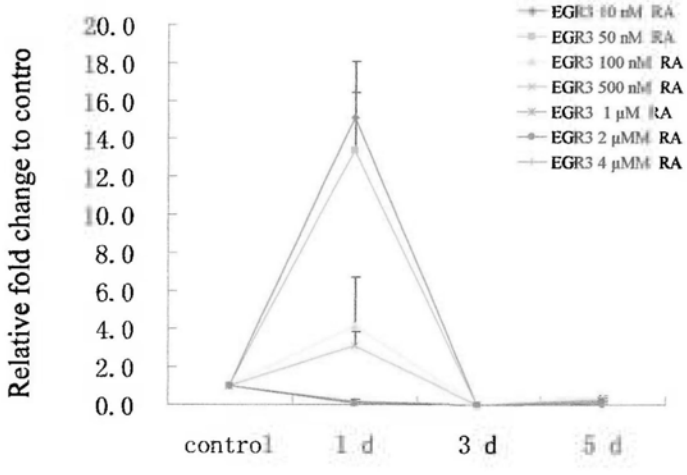
a


Relative fold change to control
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
control
1 d
3 d
5 d
BMP4 10 nM RA
BMP4 50 nM RA
BMP4 100 nM RA
BMP4 500 nM RA
BMP4 1 μM RA
BMP4 2 μMM RA
BMP4 4 μMM RA

b


Relative fold change to control
1000.0
900.0
800.0
700.0
600.0
500.0
400.0
300.0
200.0
100.0
0.0
control
1 d
3 d
5 d

c


Relative fold change to contro
20.0
18.0
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0
control
1 d
3 d
5 d
EGR3 50 nM RA
EGR3 100 nM RA
EGR3 500 nM RA
EGR3 1 μM RA
EGR3 2 μMM RA
EGR3 4 μMM RA

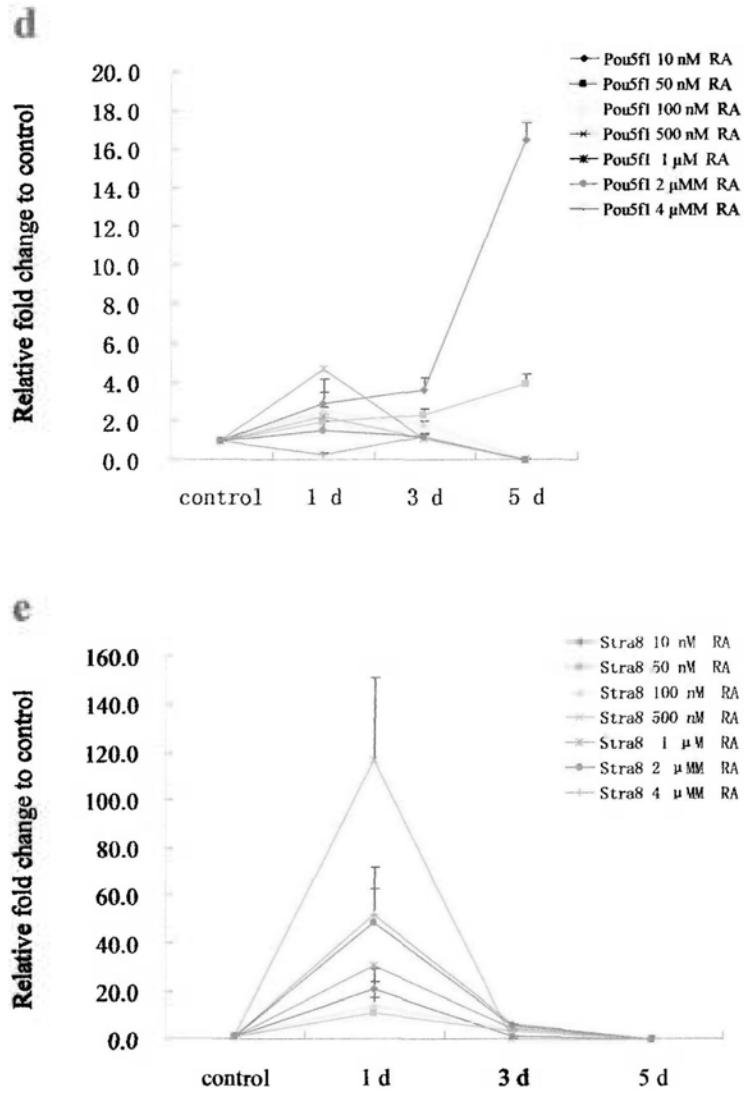


Fig 3.14. Expression dynamics of stem cell-related genes in P19 cells stimulated by RA.

P19 cells were cultured with different RA concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1 μM, 2μM and 4μM) for 1, 3 and 5 days respectively. Expression of stem cell-related genes (*Bmp4*, *Cyp26b1*, *Egr3*, *Pou5f1* and *Stra8*) was quantified by real-time PCR. Values of the vertical axis represented the expression fold change comparing with the control group. (a) Expression changes of *Bmp4* during 5 days of RA induction. (b) Expression changes of *Cyp26b1* during 5 days of RA induction. (c) Expression changes of *Egr3* during 5 days of RA induction. (d) Expression changes of *Pou5f1* during 5 days of RA induction. (e) Expression changes of *Stra8* during 5 days of RA induction.

Chapter 4

Discussion

4.1 The developmental stages of the cell lines and testes represented in this study.

In the testes of a mature male mouse, the male germ cells can be subgrouped into spermatogonia, spermatocyte, spermatids and spermatozoa according to their morphology. Two types of spermatogonia, the undifferentiated and differentiating ones, can be differentiated by their expression of specific markers and their abilities to regenerate spermatogenesis after transplantation to the deficient recipient testicular tubules. However, there is no way to separate the undifferentiated and the differentiating spermatogonia from the testes completely so far. It is difficult to define the biological characters of the undifferentiated and the differentiating spermatogonia when they are in the testis where complicated cell-cell conections are existed. Because the male germ cell lines are consistent cell populations with clear origination, the characters of the original cell population, which usually difficult to isolate, are truthfully represented in these cell lines. They are easier to maintain in the mean time. Therefore, in this study, we have selected several classic male germ cell lines together with different age mouse testes as the study objects. The cell line c18-4 and 5 dpp mouse testes (before the initiation of spermatogonia differentiation) are regarded as the undifferentiated spermatogonia group. Whereas CRL-2053 and 10 dpp mouse testes (after the initiation of spermatogonia differentiation) are regarded as the differentiating spermatogonia group. 60 dpp mouse testes represent a mixture

of the undifferentiated, the differentiating and other maturing and matured germ cells.

4.2 Expression profiles of *c-kit* in the undifferentiated and the differentiating spermatogonia

Previous studies on the expression and function of Kit in the undifferentiated SSCs are contradictory. Some studies have showed that Kit expression in the adult testis is detected by immunohistochemical analysis and in situ hybridization in the differentiating spermatogonia (type A1–A4, intermediate, and type B), as well as in the preleptotene spermatocytes and interstitial Leydig cells, but not in the undifferentiated spermatogonia and Sertoli cells (Yoshinaga et al，1991；Schrans-Stassen et al，1999). Some functional studies have showed that activation of the Kit/Kitl signaling pathway is not required for SSCs self-renewal (Kubota et al，2009；Morimoto et al，2009). However, more recent studies have demonstrated that both Kit^- and Kit^+ cells showed comparable levels of stem cell activity after germ cell transplantation (Barroca et al，2009；Morimoto et al，2009；Trefil et al，2010). As SSCs can change their phenotype according to their microenvironment, the Kit^+ SSCs are hypothesized to be at an intermediate state during the process of their self-renewal (Morimoto et al，2009). Izadyar et al. have characterized the Kit^+ SSCs and found that the $POU5F1^+/Kit^+$ subset of mouse SSCs generates cell lines expressed the pluripotent ES markers and could differentiate into multiple lineages *in vitro*. However, only the $POU5F1^+/Kit^-$ SSCs could regenerate the spermatogenesis

in the recipient tests after *in vivo* transplantation (Izadyar et al，2008). It seems that Kit cannot directly affect SSCs self-renewal ability but affect the size of SSCs pool by phenotypic transition.

Synthesis of *c-kit* mRNA and protein is concordant with the first appearance of differentiating spermatogonia (Prabhu et al，2006). The presence of Kit in the spermatogonia has been routinely regarded as a marker of the differentiating spermatogonia (Shinohara et al，1999, 2000). Function of *c-kit* gene during this stage is considered to be preparing these differentiating spermatogonia to go into meiosis.

Taken together, *c-kit* plays important roles in both the undifferentiated and the differentiating spermatogonia. Thus, it is important to understand the true expression profile changes during this transition. In this study, we have found that the undifferentiated spermatogonia express the full-length *c-kit* transcript, a SSCs-specific transcript and a group of truncated *c-kit* transcripts. The SSCs-specific transcript starts from intron 9 and included all the following exons downstream as the full-length *c-kit* transcript with a possible promoter located in intron 9. A 502 aa long protein might be produced by this transcript as predicted by ORF finder. The 502 aa protein was different from the full-length Kit on the extracellular domain. The full-length *c-kit* transcript is also detectable in the c18-4 cells and the 5 dpp testes whereas Kit was not present on the germ cells. This phenomenon has also been reported by early studies. However the mechanism is still unclear by far. One hypothesis is that *c-kit* mRNA transcription and translation are regulated individually by different micorenvironmental factors (Prabhu et al，2006).

In the differentiating spermatogonia, a shorter transcript with a truncated 3′ UTR is found in which is not expressed in the undifferentiated spermatogonia. This 4-kb-long short 3′ UTR transcript is composed by 21 exons exactly the same as the full-length transcript. The difference between them locates in the 3′ UTR, where the 4 kb transcript is 1.2 kb shorter than the full-length one. Functions of 3′ UTR involve supplying binding site for microRNAs and post-transcriptional regulation. We hypothesized that the shorter 3′ UTR transcript may lose some binding sites for the testes specific microRNAs and, as a result, Kit is more easily to be produced. The discovery of the short 3′ UTR transcript give us some clues to explain the phenomenon why the full-length *c-kit* transcript is expressed in the undifferentiated spermatogonia but the Kit as a protein is not. As the 4 kb shorter 3′ UTR transcript is only expressed in the differentiating spermatogonia, and most importantly, its presence has been confirmed in the normal testis tissues, this transcript can be used as a marker of differentiating spermatogonia. Its roles in SSCs differentiation need to be explored further.

Except the full-length protein form of Kit, a truncated form with a molecular weight around 50 kDa is expressed during spermatogenesis by protein analysis of the cell lines and the testes. The full-length form and the truncated form of protein show different expression pattern during germ cells development. The full-lenth form is expressed only in the differentiating spermatogonia and pre-meiosis spermatocytes. Therefore, its function may be preparing the germ cells to go into meiosis. The truncated form is expressed profoundly in the germ cells including the spermatogonia

(both undifferentiated and differentiating), spermatocyte and spermatids. The function of the truncated Kit may be keeping these cells proliferate.

Compare with the undifferentiated spermatogonia, expression of *c-kit* mRNA and protein in the differentiating spermatogonia increases significantly. For the first time, we have found that not only the quantity but also the component of *c-kit* mRNA and protein change during spermatogonia differentiation. These alterations may contribute the initiation of spermatogonia differentiation, which is valuable for the educidation of germ cell differentiation mechanisms.

4.3 Effect of RA on spermatogonia differentiation and *c-kit* expression.

Expression profiles of *c-kit* before and after spermatogonia differentiation are different according to our studies. Regulatory mechanisms controlling the transition are unclear. In early studies, RA, BMP4 and FSH are showed to be potential regulatory factors for this transition. RA, an active metabolite of vitamin A, is a vital signaling molecule for normal fetal development, pattern formation, cell proliferation, differentiation and apoptosis (Livera et al，2002). RA is considered to be crucial for germ cells to enter meiosis in both male and female (Bowles et al，2006；Koubova et al，2006). Testes of adult vitamin A-deficient mice/rat exhibit seminiferous tubules containing only Sertoli cells, type A spermatogonia and few preleptotene spermatocytes (Morales & Griswold, 1987). The type A spermatogonia are arrested before differentiation (before A1 stage spermatogonia) with a reduced *c-kit* expression and no Stra8 expression. Administration of vitamin A to these animals

results in a synchronized spermatogenesis emerging from type A spermatogonia and an enhanced expression of *c-kit* (van and de Rooij DG，1991). Therefore, RA is a key regulatory factor for spermatogonial differentiation and *c-kit* expression. In this study, *in vitro* tissue culture (testis tissue culture), germ cells culture and multipotential cells culture are used to investigate RA's function on spermatogonia differentiation and *c-kit* expression. RA is set as a stimulus for spermatogonia differentiation. *c-kit* is set as the focus of the transition. The differentiation-related genes including *Bmp4* (a SSCs pluripotential maintance gene), *Cyp26b1* (a RA degradation gene), *Dazl* (a germ cells marker gene), *Egr3* (an early growth response gene), *Kitl* (Kit ligand), *Rarα* (an RA receptor gene) and *Stra8* (gene stimulated by RA) were set as potential participating factors.

Testes tissue culture shows that RA could enhance *c-kit* expression. *Bmp4* is down-regulated by RA which confirms that RA can induce the differentiation of SSCs. Expression of *Rarα* and *Egr3* are not affected by RA in the postnatal testes. So, these two genes may play a parallel role with RA in SSCs differentiation. *Cyp26b1* and *Stra8* are significantly up-regulated. The function of *Cyp26b1* is to degrade excessive RA and maintain the stability of RA concentration in the male germ cell growth environment. As the exogenous RA exceeds the degradability of the *Cyp26b1* gene, expression of *Stra8* was significantly enhanced. *Stra8* may be a partner gene with *c-kit* during spermatogonial differentiation. Effect of RA on DAZL and Kitl expression is minor.

As the testis is a mixture of the germ cells at different developmental stages, the

Sertoli cells and the leydig cells, it is difficult to determine which cells population has responded to RA stimulation. Our results from RA treated germ cell line have illustrated a better understanding of the relationship between RA, *c-kit* expression and spermatogonial differentiation. It is demonstrated that RA can directively affect spermatogonial germ cells. Expression pattern of *c-kit, Pouf51* and *Egr3* in the spermatogonial germ cells are different from that in the testes after RA stimulation. We hypothesize that RA can either works directly or indirectly on *c-kit* expression in the male germ cells. Besides RA, a significant increase of *c-kit* in the differentiating spermatogonia may need other factors from the testis. RA at an appropriate concentration is essential for spermatogonial differentiation, however, too much exogenous RA will push SSCs go into abnormal differentiation and apoptosis (Snyder et al, 2011). Our results have showed that after 24 hours of 2 μM RA treatment, expression of *Cyp26b1* gene is highly stimulated which will degrade RA into the inactive form. This verifies the assumption that the testis environment tends to protect the spermatogonia from detrimental concentrations of RA.

4.4 RA can induce P19 cells towards germ cells *in vitro*.

P19 cells are a line of pluripotent embryonal carcinoma that is able to grow continuously in serum-supplemented media. Like other embryonal carcinoma and embryonic stem cells, P19 cells are developmentally pluripotent and appear to differentiate using the same mechanisms as normal embryonic cells. P19 cells are easy to grow and maintain in the undifferentiated state but they can also be

efficiently induced to differentiate by simple manipulation of the culture conditions. The genetic composition of the cells can be easily manipulated either by the selection of mutant strains or by selection of clones carrying transfected genes stably integrated into their genomes (McBurney, 1993). Obviously, P19 cells are valuable tools for developmental biology studies. Retinoic acid can effectively induce P19 cells towards the neurons, astroglia and microglia which are cell derived from the neuroectoderm (Jones-Villeneuve et al, 1982; Edwards and McBurney, 1983; Andrews, 1984; McBurney, 1993). Whether P19 can be induced into germ cells is still unknown. RA is a well known PGCs inducer and have successfully induced the ES cells towards PGCs (Vogel, 2003; Eguizabal et al, 2009). RA has also been successfully used to induce ES cells into functional spermatids expressing *Stra8* and *Mvh* (Nayernia et al, 2006). In our study, different RA concentrations have been used to induce P19 cells. Germ cell specific gene expression is checked on 1 day, 3 days and 5 days after the induction. Our results indicate that, after 3 days induction, the morphology of P19 starts to change and the germ cell-specific genes *Dazl* and *Plzf* begin to be expressed. In contrast, the stem cell-specific genes *Pou5f1* and *Egr3* have been lost. Expression of *Dazl* is kept steady during 3-5 days after induction. After 5 days of induction and 9 days of culture without RA, the germ cell-like cells can be observed. During the induction process, expression of *c-kit* begins to increase after 1 day of culture when *Dazl* and *Plzf* are still not detectable. So, *c-kit* seems to be an more up-stream factor than *Dazl* and *Plzf* during early stage of germ cell development. Expression of *c-kit* in the high concentration RA treated groups (50 nM

– 4 μM) starts to decline after 3 days of induction. This decline is more obvious after 5 days of induction. The wave-like change of *c-kit* (increase at first and then decrease) is consistent with that occurred during *in vivo* germ cell development from ES (c-kit^{low}) to PGCs (c-kit^{high}) and to SSCs (c-kit^{low}). So, we have proved that, for the first time, RA can induce the pluripotential embryonic carcinoma cells to develop towards male germ cells, at least until the stage of spermatogonia or oogonia.

Chapter 5

Conclusions

This is a preliminary study on the role of *c-kit* in spermatogonial stem cells differentiation. According to this study, following conclusions are made:

1. Before and after commitment of differentiation, spermatogonia express different sets of *c-kit* transcripts. SSCs before differentiation express the full-length, the SSCs-specific and the Tr-kit transcripts. The differentiating SSCs express the full-length, the short 3′ UTR and the Tr-kit transcripts.

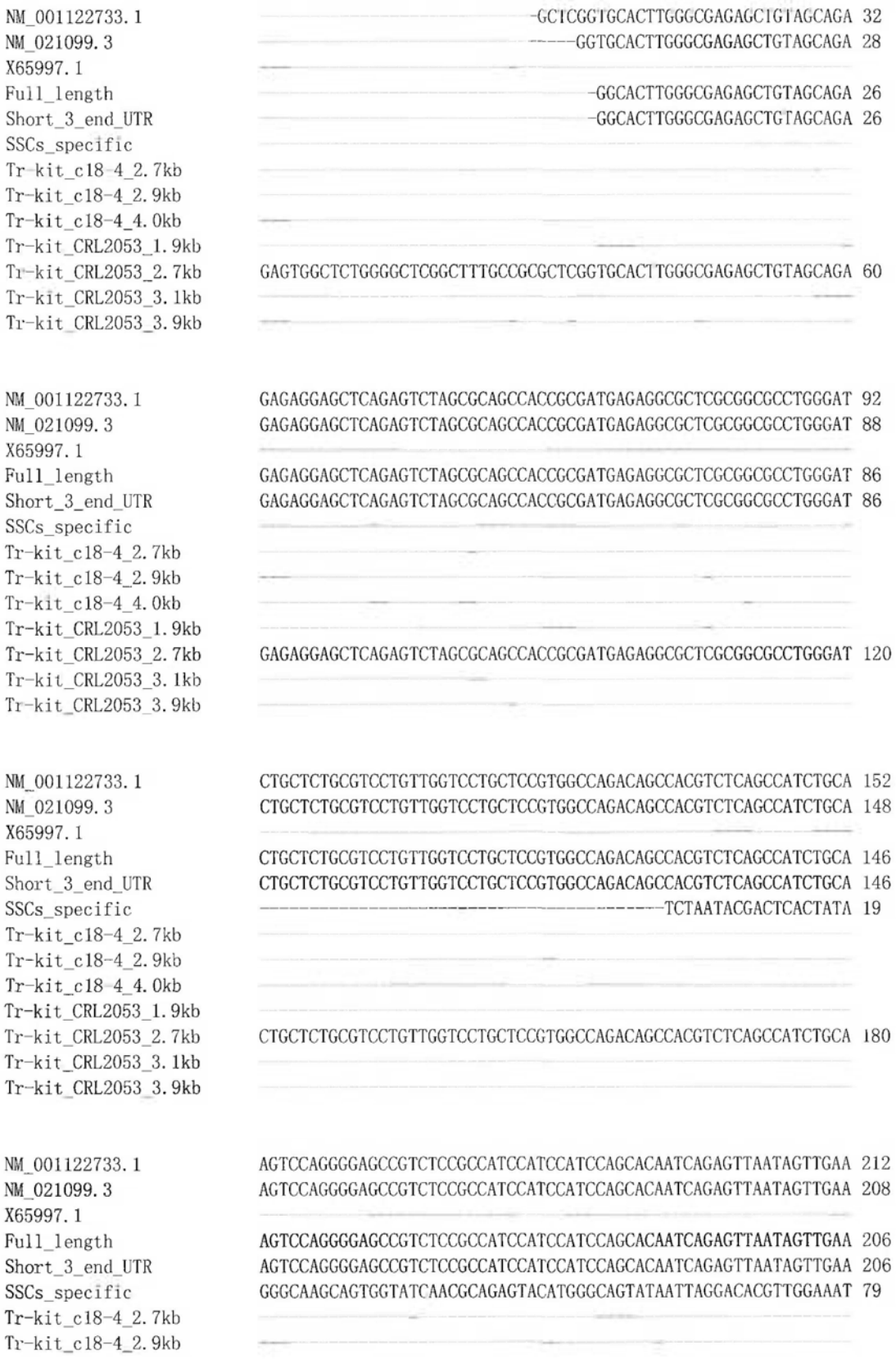
2. Kit is expressed in the differentiating but not in the undifferentiated SSCs. A new Tr-Kit (50 kDa) is expressed in both undifferentiated SSCs and differentiating spermatogonia.

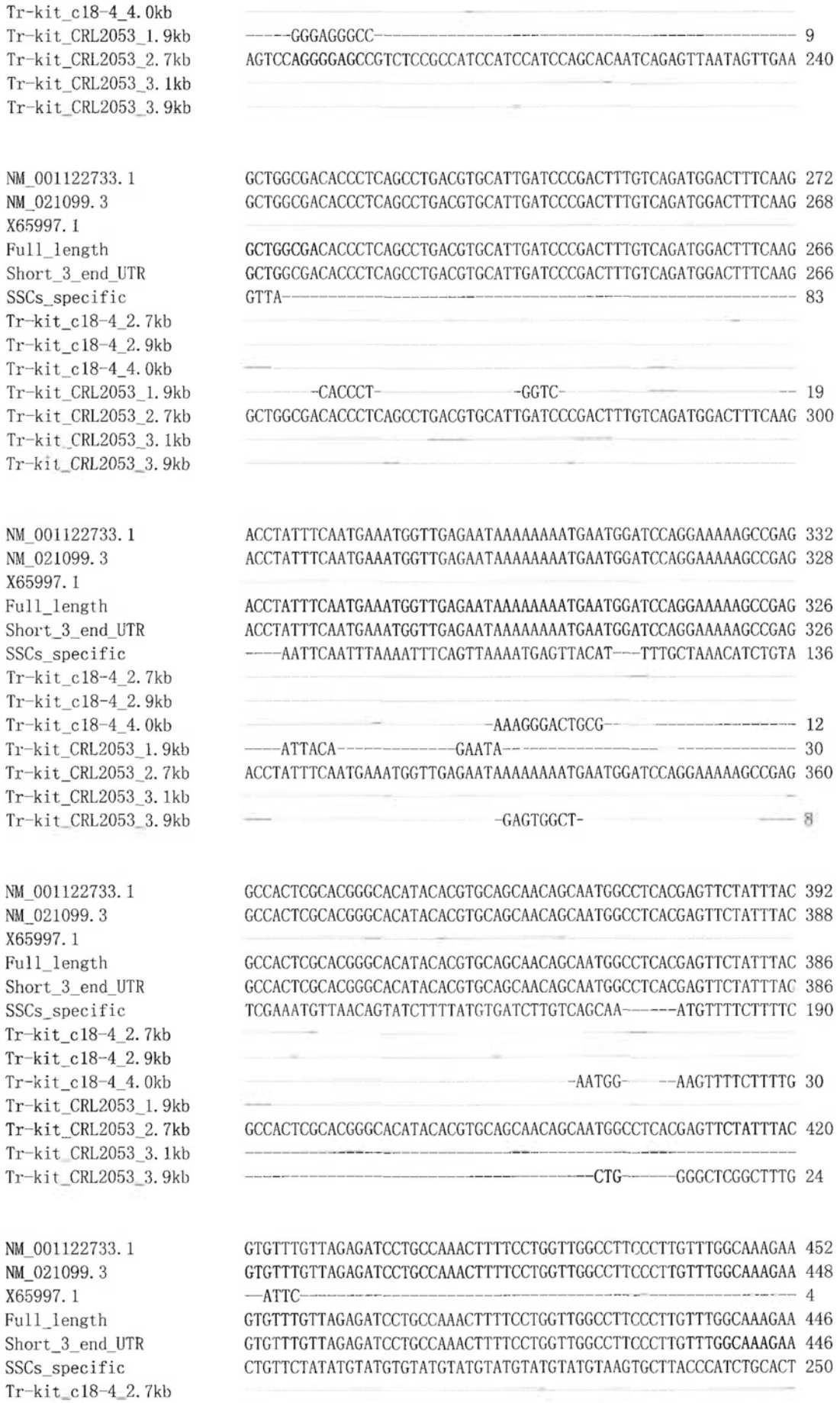
3. RA can regulate *c-kit* expression in the spermatogonial cells both directly and indirectly. RA directly works on the germ cells and decreases their *c-kit* expression, especially the full-length transcript. However, RA and other unknown factors in the testis can work together to increase *c-kit* expression. As an inducing factor, RA seems to work through somatic cells to stimulate *c-kit* expression in SSCs rather than work directly on SSCs during spermatogenesis.

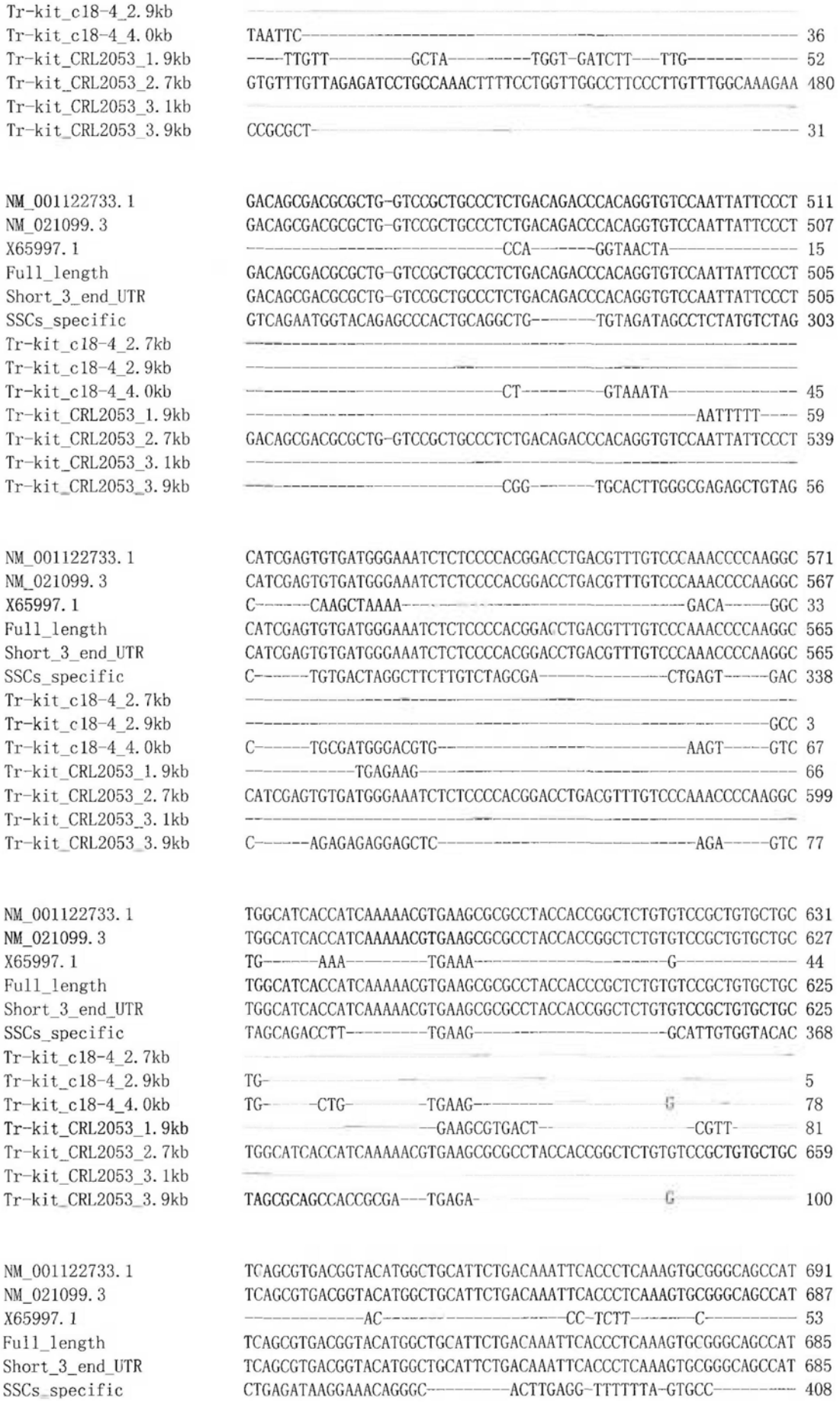
4. RA can induce the pluripotential embryonic carcinoma cells (P19 cells) to lose their characteristics of stem cells and develop towards germ cells *in vitro*. The wave-like change of *c-kit* expression during the induction process is similar to that occurred from ES cells to PGCs and to SSCs *in vivo*.

Appendix

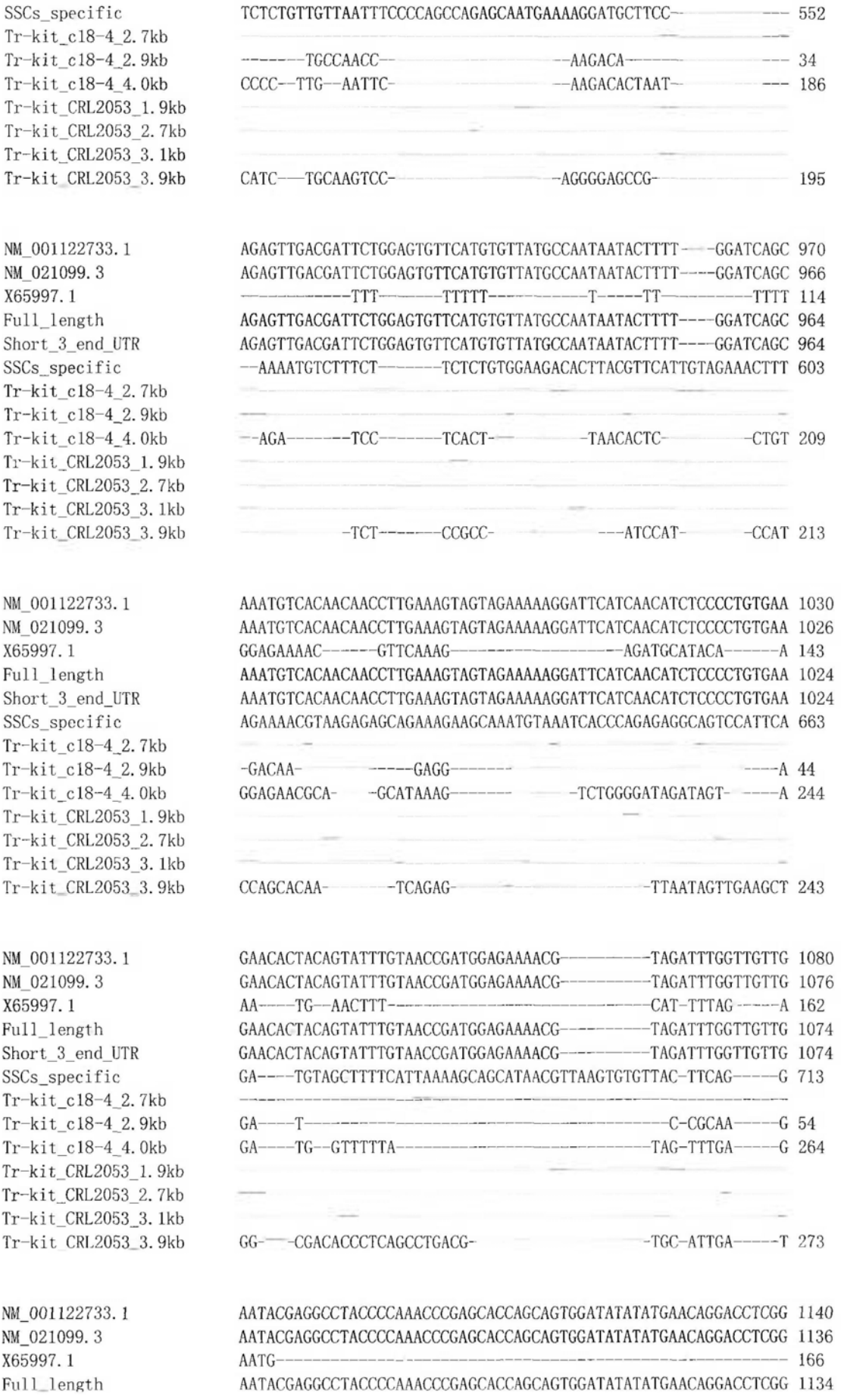
Appendix 1 Multiple sequence alignment of *c-kit* transcripts

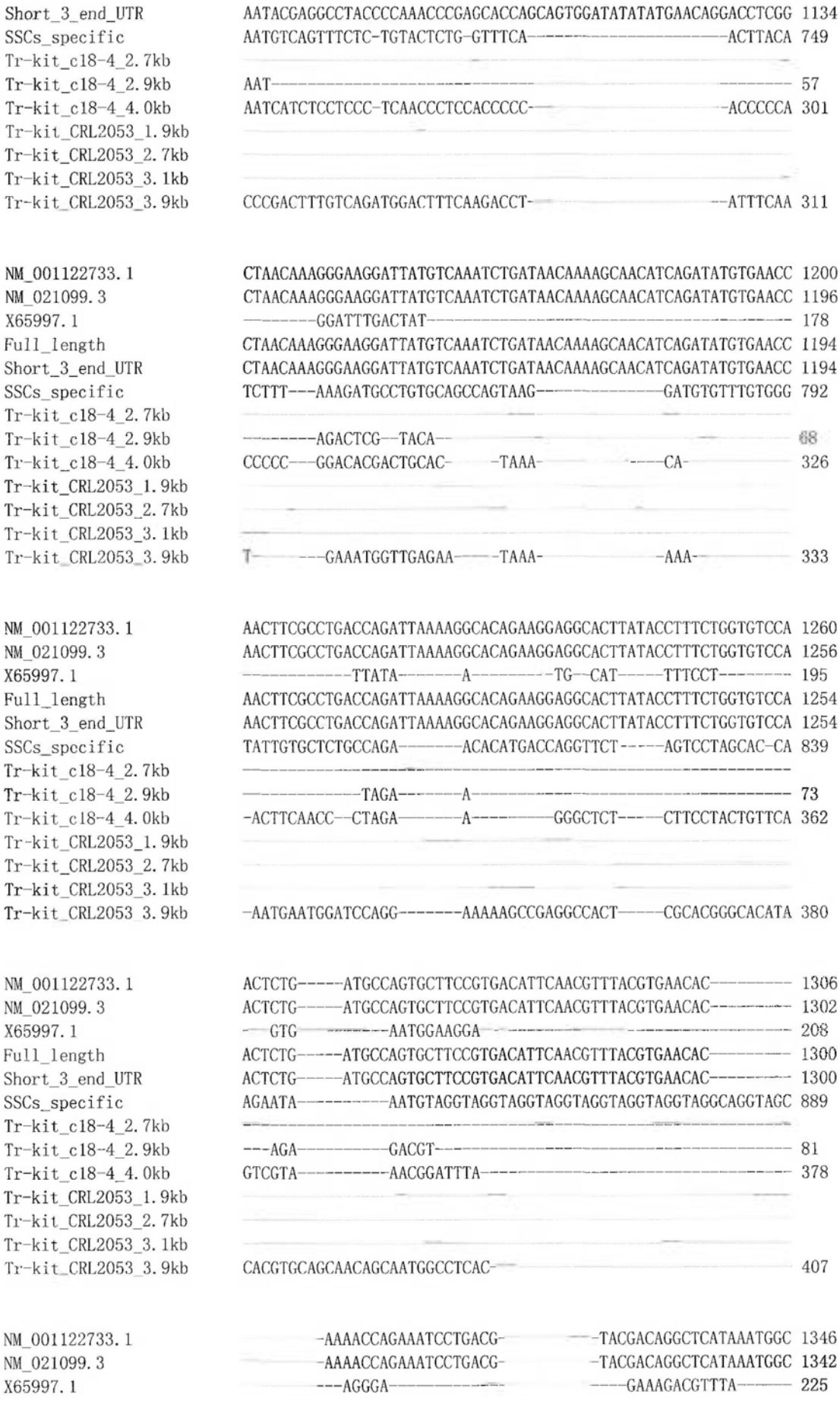


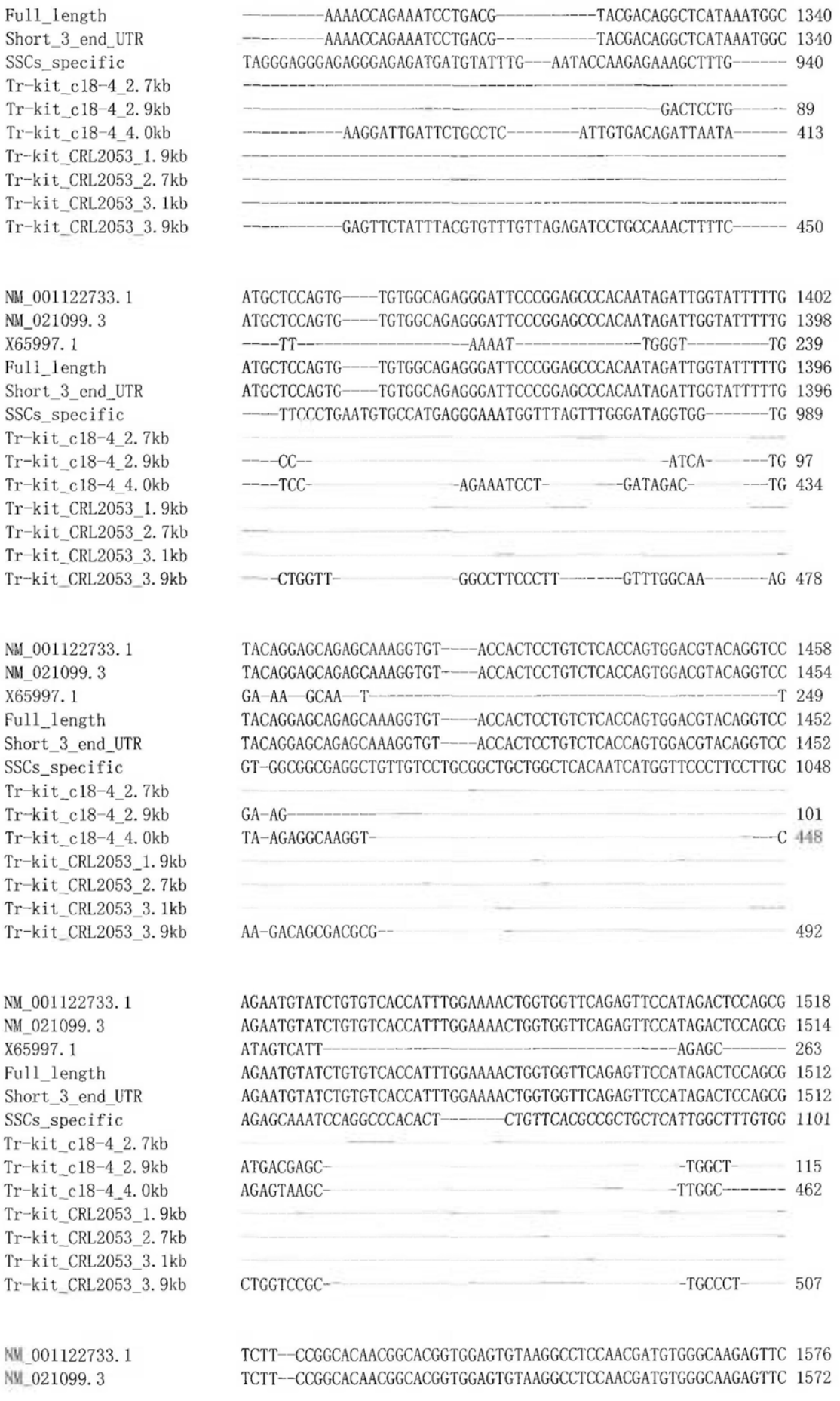

Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb -----GGGAGGGCC---------- 9
Tr-kit_CRL2053_2.7kb AGTCCAGGGGAGCCGTCTCCGCCATCCATCCATCCAGCACAATCAGAGTTAATAGTTGAA 240
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NM_001122733.1 GCTGGCGACACCCTCAGCCTGACGTGCATTGATCCCGACTTTGTCAGATGGACTTTCAAG 272
NM_021099.3 GCTGGCGACACCCTCAGCCTGACGTGCATTGATCCCGACTTTGTCAGATGGACTTTCAAG 268
X65997.1
Full_length GCTGGCGACACCCTCAGCCTGACGTGCATTGATCCCGACTTTGTCAGATGGACTTTCAAG 266
Short_3_end_UTR GCTGGCGACACCCTCAGCCTGACGTGCATTGATCCCGACTTTGTCAGATGGACTTTCAAG 266
SSCs_specific GTTA---------- 83
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb -CACCCT--------GGTC---------- 19
Tr-kit_CRL2053_2.7kb GCTGGCGACACCCTCAGCCTGACGTGCATTGATCCCGACTTTGTCAGATGGACTTTCAAG 300
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NM_001122733.1 ACCTATTTCAATGAAATGGTTGAGAATAAAAAAAATGAATGGATCCAGGAAAAAGCCGAG 332
NM_021099.3 ACCTATTTCAATGAAATGGTTGAGAATAAAAAAAATGAATGGATCCAGGAAAAAGCCGAG 328
X65997.1
Full_length ACCTATTTCAATGAAATGGTTGAGAATAAAAAAAATGAATGGATCCAGGAAAAAGCCGAG 326
Short_3_end_UTR ACCTATTTCAATGAAATGGTTGAGAATAAAAAAAATGAATGGATCCAGGAAAAAGCCGAG 326
SSCs_specific ----AATTCAATTTAAAATTTCAGTTAAAATGAGTTACAT---TTTGCTAAACATCTGTA 136
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb -AAAGGGACTGCG---------- 12
Tr-kit_CRL2053_1.9kb ----ATTACA----------GAATA---------- 30
Tr-kit_CRL2053_2.7kb ACCTATTTCAATGAAATGGTTGAGAATAAAAAAAATGAATGGATCCAGGAAAAAGCCGAG 360
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb -GAGTGGCT- 8
NM_001122733.1 GCCACTCGCACGGGCACATACACGTGCAGCAACAGCAATGGCCTCACGAGTTCTATTTAC 392
NM_021099.3 GCCACTCGCACGGGCACATACACGTGCAGCAACAGCAATGGCCTCACGAGTTCTATTTAC 388
X65997.1
Full_length GCCACTCGCACGGGCACATACACGTGCAGCAACAGCAATGGCCTCACGAGTTCTATTTAC 386
Short_3_end_UTR GCCACTCGCACGGGCACATACACGTGCAGCAACAGCAATGGCCTCACGAGTTCTATTTAC 386
SSCs_specific TCGAAATGTTAACAGTATCTTTTATGTGATCTTGTCAGCAA------ATGTTTTCTTTTC 190
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb -AATGG----AAGTTTTCTTTTG 30
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb GCCACTCGCACGGGCACATACACGTGCAGCAACAGCAATGGCCTCACGAGTTCTATTTAC 420
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb ----------CTG-------GGGCTCGGCTTTG 24
NM_001122733.1 GTGTTTGTTAGAGATCCTGCCAAACTTTTCCTGGTTGGCCTTCCCTTGTTTGGCAAAGAA 452
NM_021099.3 GTGTTTGTTAGAGATCCTGCCAAACTTTTCCTGGTTGGCCTTCCCTTGTTTGGCAAAGAA 448
X65997.1 --ATTC---------- 4
Full_length GTGTTTGTTAGAGATCCTGCCAAACTTTTCCTGGTTGGCCTTCCCTTGTTTGGCAAAGAA 446
Short_3_end_UTR GTGTTTGTTAGAGATCCTGCCAAACTTTTCCTGGTTGGCCTTCCCTTGTTTGGCAAAGAA 446
SSCs_specific CTGTTCTATATGTATGTGTATGTATGTATGTATGTAAGTGCTTACCCATCTGCACT 250
Tr-kit_c18-4_2.7kb


Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb TAATTC 36
Tr-kit_CRL2053_1.9kb ----TTGTT GCTA TGGT-GATCTT TTG 52
Tr-kit_CRL2053_2.7kb GTGTTTGTTAGAGATCCTGCCAAACTTTTCCTGGTTGGCCTTCCCTTGTTTGGCAAAGAA 480
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CCGCGCT 31
NM_001122733.1 GACAGCGACGCGCTG-GTCCGCTGCCCTCTGACAGACCCACAGGTGTCCAATTATTCCCT 511
NM_021099.3 GACAGCGACGCGCTG-GTCCGCTGCCCTCTGACAGACCCACAGGTGTCCAATTATTCCCT 507
X65997.1 CCA GGTAACTA 15
Full_length GACAGCGACGCGCTG-GTCCGCTGCCCTCTGACAGACCCACAGGTGTCCAATTATTCCCT 505
Short_3_end_UTR GACAGCGACGCGCTG-GTCCGCTGCCCTCTGACAGACCCACAGGTGTCCAATTATTCCCT 505
SSCs_specific GTCAGAATGGTACAGAGCCCACTGCAGGCTG TGTAGATAGCCTCTATGTCTAG 303
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb CT GTAAATA 45
Tr-kit_CRL2053_1.9kb AATTTTT 59
Tr-kit_CRL2053_2.7kb GACAGCGACGCGCTG-GTCCGCTGCCCTCTGACAGACCCACAGGTGTCCAATTATTCCCT 539
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CGG TGCACTTGGGCGAGAGCTGTAG 56
NM_001122733.1 CATCGAGTGTGATGGGAAATCTCTCCCCACGGACCTGACGTTTGTCCCAAACCCCAAGGC 571
NM_021099.3 CATCGAGTGTGATGGGAAATCTCTCCCCACGGACCTGACGTTTGTCCCAAACCCCAAGGC 567
X65997.1 C CAAGCTAAAA GACA GGC 33
Full_length CATCGAGTGTGATGGGAAATCTCTCCCCACGGACCTGACGTTTGTCCCAAACCCCAAGGC 565
Short_3_end_UTR CATCGAGTGTGATGGGAAATCTCTCCCCACGGACCTGACGTTTGTCCCAAACCCCAAGGC 565
SSCs_specific C TGTGACTAGGCTTCTTGTCTAGCGA CTGAGT GAC 338
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb GCC 3
Tr-kit_c18-4_4.0kb C TGCGATGGGACGTG AAGT GTC 67
Tr-kit_CRL2053_1.9kb TGAGAAG 66
Tr-kit_CRL2053_2.7kb CATCGAGTGTGATGGGAAATCTCTCCCCACGGACCTGACGTTTGTCCCAAACCCCAAGGC 599
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb C AGAGAGAGGAGCTC AGA GTC 77
NM_001122733.1 TGGCATCACCATCAAAAACGTGAAGCGCGCCTACCACCGGCTCTGTGTCCGCTGTGCTGC 631
NM_021099.3 TGGCATCACCATCAAAAACGTGAAGCGCGCCTACCACCGGCTCTGTGTCCGCTGTGCTGC 627
X65997.1 TG AAA TGAAA G 44
Full_length TGGCATCACCATCAAAAACGTGAAGCGCGCCTACCACCCGCTCTGTGTCCGCTGTGCTGC 625
Short_3_end_UTR TGGCATCACCATCAAAAACGTGAAGCGCGCCTACCACCGGCTCTGTGTCCGCTGTGCTGC 625
SSCs_specific TAGCAGACCTT TGAAG GCATTGTGGTACAC 368
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb TG 5
Tr-kit_c18-4_4.0kb TG CTG TGAAG G 78
Tr-kit_CRL2053_1.9kb GAAGCGTGACT CGTT 81
Tr-kit_CRL2053_2.7kb TGGCATCACCATCAAAAACGTGAAGCGCGCCTACCACCGGCTCTGTGTCCGCTGTGCTGC 659
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb TAGCGCAGCCACCGCGA TGAGA G 100
NM_001122733.1 TCAGCGTGACGGTACATGGCTGCATTCTGACAAATTCACCCTCAAAGTGCGGGCAGCCAT 691
NM_021099.3 TCAGCGTGACGGTACATGGCTGCATTCTGACAAATTCACCCTCAAAGTGCGGGCAGCCAT 687
X65997.1 AC CC-TCTT C 53
Full_length TCAGCGTGACGGTACATGGCTGCATTCTGACAAATTCACCCTCAAAGTGCGGGCAGCCAT 685
Short_3_end_UTR TCAGCGTGACGGTACATGGCTGCATTCTGACAAATTCACCCTCAAAGTGCGGGCAGCCAT 685
SSCs_specific CTGAGATAAGGAAACAGGGC ACTTGAGG-TTTTTTA-GTGCC 408

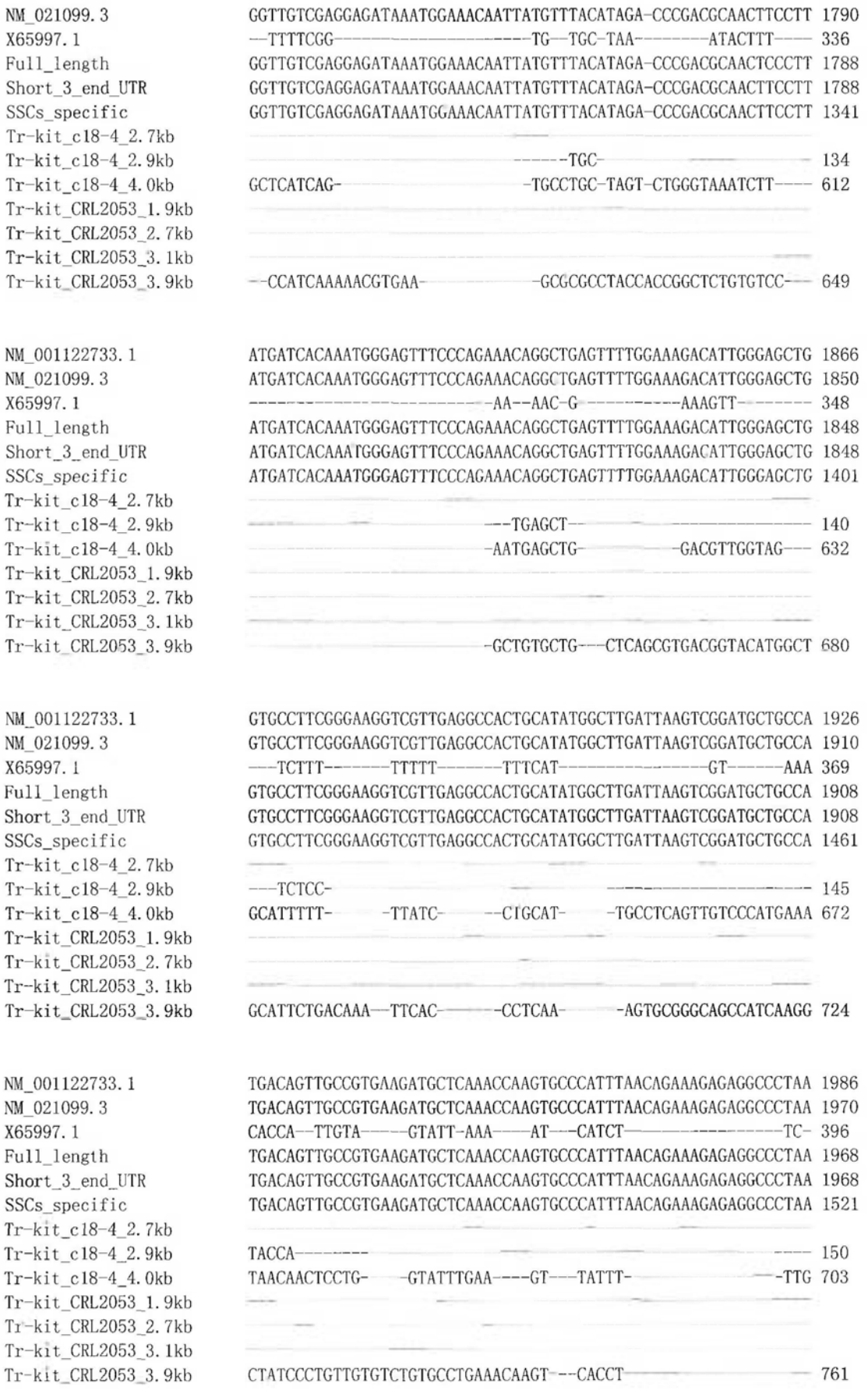
```
Tr-kit_c18-4_2.7kb             ------------------------------------------------------------
Tr-kit_c18-4_2.9kb             ------------------------------------GCG-TTTCCTA------------- 15
Tr-kit_c18-4_4.0kb             ---------------ACAAGCC-----------ACAGAGCC-TCTTTTA-CTGC------ 104
Tr-kit_CRL2053_1.9kb           --------------------------TATTTT-------------CTCAAAG-------- 94
Tr-kit_CRL2053_2.7kb           TCAGCGTGACGGTACATGGCTGCATTCTGACAAATTCACCCTCAAAGTGCGGGCAGCCAT 719
Tr-kit_CRL2053_3.1kb           ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb           ---------------GCG---C-----------TCGCGGCG-CCTGGGATCTGC------ 124

NM_001122733.1                 CAAGGCTATCCCTGTTGTGTCTGTGCCTGAAACAAGTCACCTCCTTAAGAAAGGGGACAC 751
NM_021099.3                    CAAGGCTATCCCTGTTGTGTCTGTGCCTGAAACAAGTCACCTCCTTAAGAAAGGGGACAC 747
X65997.1                       ---------CTTGTGT-------------------------------------------- 60
Full_length                    CAAGGCTATCCCTGTTGTGTCTGTGCCTGAAACAAGTCACCTCCTTAAGAAAGGGGACAC 745
Short_3_end_UTR                CAAGGCTATCCCTGTTGTGTCTGTGCCTGAAACAAGTCACCTCCTTAAGAAAGGGGACAC 745
SSCs_specific                  -----CAGTTTCTGTGGTGTTTA-----------------------------------GC 428
Tr-kit_c18-4_2.7kb             ------------------------------------------------------------
Tr-kit_c18-4_2.9kb             ------------------------------------------------------------
Tr-kit_c18-4_4.0kb             ---------TTTGTCTAGTTTTAG----------------------------------AC 121
Tr-kit_CRL2053_1.9kb           CAAG-------------------------------------------------------- 98
Tr-kit_CRL2053_2.7kb           CAAGGCTATCCCTGTTGTGTCTGTGCCTGAAACAAGTCACCTCCTTAAGAAAGGGGACAC 779
Tr-kit_CRL2053_3.1kb           ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb           ---------TCTGCGTCCTGTTGG-----------------------------------TC 141

NM_001122733.1                 ATTTACGGTGGTGTGCACCATAAAAGATGTGTCTACATCCGTGAACTCCATGTGGCTAAA 811
NM_021099.3                    ATTTACGGTGGTGTGCACCATAAAAGATGTGTCTACATCCGTGAACTCCATGTGGCTAAA 807
X65997.1                       -----CCTTGGG-----------AGA-A-------------------------------- 71
Full_length                    ATTTACGGTGGTGTGCACCATAAAAGATGTGTCTACATCCGTGAACTCCATGTGGCTAAA 805
Short_3_end_UTR                ATTTACGGTGGTGTGCACCATAAAAGATGTGTCTACATCCGTGAACTCCATGTGGCTAAA 805
SSCs_specific                  ATTGGCCATAGGACTCATGCCAAAA-A---GCTGTATATCCTCAGTTTCT---------G 475
Tr-kit_c18-4_2.7kb             ------------------------------------------------------------
Tr-kit_c18-4_2.9kb             ------CGTGG------------------------------------------------- 20
Tr-kit_c18-4_4.0kb             A---ACCCTAGATTTCA--CTTAGA-A---GTT------------AATTTTT-------A 153
Tr-kit_CRL2053_1.9kb           ------------------------------------------------------------
Tr-kit_CRL2053_2.7kb           ATTTACGGTGGTGTGCACCATAAAAGATGTGTCTACATCCGTGAACTCCATGTGGCTAAA 839
Tr-kit_CRL2053_3.1kb           ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb           CTGCTCCGTGG----------CCAGACA---GCC-------------------------A 163

NM_001122733.1                 GATGAACCCTCAGCCTCAGCACATAGCCCAGGTAAAGCACAATAGCTGGCACCGGGGTGA 871
NM_021099.3                    GATGAACCCTCAGCCTCAGCACATAGCCCAGGTAAAGCACAATAGCTGGCACCGGGGTGA 867
X65997.1                       GAC------------------GTCAAG--------------------------------- 80
Full_length                    GATGAACCCTCAGCCTCAGCACATAGCCCAGGTAAAGCACAATAGCTGGCACCGGGGTGA 865
Short_3_end_UTR                GATGAACCCTCAGCCTCAGCACATAGCCCAGGTAAAGCACAATAGCTGGCACCGGGGTGA 865
SSCs_specific                  GATGAGGAAGGGGACGCATCA--TAACTCTGC---------------------------- 505
Tr-kit_c18-4_2.7kb             ------------------------------------------------------------
Tr-kit_c18-4_2.9kb             ------------------------------------------------------------
Tr-kit_c18-4_4.0kb             GAT-------------------ACCCAGG------------------------------- 163
Tr-kit_CRL2053_1.9kb           ------------------------------------------------------------
Tr-kit_CRL2053_2.7kb           GATGAACCCTCAGCCTCAGCACATAGCCCAG----------------------------- 870
Tr-kit_CRL2053_3.1kb           ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb           CGT-------------------CTCAGC-------------------------------- 172

NM_001122733.1                 CTTCAATTATGAACGCC---------------AGGAGACGCTGACTATCAGCTCGGCA 914
NM_021099.3                    CTTCAATTATGAACGCC---------------AGGAGACGCTGACTATCAGCTCGGCA 910
X65997.1                       ------TTG--AAGCT------------------GTGAAACTTTT------------- 99
Full_length                    CTTCAATTATGAACGCC---------------AGGAGACGCTGACTATCAGCTCGGCA 908
Short_3_end_UTR                CTTCAATTATGAACGCC---------------AGGAGACGCTGACTATCAGCTCGGCA 908
```

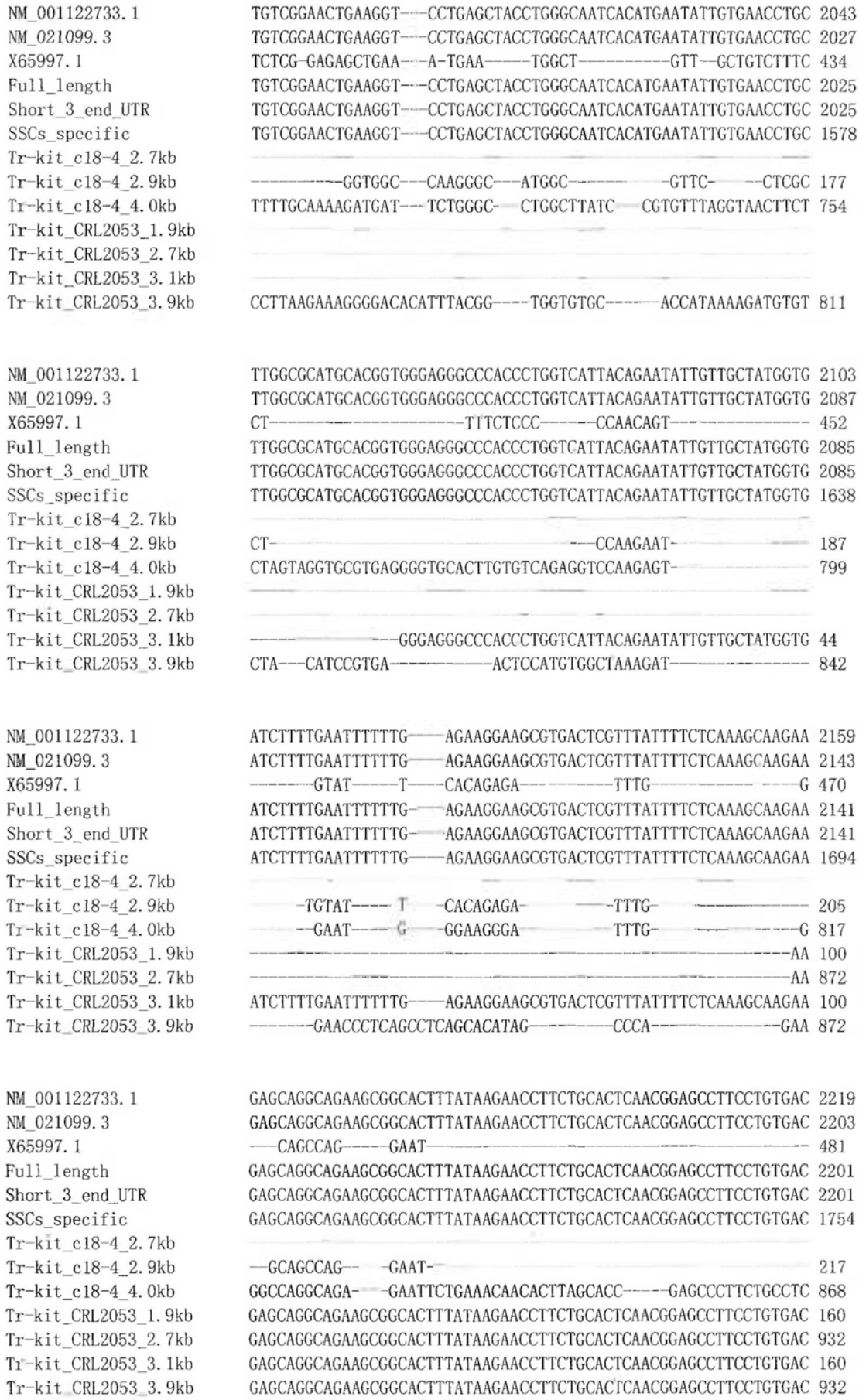

SSCs_specific TCTCTGTTGTTAATTTCCCCAGCCAGAGCAATGAAAAGGATGCTTCC 552
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb TGCCAACC AAGACA 34
Tr-kit_c18-4_4.0kb CCCC TTG AATTC AAGACACTAAT 186
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CATC TGCAAGTCC AGGGGAGCCG 195
NM_001122733.1 AGAGTTGACGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTT GGATCAGC 970
NM_021099.3 AGAGTTGACGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTT GGATCAGC 966
X65997.1 TTT TTTTT T TT TTTT 114
Full_length AGAGTTGACGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTT GGATCAGC 964
Short_3_end_UTR AGAGTTGACGATTCTGGAGTGTTCATGTGTTATGCCAATAATACTTTT GGATCAGC 964
SSCs_specific AAAATGTCTTTCT TCTCTGTGGAAGACACTTACGTTCATTGTAGAAACTTT 603
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb AGA TCC TCACT TAACACTC CTGT 209
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb TCT CCGCC ATCCAT CCAT 213
NM_001122733.1 AAATGTCACAACAACCTTGAAAGTAGTAGAAAAAGGATTCATCAACATCTCCCCTGTGAA 1030
NM_021099.3 AAATGTCACAACAACCTTGAAAGTAGTAGAAAAAGGATTCATCAACATCTCCCCTGTGAA 1026
X65997.1 GGAGAAAAC GTTCAAAG AGATGCATACA A 143
Full_length AAATGTCACAACAACCTTGAAAGTAGTAGAAAAAGGATTCATCAACATCTCCCCTGTGAA 1024
Short_3_end_UTR AAATGTCACAACAACCTTGAAAGTAGTAGAAAAAGGATTCATCAACATCTCCCCTGTGAA 1024
SSCs_specific AGAAAACGTAAGAGAGCAGAAAGAAGCAAATGTAAATCACCCAGAGAGGCAGTCCATTCA 663
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb GACAA GAGG A 44
Tr-kit_c18-4_4.0kb GGAGAACGCA GCATAAAG TCTGGGGATAGATAGT A 244
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CCAGCACAA TCAGAG TTAATAGTTGAAGCT 243
NM_001122733.1 GAACACTACAGTATTTGTAACCGATGGAGAAAACG TAGATTTGGTTGTTG 1080
NM_021099.3 GAACACTACAGTATTTGTAACCGATGGAGAAAACG TAGATTTGGTTGTTG 1076
X65997.1 AA TG AACTTT CAT TTTAG A 162
Full_length GAACACTACAGTATTTGTAACCGATGGAGAAAACG TAGATTTGGTTGTTG 1074
Short_3_end_UTR GAACACTACAGTATTTGTAACCGATGGAGAAAACG TAGATTTGGTTGTTG 1074
SSCs_specific GA TGTAGCTTTTCATTAAAAGCAGCATAACGTTAAGTGTGTTAC TTCAG G 713
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb GA T C CGCAA G 54
Tr-kit_c18-4_4.0kb GA TG GTTTTTA TAG TTTGA G 264
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb GG CGACACCCTCAGCCTGACG TGC ATTGA T 273
NM_001122733.1 AATACGAGGCCTACCCCAAACCCGAGCACCAGCAGTGGATATATATGAACAGGACCTCGG 1140
NM_021099.3 AATACGAGGCCTACCCCAAACCCGAGCACCAGCAGTGGATATATATGAACAGGACCTCGG 1136
X65997.1 AATG 166
Full_length AATACGAGGCCTACCCCAAACCCGAGCACCAGCAGTGGATATATATGAACAGGACCTCGG 1134

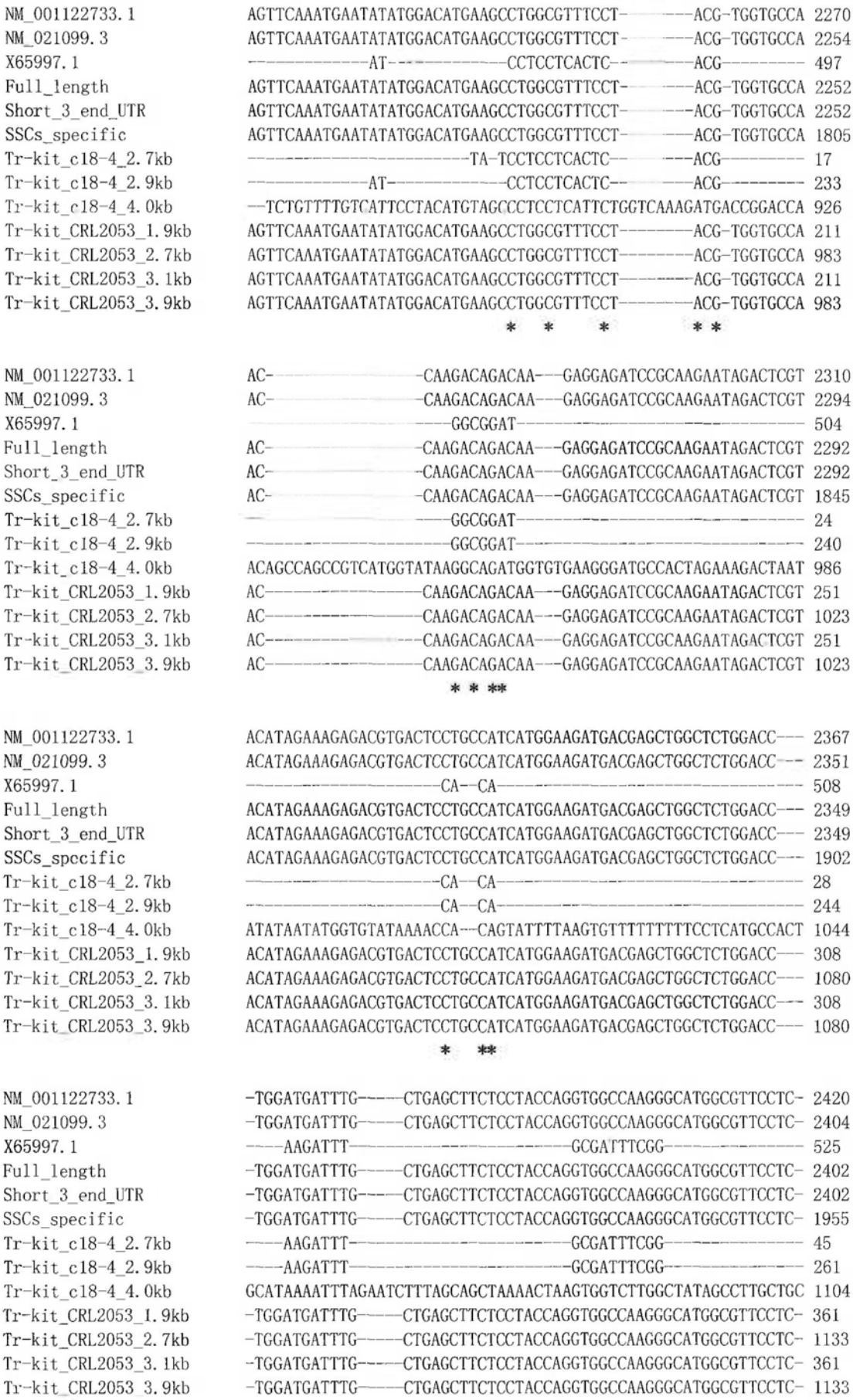

Short_3_end_UTR AATACGAGGCCTACCCCAAACCCGAGCACCAGCAGTGGATATATATGAACAGGACCTCGG 1134
SSCs_specific AATGTCAGTTTCTC-TGTACTCTG-GTTTCA----------ACTTACA 749
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb AAT---------- 57
Tr-kit_c18-4_4.0kb AATCATCTCCTCCC-TCAACCCTCCACCCCC-----ACCCCCA 301
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CCCGACTTTGTCAGATGGACTTTCAAGACCT-----ATTTCAA 311
NM_001122733.1 CTAACAAAGGGAAGGATTATGTCAAATCTGATAACAAAAGCAACATCAGATATGTGAACC 1200
NM_021099.3 CTAACAAAGGGAAGGATTATGTCAAATCTGATAACAAAAGCAACATCAGATATGTGAACC 1196
X65997.1 --------GGATTTGACTAT---------- 178
Full_length CTAACAAAGGGAAGGATTATGTCAAATCTGATAACAAAAGCAACATCAGATATGTGAACC 1194
Short_3_end_UTR CTAACAAAGGGAAGGATTATGTCAAATCTGATAACAAAAGCAACATCAGATATGTGAACC 1194
SSCs_specific TCTTT---AAAGATGCCTGTGCAGCCAGTAAG----------GATGTGTTTGTGGG 792
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb --------AGACTCG--TACA-- 68
Tr-kit_c18-4_4.0kb CCCCC---GGACACGACTGCAC-----TAAA-----CA- 326
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb T-----GAAATGGTTGAGAA-----TAAA-----AAA- 333
NM_001122733.1 AACTTCGCCTGACCAGATTAAAAGGCACAGAAGGAGGCACTTATACCTTTCTGGTGTCCA 1260
NM_021099.3 AACTTCGCCTGACCAGATTAAAAGGCACAGAAGGAGGCACTTATACCTTTCTGGTGTCCA 1256
X65997.1 ----------TTATA--------A---------TG--CAT-----TTTCCT--------- 195
Full_length AACTTCGCCTGACCAGATTAAAAGGCACAGAAGGAGGCACTTATACCTTTCTGGTGTCCA 1254
Short_3_end_UTR AACTTCGCCTGACCAGATTAAAAGGCACAGAAGGAGGCACTTATACCTTTCTGGTGTCCA 1254
SSCs_specific TATTGTGCTCTGCCAGA--------ACACATGACCAGGTTCT-----AGTCCTAGCAC-CA 839
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb ----------TAGA--------A---------- 73
Tr-kit_c18-4_4.0kb -ACTTCAACC--CTAGA--------A----------GGGCTCT-----CTTCCTACTGTTCA 362
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb -AATGAATGGATCCAGG-------AAAAAGCCGAGGCCACT-----CGCACGGGCACATA 380
NM_001122733.1 ACTCTG-----ATGCCAGTGCTTCCGTGACATTCAACGTTTACGTGAACAC---------- 1306
NM_021099.3 ACTCTG-----ATGCCAGTGCTTCCGTGACATTCAACGTTTACGTGAACAC---------- 1302
X65997.1 --GTG--------AATGGAAGGA---------- 208
Full_length ACTCTG-----ATGCCAGTGCTTCCGTGACATTCAACGTTTACGTGAACAC---------- 1300
Short_3_end_UTR ACTCTG-----ATGCCAGTGCTTCCGTGACATTCAACGTTTACGTGAACAC---------- 1300
SSCs_specific AGAATA----------AATGTAGGTAGGTAGGTAGGTAGGTAGGTAGGCAGGTAGC 889
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb ---AGA----------GACGT---------- 81
Tr-kit_c18-4_4.0kb GTCGTA----------AACGGATTTA---------- 378
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CACGTGCAGCAACAGCAATGGCCTCAC----- 407
NM_001122733.1 -AAAACCAGAAATCCTGACG-----TACGACAGGCTCATAAATGGC 1346
NM_021099.3 -AAAACCAGAAATCCTGACG-----TACGACAGGCTCATAAATGGC 1342
X65997.1 ---AGGGA----------GAAAGACGTTTA-------- 225

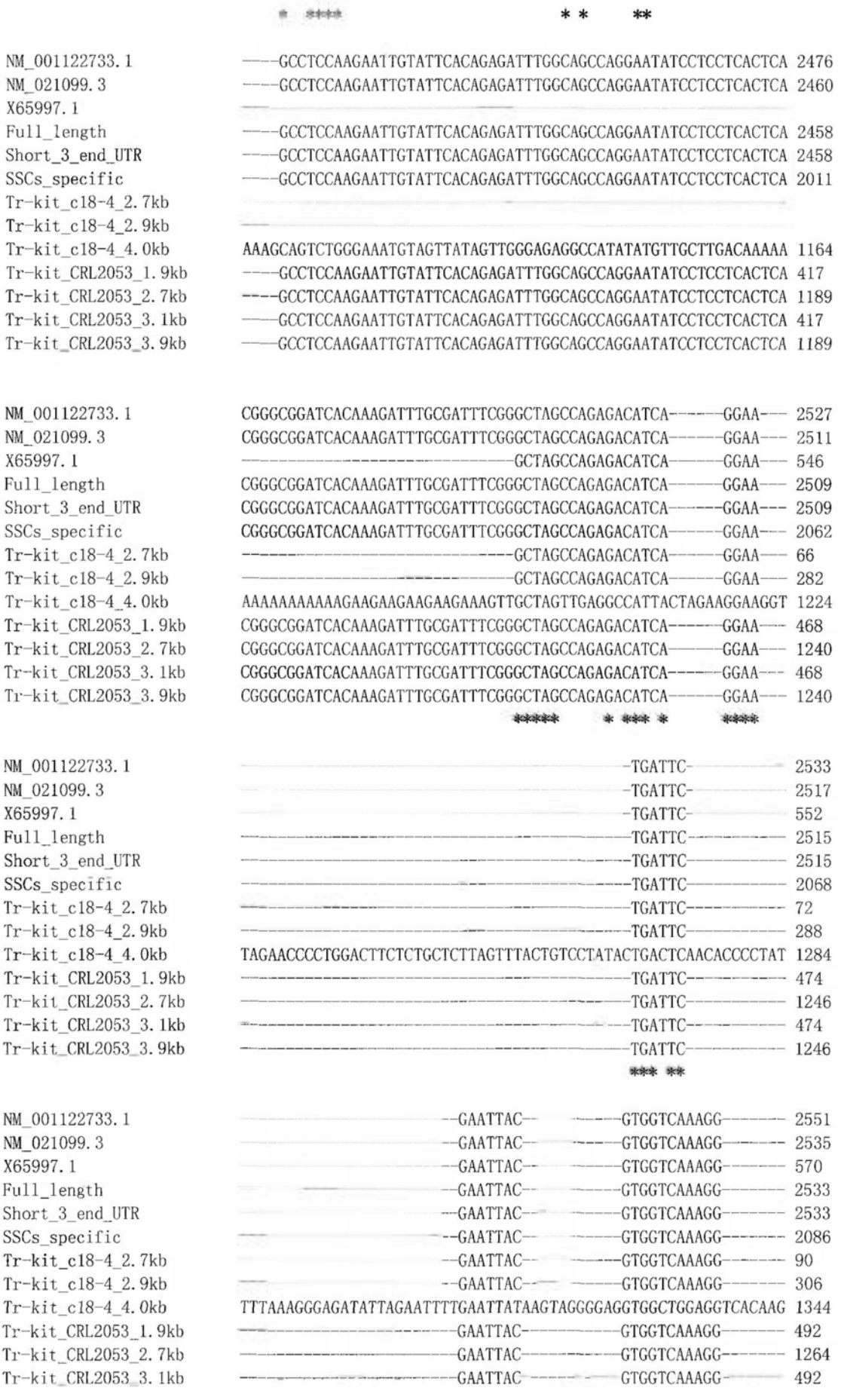

Full_length
Short_3_end_UTR
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NM_001122733.1
NM_021099.3
X65997.1

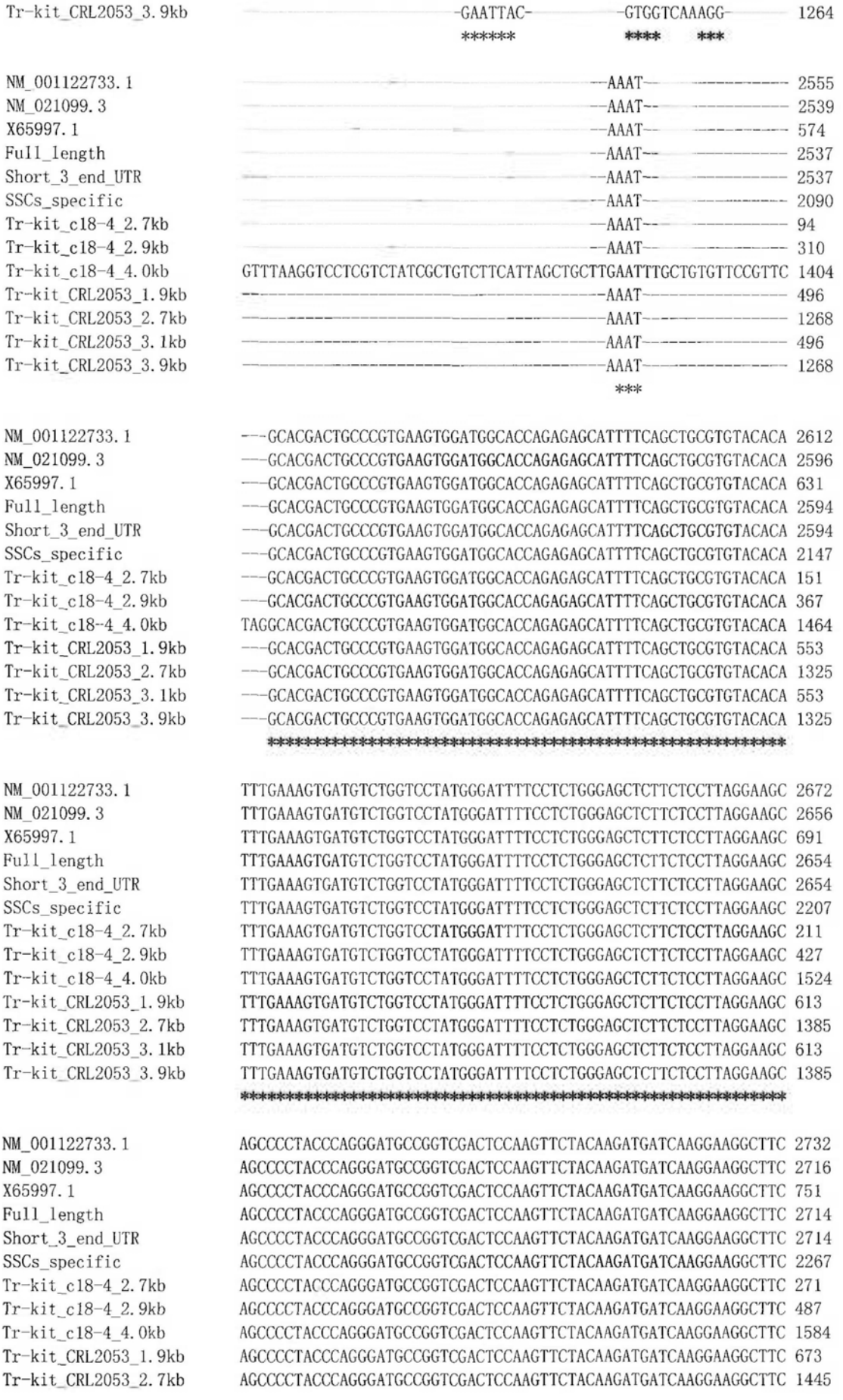
```
X65997.1                        -C---C-C-           -GA-                                    268
Full_length                 TCTT--CCGGCACAACGGCACGGTGGAGTGTAAGGCCTCCAACGATGTGGGCAAGAGTTC 1570
Short_3_end_UTR             TCTT--CCGGCACAACGGCACGGTGGAGTGTAAGGCCTCCAACGATGTGGGCAAGAGTTC 1570
SSCs_specific               TCGCAGCTGGCGC------------GA-TGGGGATC--------------ATTGTGATGG 1134
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb              -CTG-----          -GA--- ------           ------------ 120
Tr-kit_c18-4_4.0kb              -CTGAC-C-          -GAGTGGAAGGA-            -ACTGGAATG- 488
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb            -CTGAC-A-          -GACCCACAGGT-            -GTCCAATTA- 533

NM_001122733.1              CGCCTTCTTTAACTTTGCATTTAAAGGTAACAACAAAGAGCAAATCCAGGCCCACACTCT 1636
NM_021099.3                 CGCCTTCTTTAACTT---TGCATTTAAAG-       ----AGCAAATCCAGGCCCACACTCT 1620
X65997.1                    -----TCCTG--------------TGA---       ----AACA---CAAA----------- 284
Full_length                 CGCCTTCTTTAACTT---TGCATTTAAAG-       ----AGCAAATCCAGGCCCACACTCT 1618
Short_3_end_UTR             CGCCTTCTTTAACTT------TGCATTTA-      -AAGAGCAAATCCAGGCCCACACTCT 1618
SSCs_specific               TGCTCACCTACAAAT---ATTTGCAGGTG-       ----AGCA---TTGAATTGTTCT--- 1176
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb          -----CCTG-                                              124
Tr-kit_c18-4_4.0kb          ---TCTCCTG-           -ACATA----     -AGCA---CTAACTGTGCGC--- 516
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb        ---TTCCCT-                          -CA---TCGAGTGTG-      550

NM_001122733.1              GTTCACGCCGCTGCTCATTGGCTTTGTGGTTGCAGCTGGCGCGATGGGGATCATTGTGAT 1696
NM_021099.3                 GTTCACGCCGCTGCTCATTGGCTTTGTGGTTGCAGCTGGCGCGATGGGGATCATTGTGAT 1680
X65997.1                    ----------------AC-------------GGG---------AATA------------- 293
Full_length                 GTTCACGCCGCTGCTCATTGGCTTTGTGGTCGCAGCTGGCGCGATGGGGATCATTGTGAT 1678
Short_3_end_UTR             GTTCACGCCGCTGCTCATTGGCTTTGTGGTTGCAGCTGGCGCGATGGGGATCATTGTGAT 1678
SSCs_specific               CTTCCTGGGGACGCCAAG--------------GCGGCAGGGCAGGCACTGATTGTTCAGCG 1223
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb          GCTCGCGCGCGTGCACAC-          -GCGGCGCGCACAACA-             549
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb        ---ATGGGAAATCTCTC-           -CCCACGGACCTGACG-             579

NM_001122733.1              GGTGCTCACCTACAAATATT----------TGCAGAAACCCATGTATGAAGTACAATGGAA 1747
NM_021099.3                 GGTGCTCACCTACAAATATT----------TGCAGAAACCCATGTATGAAGTACAATGGAA 1731
X65997.1                    -----TCACTT------------------------GCAC---CATAATTTTTA------ 314
Full_length                 GGTGCTCACCTACAAATATT----------TGCAGAAACCCATGTATGAAGTACAATGGAA 1729
Short_3_end_UTR             GGTGCTCACCTACAAATATT----------TGCAGAAACCCATGTATGAAGTACAATGGAA 1729
SSCs_specific               GGTG-ACACATCTTTCTTTTCCTTTCTCCTCCAGAAACCCATGTATGAAGTACAATGGAA 1282
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb                                        ------------GATGATT------- 131
Tr-kit_c18-4_4.0kb              -ACACAC-                    -ACACACACACAGAGATCAGACACA 579
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb            -TTTGT-                     -CCCAAACCCCAAGGCTGGCATCA 607

NM_001122733.1              GGTTGTCGAGGAGATAAATGGAAACAATTATGTTTACATAGA-CCCGACGCAACTTCCTT 1806
```


NM_021099.3 GGTTGTCGAGGAGATAAATGGAAACAATTATGTTTACATAGA-CCCGACGCAACTTCCTT 1790
X65997.1 --TTTTCGG----------TG--TGC-TAA--------ATACTTT---- 336
Full_length GGTTGTCGAGGAGATAAATGGAAACAATTATGTTTACATAGA-CCCGACGCAACTCCCTT 1788
Short_3_end_UTR GGTTGTCGAGGAGATAAATGGAAACAATTATGTTTACATAGA-CCCGACGCAACTTCCTT 1788
SSCs_specific GGTTGTCGAGGAGATAAATGGAAACAATTATGTTTACATAGA-CCCGACGCAACTTCCTT 1341
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb -----TGC- 134
Tr-kit_c18-4_4.0kb GCTCATCAG- -TGCCTGC-TAGT-CTGGGTAAATCTT---- 612
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb --CCATCAAAAACGTGAA- -GCGCGCCTACCACCGGCTCTGTGTCC--- 649
NM_001122733.1 ATGATCACAAATGGGAGTTTCCCAGAAACAGGCTGAGTTTTGGAAAGACATTGGGAGCTG 1866
NM_021099.3 ATGATCACAAATGGGAGTTTCCCAGAAACAGGCTGAGTTTTGGAAAGACATTGGGAGCTG 1850
X65997.1 ----------AA--AAC-G----------AAAGTT-------- 348
Full_length ATGATCACAAATGGGAGTTTCCCAGAAACAGGCTGAGTTTTGGAAAGACATTGGGAGCTG 1848
Short_3_end_UTR ATGATCACAAATGGGAGTTTCCCAGAAACAGGCTGAGTTTTGGAAAGACATTGGGAGCTG 1848
SSCs_specific ATGATCACAAATGGGAGTTTCCCAGAAACAGGCTGAGTTTTGGAAAGACATTGGGAGCTG 1401
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb ---TGAGCT-- 140
Tr-kit_c18-4_4.0kb -AATGAGCTG- -GACGTTGGTAG--- 632
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb -GCTGTGCTG---CTCAGCGTGACGGTACATGGCT 680
NM_001122733.1 GTGCCTTCGGGAAGGTCGTTGAGGCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCA 1926
NM_021099.3 GTGCCTTCGGGAAGGTCGTTGAGGCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCA 1910
X65997.1 ---TCTTT--------TTTTT--------TTTCAT----------GT------AAA 369
Full_length GTGCCTTCGGGAAGGTCGTTGAGGCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCA 1908
Short_3_end_UTR GTGCCTTCGGGAAGGTCGTTGAGGCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCA 1908
SSCs_specific GTGCCTTCGGGAAGGTCGTTGAGGCCACTGCATATGGCTTGATTAAGTCGGATGCTGCCA 1461
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb ---TCTCC- 145
Tr-kit_c18-4_4.0kb GCATTTTT- -TTATC- --CTGCAT- -TGCCTCAGTTGTCCCATGAAA 672
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb GCATTCTGACAAA--TTCAC--------CCTCAA- -AGTGCGGGCAGCCATCAAGG 724
NM_001122733.1 TGACAGTTGCCGTGAAGATGCTCAAACCAAGTGCCCATTTAACAGAAAGAGAGGCCCTAA 1986
NM_021099.3 TGACAGTTGCCGTGAAGATGCTCAAACCAAGTGCCCATTTAACAGAAAGAGAGGCCCTAA 1970
X65997.1 CACCA--TTGTA------GTATT-AAA-----AT---CATCT----------TC- 396
Full_length TGACAGTTGCCGTGAAGATGCTCAAACCAAGTGCCCATTTAACAGAAAGAGAGGCCCTAA 1968
Short_3_end_UTR TGACAGTTGCCGTGAAGATGCTCAAACCAAGTGCCCATTTAACAGAAAGAGAGGCCCTAA 1968
SSCs_specific TGACAGTTGCCGTGAAGATGCTCAAACCAAGTGCCCATTTAACAGAAAGAGAGGCCCTAA 1521
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb TACCA--------- ----- 150
Tr-kit_c18-4_4.0kb TAACAACTCCTG- -GTATTTGAA----GT---TATTT- --TTG 703
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CTATCCCTGTTGTGTCTGTGCCTGAAACAAGT---CACCT---------- 761


NM_001122733.1 TGTCGGAACTGAAGGT---CCTGAGCTACCTGGGCAATCACATGAATATTGTGAACCTGC 2043
NM_021099.3 TGTCGGAACTGAAGGT---CCTGAGCTACCTGGGCAATCACATGAATATTGTGAACCTGC 2027
X65997.1 TCTCG-GAGAGCTGAA---A-TGAA-----TGGCT----------GTT--GCTGTCTTTC 434
Full_length TGTCGGAACTGAAGGT---CCTGAGCTACCTGGGCAATCACATGAATATTGTGAACCTGC 2025
Short_3_end_UTR TGTCGGAACTGAAGGT---CCTGAGCTACCTGGGCAATCACATGAATATTGTGAACCTGC 2025
SSCs_specific TGTCGGAACTGAAGGT---CCTGAGCTACCTGGGCAATCACATGAATATTGTGAACCTGC 1578
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb ----------GGTGGC---CAAGGGC---ATGGC-------GTTC-----CTCGC 177
Tr-kit_c18-4_4.0kb TTTTGCAAAAGATGAT---TCTGGGC---CTGGCTTATC---CGTGTTTAGGTAACTTCT 754
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb CCTTAAGAAAGGGGACACATTTACGG----TGGTGTGC------ACCATAAAAGATGTGT 811
NM_001122733.1 TTGGCGCATGCACGGTGGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 2103
NM_021099.3 TTGGCGCATGCACGGTGGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 2087
X65997.1 CT----------TTTCTCCC------CCAACAGT---------- 452
Full_length TTGGCGCATGCACGGTGGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 2085
Short_3_end_UTR TTGGCGCATGCACGGTGGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 2085
SSCs_specific TTGGCGCATGCACGGTGGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 1638
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb CT-----------CCAAGAAT-------- 187
Tr-kit_c18-4_4.0kb CTAGTAGGTGCGTGAGGGGTGCACTTGTGTCAGAGGTCCAAGAGT-------- 799
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb ----------GGGAGGGCCCACCCTGGTCATTACAGAATATTGTTGCTATGGTG 44
Tr-kit_CRL2053_3.9kb CTA---CATCCGTGA----------ACTCCATGTGGCTAAAGAT---------- 842
NM_001122733.1 ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 2159
NM_021099.3 ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 2143
X65997.1 -------GTAT-----T----CACAGAGA----------TTTG----------G 470
Full_length ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 2141
Short_3_end_UTR ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 2141
SSCs_specific ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 1694
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb -TGTAT-----T---CACAGAGA-------TTTG---------- 205
Tr-kit_c18-4_4.0kb --GAAT----G---GGAAGGGA-------TTTG----------G 817
Tr-kit_CRL2053_1.9kb ----------AA 100
Tr-kit_CRL2053_2.7kb ----------AA 872
Tr-kit_CRL2053_3.1kb ATCTTTTGAATTTTTTG----AGAAGGAAGCGTGACTCGTTTATTTTCTCAAAGCAAGAA 100
Tr-kit_CRL2053_3.9kb -------GAACCCTCAGCCTCAGCACATAG---------CCCA----------GAA 872
NM_001122733.1 GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 2219
NM_021099.3 GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 2203
X65997.1 ---CAGCCAG-----GAAT---------- 481
Full_length GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 2201
Short_3_end_UTR GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 2201
SSCs_specific GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 1754
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb --GCAGCCAG---GAAT-- 217
Tr-kit_c18-4_4.0kb GGCCAGGCAGA---GAATTCTGAAACAACACTTAGCACC-----GAGCCCTTCTGCCTC 868
Tr-kit_CRL2053_1.9kb GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 160
Tr-kit_CRL2053_2.7kb GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 932
Tr-kit_CRL2053_3.1kb GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 160
Tr-kit_CRL2053_3.9kb GAGCAGGCAGAAGCGGCACTTTATAAGAACCTTCTGCACTCAACGGAGCCTTCCTGTGAC 932


NM_001122733.1 AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 2270
NM_021099.3 AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 2254
X65997.1 AT----CCTCCTCACTC----ACG 497
Full_length AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 2252
Short_3_end_UTR AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 2252
SSCs_specific AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 1805
Tr-kit_c18-4_2.7kb TA-TCCTCCTCACTC----ACG 17
Tr-kit_c18-4_2.9kb AT----CCTCCTCACTC----ACG 233
Tr-kit_c18-4_4.0kb TCTGTTTTGTCATTCCTACATGTAGCCCTCCTCATTCTGGTCAAAGATGACCGGACCA 926
Tr-kit_CRL2053_1.9kb AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 211
Tr-kit_CRL2053_2.7kb AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 983
Tr-kit_CRL2053_3.1kb AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 211
Tr-kit_CRL2053_3.9kb AGTTCAAATGAATATATGGACATGAAGCCTGGCGTTTCCT----ACG-TGGTGCCA 983
NM_001122733.1 AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 2310
NM_021099.3 AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 2294
X65997.1 GGCGGAT 504
Full_length AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 2292
Short_3_end_UTR AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 2292
SSCs_specific AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 1845
Tr-kit_c18-4_2.7kb GGCGGAT 24
Tr-kit_c18-4_2.9kb GGCGGAT 240
Tr-kit_c18-4_4.0kb ACAGCCAGCCGTCATGGTATAAGGCAGATGGTGTGAAGGGATGCCACTAGAAAGACTAAT 986
Tr-kit_CRL2053_1.9kb AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 251
Tr-kit_CRL2053_2.7kb AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 1023
Tr-kit_CRL2053_3.1kb AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 251
Tr-kit_CRL2053_3.9kb AC----CAAGACAGACAA---GAGGAGATCCGCAAGAATAGACTCGT 1023
NM_001122733.1 ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 2367
NM_021099.3 ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 2351
X65997.1 CA--CA 508
Full_length ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 2349
Short_3_end_UTR ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 2349
SSCs_specific ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 1902
Tr-kit_c18-4_2.7kb CA--CA 28
Tr-kit_c18-4_2.9kb CA--CA 244
Tr-kit_c18-4_4.0kb ATATAATATGGTGTATAAAACCA---CAGTATTTTAAGTGTTTTTTTTTCCTCATGCCACT 1044
Tr-kit_CRL2053_1.9kb ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 308
Tr-kit_CRL2053_2.7kb ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 1080
Tr-kit_CRL2053_3.1kb ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 308
Tr-kit_CRL2053_3.9kb ACATAGAAAGAGACGTGACTCCTGCCATCATGGAAGATGACGAGCTGGCTCTGGACC--- 1080
NM_001122733.1 -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 2420
NM_021099.3 -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 2404
X65997.1 AAGATTT GCGATTTCGG 525
Full_length -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 2402
Short_3_end_UTR -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 2402
SSCs_specific -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 1955
Tr-kit_c18-4_2.7kb AAGATTT GCGATTTCGG 45
Tr-kit_c18-4_2.9kb AAGATTT GCGATTTCGG 261
Tr-kit_c18-4_4.0kb GCATAAAATTTAGAATCTTTAGCAGCTAAAACTAAGTGGTCTTGGCTATAGCCTTGCTGC 1104
Tr-kit_CRL2053_1.9kb -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 361
Tr-kit_CRL2053_2.7kb -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 1133
Tr-kit_CRL2053_3.1kb -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 361
Tr-kit_CRL2053_3.9kb -TGGATGATTTG-----CTGAGCTTCTCCTACCAGGTGGCCAAGGGCATGGCGTTCCTC- 1133


NM_001122733.1
NM_021099.3
X65997.1
Full_length
Short_3_end_UTR
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb


Tr-kit_CRL2053_3.9kb -GAATTAC- -GTGGTCAAAGG- 1264
NM_001122733.1 --AAAT-- 2555
NM_021099.3 --AAAT-- 2539
X65997.1 --AAAT-- 574
Full_length --AAAT-- 2537
Short_3_end_UTR --AAAT-- 2537
SSCs_specific --AAAT-- 2090
Tr-kit_c18-4_2.7kb --AAAT-- 94
Tr-kit_c18-4_2.9kb --AAAT-- 310
Tr-kit_c18-4_4.0kb GTTTAAGGTCCTCGTCTATCGCTGTCTTCATTAGCTGCTTGAATTTGCTGTGTTCCGTTC 1404
Tr-kit_CRL2053_1.9kb --AAAT-- 496
Tr-kit_CRL2053_2.7kb --AAAT-- 1268
Tr-kit_CRL2053_3.1kb --AAAT-- 496
Tr-kit_CRL2053_3.9kb --AAAT-- 1268
NM_001122733.1 ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 2612
NM_021099.3 ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 2596
X65997.1 ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 631
Full_length ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 2594
Short_3_end_UTR ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 2594
SSCs_specific ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 2147
Tr-kit_c18-4_2.7kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 151
Tr-kit_c18-4_2.9kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 367
Tr-kit_c18-4_4.0kb TAGGCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 1464
Tr-kit_CRL2053_1.9kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 553
Tr-kit_CRL2053_2.7kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 1325
Tr-kit_CRL2053_3.1kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 553
Tr-kit_CRL2053_3.9kb ---GCACGACTGCCCGTGAAGTGGATGGCACCAGAGAGCATTTTCAGCTGCGTGTACACA 1325
NM_001122733.1 TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 2672
NM_021099.3 TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 2656
X65997.1 TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 691
Full_length TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 2654
Short_3_end_UTR TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 2654
SSCs_specific TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 2207
Tr-kit_c18-4_2.7kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 211
Tr-kit_c18-4_2.9kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 427
Tr-kit_c18-4_4.0kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 1524
Tr-kit_CRL2053_1.9kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 613
Tr-kit_CRL2053_2.7kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 1385
Tr-kit_CRL2053_3.1kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 613
Tr-kit_CRL2053_3.9kb TTTGAAAGTGATGTCTGGTCCTATGGGATTTTCCTCTGGGAGCTCTTCTCCTTAGGAAGC 1385
NM_001122733.1 AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 2732
NM_021099.3 AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 2716
X65997.1 AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 751
Full_length AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 2714
Short_3_end_UTR AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 2714
SSCs_specific AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 2267
Tr-kit_c18-4_2.7kb AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 271
Tr-kit_c18-4_2.9kb AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 487
Tr-kit_c18-4_4.0kb AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 1584
Tr-kit_CRL2053_1.9kb AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 673
Tr-kit_CRL2053_2.7kb AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 1445

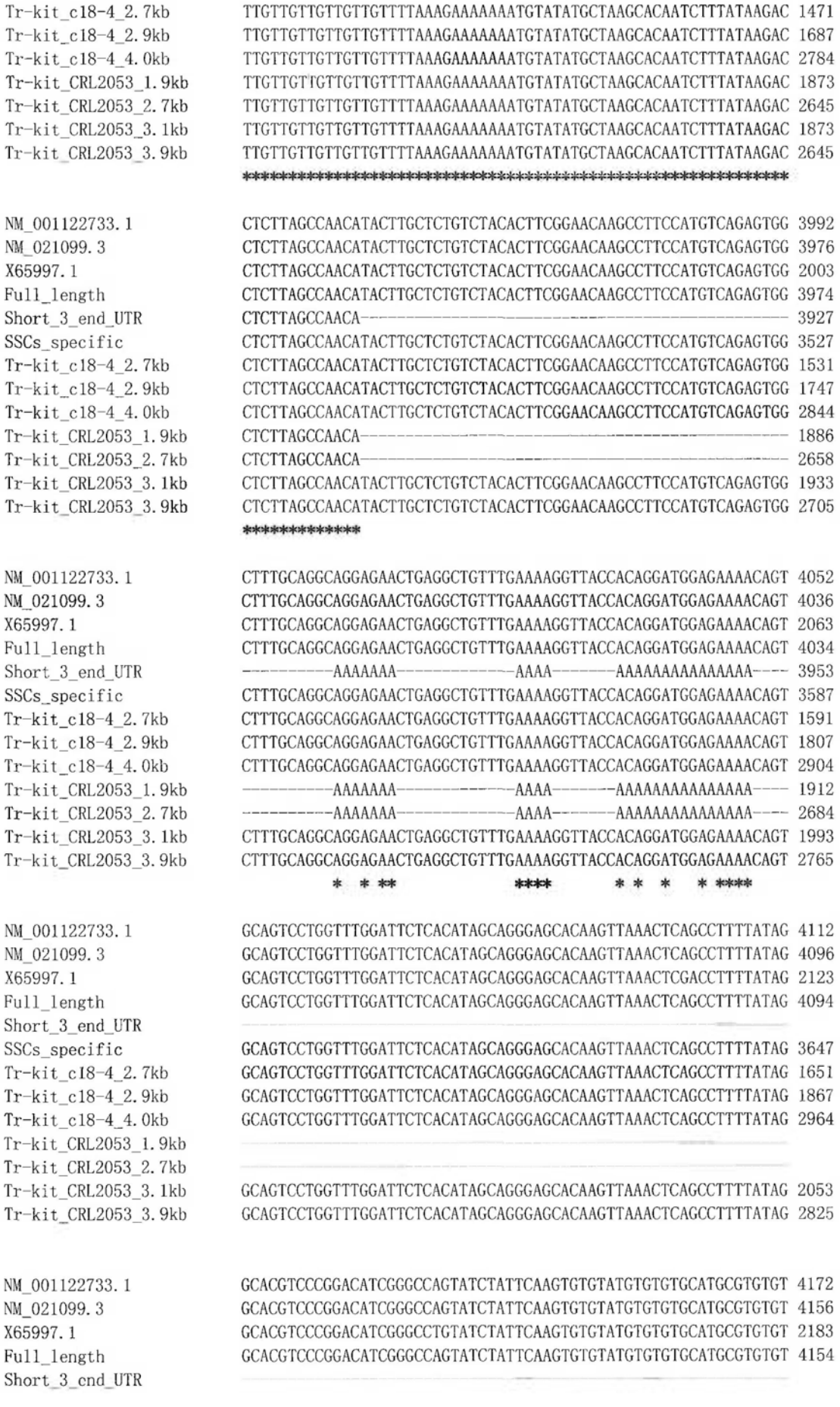
```
Tr-kit_CRL2053_3.1kb      AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 673
Tr-kit_CRL2053_3.9kb      AGCCCCTACCCAGGGATGCCGGTCGACTCCAAGTTCTACAAGATGATCAAGGAAGGCTTC 1445
                          ************************************************************

NM_001122733.1            CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 2792
NM_021099.3               CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 2776
X65997.1                  CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 811
Full_length               CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 2774
Short_3_end_UTR           CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 2774
SSCs_specific             CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 2327
Tr-kit_c18-4_2.7kb        CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 331
Tr-kit_c18-4_2.9kb        CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 547
Tr-kit_c18-4_4.0kb        CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 1644
Tr-kit_CRL2053_1.9kb      CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 733
Tr-kit_CRL2053_2.7kb      CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 1505
Tr-kit_CRL2053_3.1kb      CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 733
Tr-kit_CRL2053_3.9kb      CGGATGGTCAGCCCGGAGCACGCGCCTGCCGAAATGTATGACGTCATGAAGACTTGCTGG 1505
                          ************************************************************

NM_001122733.1            GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 2852
NM_021099.3               GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 2836
X65997.1                  GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 871
Full_length               GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 2834
Short_3_end_UTR           GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 2834
SSCs_specific             GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 2387
Tr-kit_c18-4_2.7kb        GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 391
Tr-kit_c18-4_2.9kb        GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 607
Tr-kit_c18-4_4.0kb        GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 1704
Tr-kit_CRL2053_1.9kb      GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 793
Tr-kit_CRL2053_2.7kb      GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 1565
Tr-kit_CRL2053_3.1kb      GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 793
Tr-kit_CRL2053_3.9kb      GACGCTGACCCCTTGAAAAGGCCAACATTCAAGCAGGTTGTCCAACTTATTGAGAAGCAG 1565
                          ************************************************************

NM_001122733.1            ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 2912
NM_021099.3               ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 2896
X65997.1                  ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 931
Full_length               ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 2894
Short_3_end_UTR           ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 2894
SSCs_specific             ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 2447
Tr-kit_c18-4_2.7kb        ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 451
Tr-kit_c18-4_2.9kb        ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 667
Tr-kit_c18-4_4.0kb        ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 1764
Tr-kit_CRL2053_1.9kb      ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 853
Tr-kit_CRL2053_2.7kb      ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 1625
Tr-kit_CRL2053_3.1kb      ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 853
Tr-kit_CRL2053_3.9kb      ATCTCGGACAGCACCAAGCACATTTACTCCAACTTGGCAAACTGCAACCCCAACCCAGAG 1625
                          ************************************************************

NM_001122733.1            AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 2972
NM_021099.3               AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 2956
X65997.1                  AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 991
Full_length               AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 2954
Short_3_end_UTR           AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 2954
SSCs_specific             AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 2507
Tr-kit_c18-4_2.7kb        AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 511
Tr-kit_c18-4_2.9kb        AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 727
Tr-kit_c18-4_4.0kb        AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 1824
Tr-kit_CRL2053_1.9kb      AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 913
```

```
Tr-kit_CRL2053_2.7kb      AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 1685
Tr-kit_CRL2053_3.1kb      AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 913
Tr-kit_CRL2053_3.9kb      AACCCCGTGGTGGTGGACCATTCCGTGAGGGTCAACTCGGTGGGCAGCAGCGCCTCTTCT 1685
                          ************************************************************

NM_001122733.1            ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 3032
NM_021099.3               ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 3016
X65997.1                  ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 1051
Full_length               ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 3014
Short_3_end_UTR           ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 3014
SSCs_specific             ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 2567
Tr-kit_c18-4_2.7kb        ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 571
Tr-kit_c18-4_2.9kb        ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 787
Tr-kit_c18-4_4.0kb        ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 1884
Tr-kit_CRL2053_1.9kb      ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 973
Tr-kit_CRL2053_2.7kb      ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 1745
Tr-kit_CRL2053_3.1kb      ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 973
Tr-kit_CRL2053_3.9kb      ACGCAGCCCCTGCTCGTGCACGAAGATGCCTGAGCAGAAACCCAAGTCCAACAGGCTTTG 1745
                          ************************************************************

NM_001122733.1            CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 3092
NM_021099.3               CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 3076
X65997.1                  CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1111
Full_length               CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 3074
Short_3_end_UTR           CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 3074
SSCs_specific             CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 2627
Tr-kit_c18-4_2.7kb        CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 631
Tr-kit_c18-4_2.9kb        CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 847
Tr-kit_c18-4_4.0kb        CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1944
Tr-kit_CRL2053_1.9kb      CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1033
Tr-kit_CRL2053_2.7kb      CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1805
Tr-kit_CRL2053_3.1kb      CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1033
Tr-kit_CRL2053_3.9kb      CTGCTGTCTCCGACCCCGTCCTTCTGGCTTCTGTGATGGTTACTTGGTTTCCCTTTGACT 1805
                          ************************************************************

NM_001122733.1            TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 3152
NM_021099.3               TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 3136
X65997.1                  TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 1171
Full_length               TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 3134
Short_3_end_UTR           TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 3134
SSCs_specific             TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 2687
Tr-kit_c18-4_2.7kb        TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 691
Tr-kit_c18-4_2.9kb        TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 907
Tr-kit_c18-4_4.0kb        TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 2004
Tr-kit_CRL2053_1.9kb      TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 1093
Tr-kit_CRL2053_2.7kb      TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 1865
Tr-kit_CRL2053_3.1kb      TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 1093
Tr-kit_CRL2053_3.9kb      TGCATCCTATTCCAGGGTAGCGAGTTCCCCACCCCACCTCCAACCCCACTGTGATTCCGC 1865
                          ************************************************************

NM_001122733.1            CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 3212
NM_021099.3               CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 3196
X65997.1                  CTTTACGAGCACACACTTTAGTGGCCGATGGC-TTTTCTTTTCTGCCATCAGCCACCGTC 1230
Full_length               CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 3194
Short_3_end_UTR           CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 3194
SSCs_specific             CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 2747
Tr-kit_c18-4_2.7kb        CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 751
Tr-kit_c18-4_2.9kb        CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 967
Tr-kit_c18-4_4.0kb        CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 2064
```

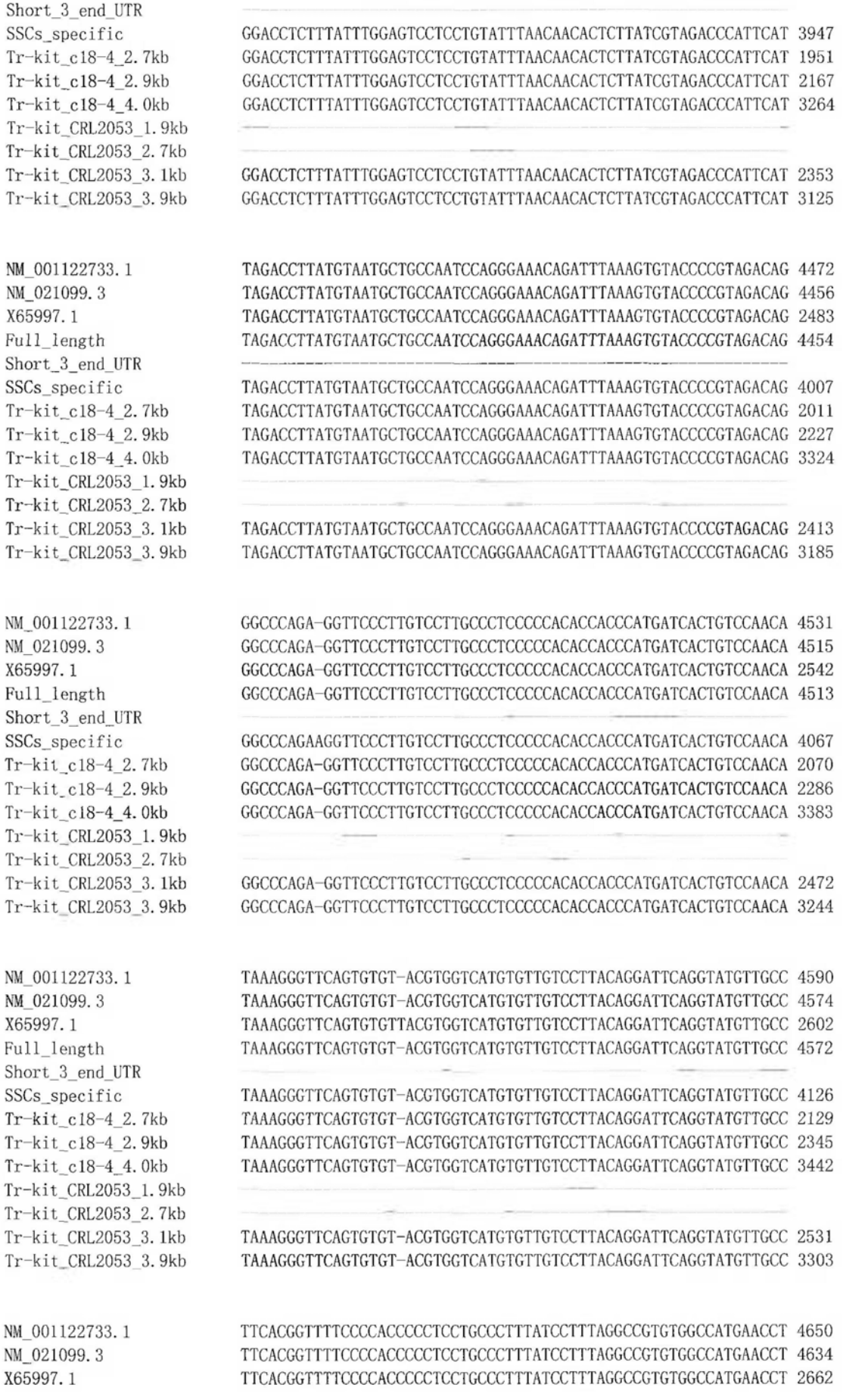
```
Tr-kit_CRL2053_1.9kb      CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 1153
Tr-kit_CRL2053_2.7kb      CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 1925
Tr-kit_CRL2053_3.1kb      CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 1153
Tr-kit_CRL2053_3.9kb      CTTTACGAGCACACACTTTAGTGGCCGATGGCCTTTTCTTTTCTGCCATCAGCCACCGTC 1925
                          ********************************* ***************************

NM_001122733.1            CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 3272
NM_021099.3               CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 3256
X65997.1                  CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1290
Full_length               CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 3254
Short_3_end_UTR           CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 3254
SSCs_specific             CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 2807
Tr-kit_c18-4_2.7kb        CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 811
Tr-kit_c18-4_2.9kb        CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1027
Tr-kit_c18-4_4.0kb        CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 2124
Tr-kit_CRL2053_1.9kb      CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1213
Tr-kit_CRL2053_2.7kb      CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1985
Tr-kit_CRL2053_3.1kb      CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1213
Tr-kit_CRL2053_3.9kb      CCGCTGCGAAGGTCCGAACTGTATGTATATATTTTCCCAATAGCAAAGTAGCTCCTACTG 1985
                          ************************************************************

NM_001122733.1            TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 3332
NM_021099.3               TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 3316
X65997.1                  TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 1350
Full_length               TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 3314
Short_3_end_UTR           TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 3314
SSCs_specific             TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 2867
Tr-kit_c18-4_2.7kb        TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 871
Tr-kit_c18-4_2.9kb        TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 1087
Tr-kit_c18-4_4.0kb        TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 2184
Tr-kit_CRL2053_1.9kb      TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 1273
Tr-kit_CRL2053_2.7kb      TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 2045
Tr-kit_CRL2053_3.1kb      TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 1273
Tr-kit_CRL2053_3.9kb      TAAACAGAAGGACTCCTCCTGCTTTAGAGGAGAAGGGAAGGGCGGGGTGAAACTGGATGC 2045
                          ************************************************************

NM_001122733.1            CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 3392
NM_021099.3               CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 3376
X65997.1                  CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 1410
Full_length               CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 3374
Short_3_end_UTR           CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 3374
SSCs_specific             CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 2927
Tr-kit_c18-4_2.7kb        CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 931
Tr-kit_c18-4_2.9kb        CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 1147
Tr-kit_c18-4_4.0kb        CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 2244
Tr-kit_CRL2053_1.9kb      CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 1333
Tr-kit_CRL2053_2.7kb      CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 2105
Tr-kit_CRL2053_3.1kb      CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 1333
Tr-kit_CRL2053_3.9kb      CCAGAGTTCTTCCCCCAGTGCTCCCCTGAGTGTATTTGAAAAGTATGGCCAGTAGTTCAC 2105
                          ************************************************************

NM_001122733.1            TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 3452
NM_021099.3               TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 3436
X65997.1                  TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 1470
Full_length               TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 3434
Short_3_end_UTR           TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 3434
SSCs_specific             TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 2987
Tr-kit_c18-4_2.7kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 991
Tr-kit_c18-4_2.9kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 1207
```

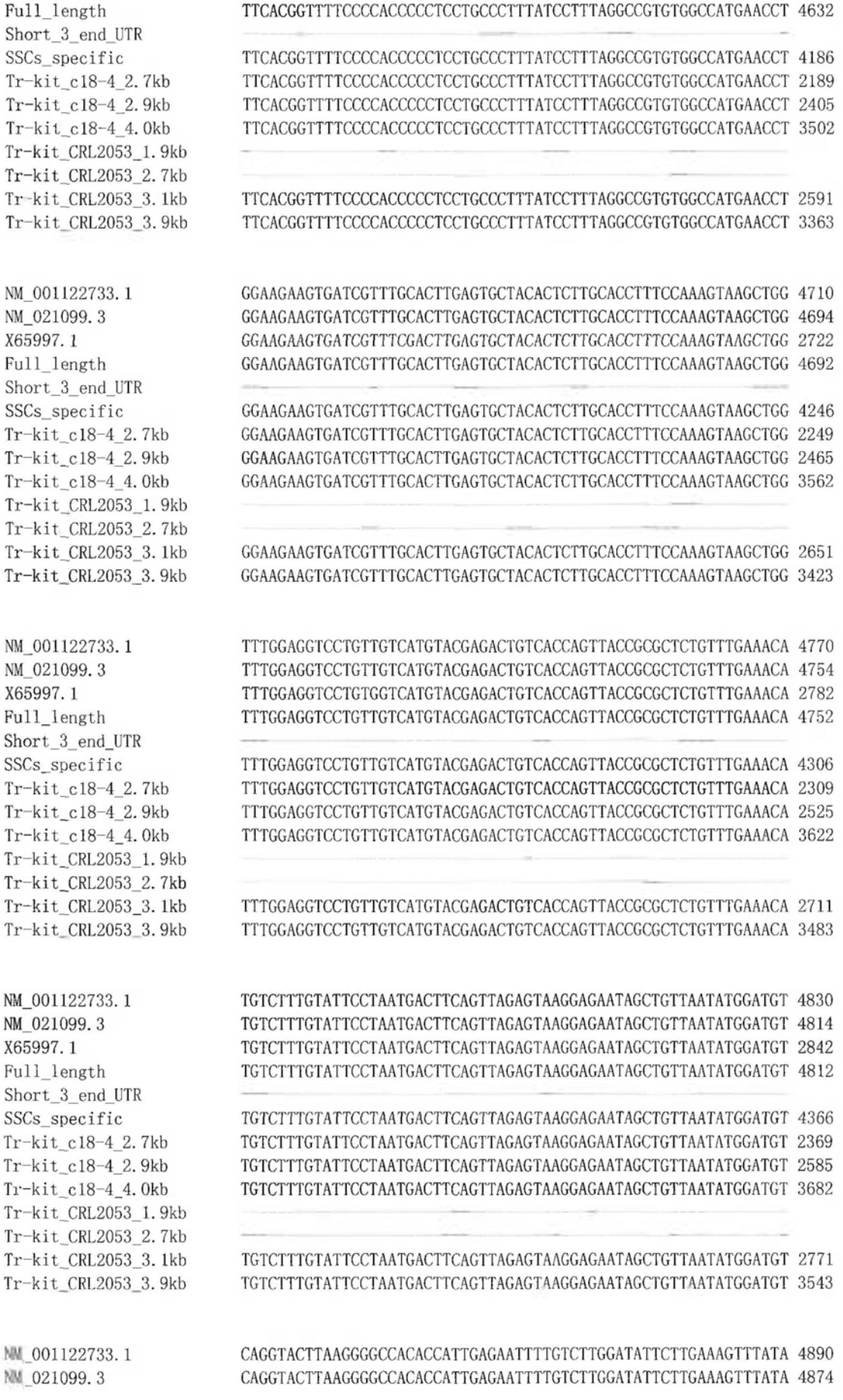
```
Tr-kit_c18-4_4.0kb          TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 2304
Tr-kit_CRL2053_1.9kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 1393
Tr-kit_CRL2053_2.7kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 2165
Tr-kit_CRL2053_3.1kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 1393
Tr-kit_CRL2053_3.9kb        TTGAAGAATAGATGTAGTCCCATTTGGCCCTGAGAGCCATCCTTAATGATGGGAGATATA 2165
                            ************************************************************

NM_001122733.1              TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 3512
NM_021099.3                 TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 3496
X65997.1                    TGTAGCAAGACTA------GAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 1524
Full_length                 TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 3494
Short_3_end_UTR             TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 3494
SSCs_specific               TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 3047
Tr-kit_c18-4_2.7kb          TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 1051
Tr-kit_c18-4_2.9kb          TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 1267
Tr-kit_c18-4_4.0kb          TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 2364
Tr-kit_CRL2053_1.9kb        TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 1453
Tr-kit_CRL2053_2.7kb        TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 2225
Tr-kit_CRL2053_3.1kb        TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 1453
Tr-kit_CRL2053_3.9kb        TGTAGCAAGACTAGAAAAGGAAAGCCAAGCCCTTTGTGTAGAAAGCAGACCATTCTTAGA 2225
                            ************* ****************************************

NM_001122733.1              ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 3572
NM_021099.3                 ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 3556
X65997.1                    ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 1584
Full_length                 ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 3554
Short_3_end_UTR             ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 3554
SSCs_specific               ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 3107
Tr-kit_c18-4_2.7kb          ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 1111
Tr-kit_c18-4_2.9kb          ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 1327
Tr-kit_c18-4_4.0kb          ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 2424
Tr-kit_CRL2053_1.9kb        ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 1513
Tr-kit_CRL2053_2.7kb        ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 2285
Tr-kit_CRL2053_3.1kb        ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 1513
Tr-kit_CRL2053_3.9kb        ACAGAGGGCAACGGGGCATCGGAAGTCTGGTCACGCTAAGAAGACCGAGGCTGAGAAGGA 2285
                            ************************************************************

NM_001122733.1              ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 3632
NM_021099.3                 ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 3616
X65997.1                    ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 1644
Full_length                 ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 3614
Short_3_end_UTR             ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 3614
SSCs_specific               ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 3167
Tr-kit_c18-4_2.7kb          ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 1171
Tr-kit_c18-4_2.9kb          ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 1387
Tr-kit_c18-4_4.0kb          ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 2484
Tr-kit_CRL2053_1.9kb        ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 1573
Tr-kit_CRL2053_2.7kb        ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 2345
Tr-kit_CRL2053_3.1kb        ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 1573
Tr-kit_CRL2053_3.9kb        ACAAGCCAGGGGAAGCGTGAACAATGATGCTCTGCTCTGGGCTGCCGCTCGGGCTTCTGT 2345
                            ************************************************************

NM_001122733.1              ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 3692
NM_021099.3                 ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 3676
X65997.1                    ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 1704
Full_length                 ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 3674
Short_3_end_UTR             ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 3674
SSCs_specific               ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 3227
Tr-kit_c18-4_2.7kb          ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 1231
```

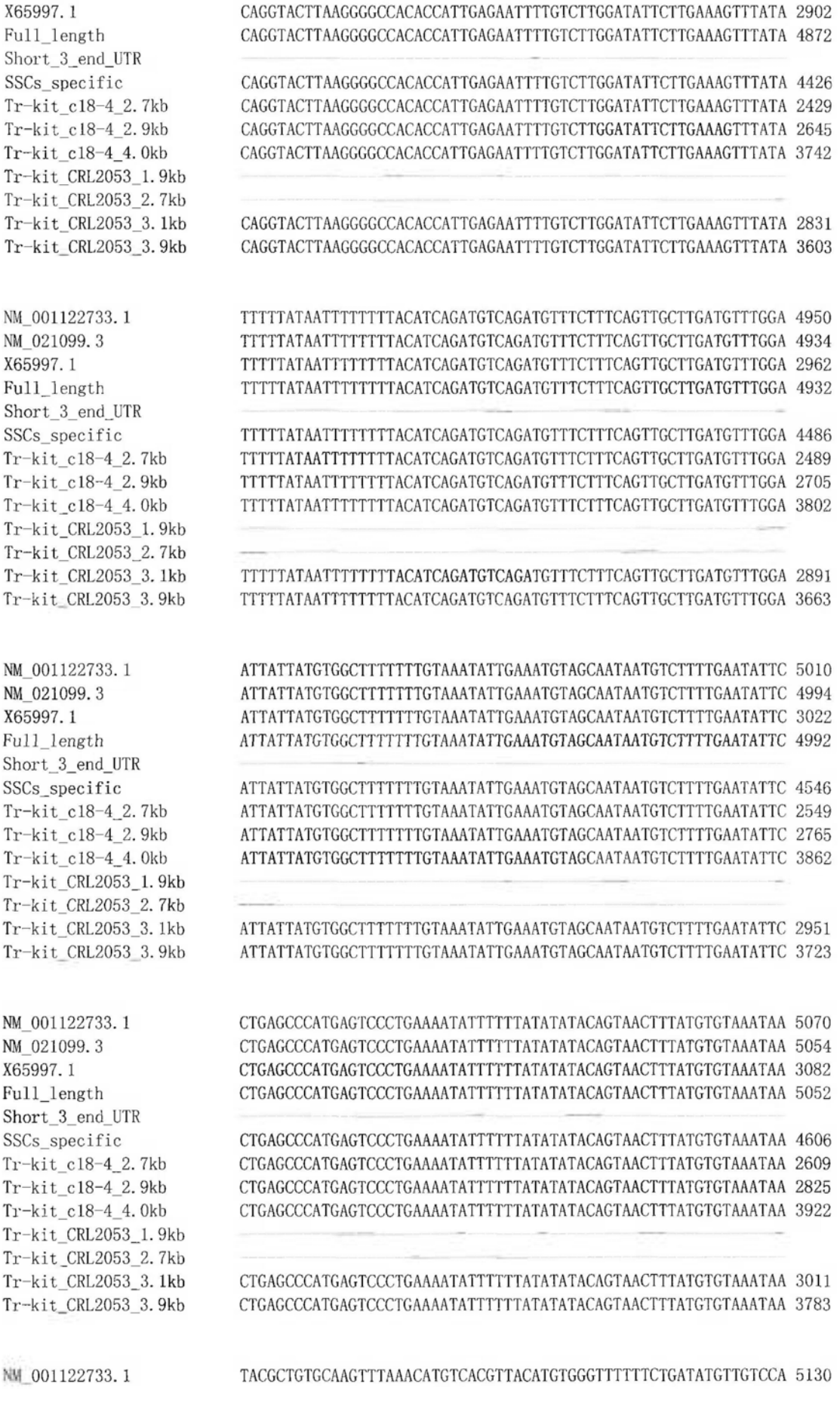
```
Tr-kit_c18-4_2.9kb      ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 1447
Tr-kit_c18-4_4.0kb      ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 2544
Tr-kit_CRL2053_1.9kb    ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 1633
Tr-kit_CRL2053_2.7kb    ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 2405
Tr-kit_CRL2053_3.1kb    ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 1633
Tr-kit_CRL2053_3.9kb    ACAACTGACCTGGTTTCTCAGTACTTTGCTGTCTGGGAGTAGCATTGGAATCAAGGCCTC 2405
                        ************************************************************

NM_001122733.1          CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 3752
NM_021099.3             CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 3736
X65997.1                CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 1764
Full_length             CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 3734
Short_3_end_UTR         CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 3734
SSCs_specific           CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 3287
Tr-kit_c18-4_2.7kb      CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 1291
Tr-kit_c18-4_2.9kb      CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 1507
Tr-kit_c18-4_4.0kb      CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 2604
Tr-kit_CRL2053_1.9kb    CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 1693
Tr-kit_CRL2053_2.7kb    CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 2465
Tr-kit_CRL2053_3.1kb    CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 1693
Tr-kit_CRL2053_3.9kb    CTCCCTAGTCAGCCTTTGTATATACTCATCTATACGTTGTATGCGTTCATACTTTGGAGG 2465
                        ************************************************************

NM_001122733.1          AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 3812
NM_021099.3             AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 3796
X65997.1                AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGG-ATTAGACCTACTGTGTGT 1823
Full_length             AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 3794
Short_3_end_UTR         AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 3794
SSCs_specific           AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 3347
Tr-kit_c18-4_2.7kb      AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 1351
Tr-kit_c18-4_2.9kb      AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 1567
Tr-kit_c18-4_4.0kb      AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 2664
Tr-kit_CRL2053_1.9kb    AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 1753
Tr-kit_CRL2053_2.7kb    AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 2525
Tr-kit_CRL2053_3.1kb    AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 1753
Tr-kit_CRL2053_3.9kb    AGGGATTTCCCACAAGCTTTCGTTTCTGTGTACAGCCCTGGCATTAGACCTACTGTGTGT 2525
                        ************************************** ********************

NM_001122733.1          AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 3872
NM_021099.3             AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 3856
X65997.1                AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 1883
Full_length             AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 3854
Short_3_end_UTR         AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 3854
SSCs_specific           AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 3407
Tr-kit_c18-4_2.7kb      AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 1411
Tr-kit_c18-4_2.9kb      AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 1627
Tr-kit_c18-4_4.0kb      AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 2724
Tr-kit_CRL2053_1.9kb    AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 1813
Tr-kit_CRL2053_2.7kb    AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 2585
Tr-kit_CRL2053_3.1kb    AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 1813
Tr-kit_CRL2053_3.9kb    AAGAATAGATTAAGAGCCATACATATTTGAAGGAAACAGTTAAATGTTTTTTGGTTGTGG 2585
                        ************************************************************

NM_001122733.1          TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 3932
NM_021099.3             TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 3916
X65997.1                TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 1943
Full_length             TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 3914
Short_3_end_UTR         TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 3914
SSCs_specific           TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 3467
```

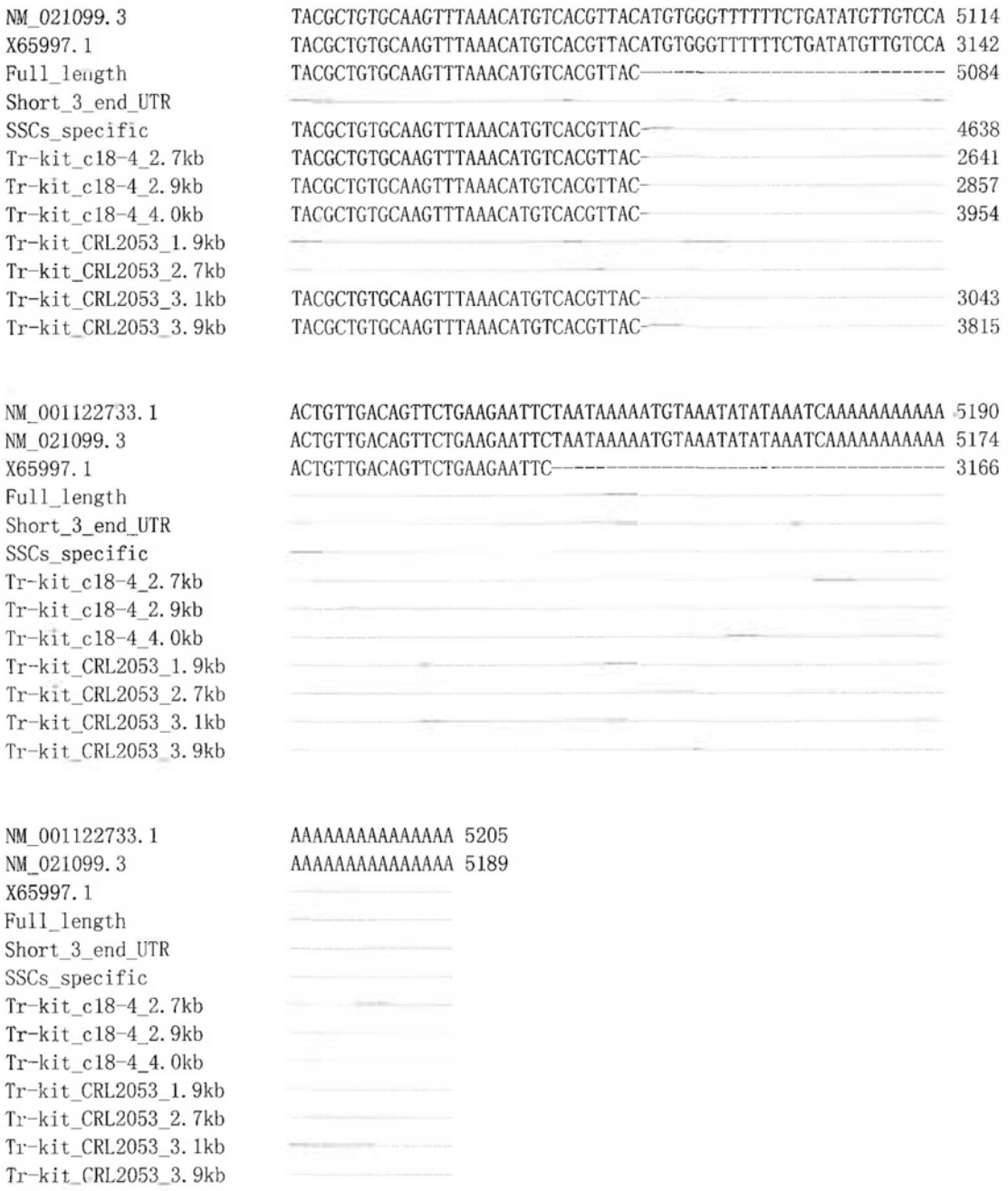

Tr-kit_c18-4_2.7kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 1471
Tr-kit_c18-4_2.9kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 1687
Tr-kit_c18-4_4.0kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 2784
Tr-kit_CRL2053_1.9kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 1873
Tr-kit_CRL2053_2.7kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 2645
Tr-kit_CRL2053_3.1kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 1873
Tr-kit_CRL2053_3.9kb TTGTTGTTGTTGTTGTTTTAAAGAAAAAAATGTATATGCTAAGCACAATCTTTATAAGAC 2645
NM_001122733.1 CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 3992
NM_021099.3 CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 3976
X65997.1 CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 2003
Full_length CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 3974
Short_3_end_UTR CTCTTAGCCAACA 3927
SSCs_specific CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 3527
Tr-kit_c18-4_2.7kb CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 1531
Tr-kit_c18-4_2.9kb CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 1747
Tr-kit_c18-4_4.0kb CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 2844
Tr-kit_CRL2053_1.9kb CTCTTAGCCAACA 1886
Tr-kit_CRL2053_2.7kb CTCTTAGCCAACA 2658
Tr-kit_CRL2053_3.1kb CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 1933
Tr-kit_CRL2053_3.9kb CTCTTAGCCAACATACTTGCTCTGTCTACACTTCGGAACAAGCCTTCCATGTCAGAGTGG 2705
NM_001122733.1 CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 4052
NM_021099.3 CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 4036
X65997.1 CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 2063
Full_length CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 4034
Short_3_end_UTR AAAAAAA AAAA AAAAAAAAAAAAAAA 3953
SSCs_specific CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 3587
Tr-kit_c18-4_2.7kb CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 1591
Tr-kit_c18-4_2.9kb CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 1807
Tr-kit_c18-4_4.0kb CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 2904
Tr-kit_CRL2053_1.9kb AAAAAAA AAAA AAAAAAAAAAAAAAA 1912
Tr-kit_CRL2053_2.7kb AAAAAAA AAAA AAAAAAAAAAAAAAA 2684
Tr-kit_CRL2053_3.1kb CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 1993
Tr-kit_CRL2053_3.9kb CTTTGCAGGCAGGAGAACTGAGGCTGTTTGAAAAGGTTACCACAGGATGGAGAAAACAGT 2765
NM_001122733.1 GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 4112
NM_021099.3 GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 4096
X65997.1 GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCGACCTTTTATAG 2123
Full_length GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 4094
Short_3_end_UTR
SSCs_specific GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 3647
Tr-kit_c18-4_2.7kb GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 1651
Tr-kit_c18-4_2.9kb GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 1867
Tr-kit_c18-4_4.0kb GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 2964
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 2053
Tr-kit_CRL2053_3.9kb GCAGTCCTGGTTTGGATTCTCACATAGCAGGGAGCACAAGTTAAACTCAGCCTTTTATAG 2825
NM_001122733.1 GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 4172
NM_021099.3 GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 4156
X65997.1 GCACGTCCCGGACATCGGGCCTGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 2183
Full_length GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 4154
Short_3_end_UTR

```
SSCs_specific                GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 3707
Tr-kit_c18-4_2.7kb           GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 1711
Tr-kit_c18-4_2.9kb           GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 1927
Tr-kit_c18-4_4.0kb           GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 3024
Tr-kit_CRL2053_1.9kb         ----------------------------------------------------------
Tr-kit_CRL2053_2.7kb         ----------------------------------------------------------
Tr-kit_CRL2053_3.1kb         GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 2113
Tr-kit_CRL2053_3.9kb         GCACGTCCCGGACATCGGGCCAGTATCTATTCAAGTGTGTATGTGTGTGCATGCGTGTGT 2885

NM_001122733.1               CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 4232
NM_021099.3                  CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 4216
X65997.1                     CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 2243
Full_length                  CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 4214
Short_3_end_UTR              ----------------------------------------------------------
SSCs_specific                CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 3767
Tr-kit_c18-4_2.7kb           CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 1771
Tr-kit_c18-4_2.9kb           CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 1987
Tr-kit_c18-4_4.0kb           CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 3084
Tr-kit_CRL2053_1.9kb         ----------------------------------------------------------
Tr-kit_CRL2053_2.7kb         ----------------------------------------------------------
Tr-kit_CRL2053_3.1kb         CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 2173
Tr-kit_CRL2053_3.9kb         CTATGCGTGTGGGTGAGTTGTGTTGGGAAACTTGCCCTGCATCCCTGAGGGTCCTCCTTC 2945

NM_001122733.1               AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 4292
NM_021099.3                  AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 4276
X65997.1                     AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 2303
Full_length                  AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 4274
Short_3_end_UTR              ----------------------------------------------------------
SSCs_specific                AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 3827
Tr-kit_c18-4_2.7kb           AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 1831
Tr-kit_c18-4_2.9kb           AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 2047
Tr-kit_c18-4_4.0kb           AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 3144
Tr-kit_CRL2053_1.9kb         ----------------------------------------------------------
Tr-kit_CRL2053_2.7kb         ----------------------------------------------------------
Tr-kit_CRL2053_3.1kb         AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 2233
Tr-kit_CRL2053_3.9kb         AGGACCCAAGACGTAACAGCTTCTGTCACCGCTCCTGTCTCTCCAGTTTCCCTGCATGTC 3005

NM_001122733.1               GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 4352
NM_021099.3                  GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 4336
X65997.1                     GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 2363
Full_length                  GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 4334
Short_3_end_UTR              ----------------------------------------------------------
SSCs_specific                GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 3887
Tr-kit_c18-4_2.7kb           GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 1891
Tr-kit_c18-4_2.9kb           GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 2107
Tr-kit_c18-4_4.0kb           GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 3204
Tr-kit_CRL2053_1.9kb         ----------------------------------------------------------
Tr-kit_CRL2053_2.7kb         ----------------------------------------------------------
Tr-kit_CRL2053_3.1kb         GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 2293
Tr-kit_CRL2053_3.9kb         GCTCACTGTCTAGAATTTACTCAAAGCCGCCACAGAGGCTTAGCGGAGTGAAGTGCCGAA 3065

NM_001122733.1               GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 4412
NM_021099.3                  GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 4396
X65997.1                     GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 2423
Full_length                  GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 4394
```


Short_3_end_UTR
SSCs_specific GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 3947
Tr-kit_c18-4_2.7kb GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 1951
Tr-kit_c18-4_2.9kb GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 2167
Tr-kit_c18-4_4.0kb GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 3264
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 2353
Tr-kit_CRL2053_3.9kb GGACCTCTTTATTTGGAGTCCTCCTGTATTTAACAACACTCTTATCGTAGACCCATTCAT 3125
NM_001122733.1 TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 4472
NM_021099.3 TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 4456
X65997.1 TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 2483
Full_length TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 4454
Short_3_end_UTR
SSCs_specific TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 4007
Tr-kit_c18-4_2.7kb TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 2011
Tr-kit_c18-4_2.9kb TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 2227
Tr-kit_c18-4_4.0kb TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 3324
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 2413
Tr-kit_CRL2053_3.9kb TAGACCTTATGTAATGCTGCCAATCCAGGGAAACAGATTTAAAGTGTACCCCGTAGACAG 3185
NM_001122733.1 GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 4531
NM_021099.3 GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 4515
X65997.1 GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 2542
Full_length GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 4513
Short_3_end_UTR
SSCs_specific GGCCCAGAAGGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 4067
Tr-kit_c18-4_2.7kb GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 2070
Tr-kit_c18-4_2.9kb GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 2286
Tr-kit_c18-4_4.0kb GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 3383
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 2472
Tr-kit_CRL2053_3.9kb GGCCCAGA-GGTTCCCTTGTCCTTGCCCTCCCCCACACCACCCATGATCACTGTCCAACA 3244
NM_001122733.1 TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 4590
NM_021099.3 TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 4574
X65997.1 TAAAGGGTTCAGTGTGTTACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 2602
Full_length TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 4572
Short_3_end_UTR
SSCs_specific TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 4126
Tr-kit_c18-4_2.7kb TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 2129
Tr-kit_c18-4_2.9kb TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 2345
Tr-kit_c18-4_4.0kb TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 3442
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 2531
Tr-kit_CRL2053_3.9kb TAAAGGGTTCAGTGTGT-ACGTGGTCATGTGTTGTCCTTACAGGATTCAGGTATGTTGCC 3303
NM_001122733.1 TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 4650
NM_021099.3 TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 4634
X65997.1 TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 2662


Full_length TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 4632
Short_3_end_UTR
SSCs_specific TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 4186
Tr-kit_c18-4_2.7kb TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 2189
Tr-kit_c18-4_2.9kb TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 2405
Tr-kit_c18-4_4.0kb TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 3502
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 2591
Tr-kit_CRL2053_3.9kb TTCACGGTTTTCCCCACCCCCTCCTGCCCTTTATCCTTTAGGCCGTGTGGCCATGAACCT 3363
NM_001122733.1 GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 4710
NM_021099.3 GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 4694
X65997.1 GGAAGAAGTGATCGTTTCGACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 2722
Full_length GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 4692
Short_3_end_UTR
SSCs_specific GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 4246
Tr-kit_c18-4_2.7kb GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 2249
Tr-kit_c18-4_2.9kb GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 2465
Tr-kit_c18-4_4.0kb GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 3562
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 2651
Tr-kit_CRL2053_3.9kb GGAAGAAGTGATCGTTTGCACTTGAGTGCTACACTCTTGCACCTTTCCAAAGTAAGCTGG 3423
NM_001122733.1 TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 4770
NM_021099.3 TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 4754
X65997.1 TTTGGAGGTCCTGTGGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 2782
Full_length TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 4752
Short_3_end_UTR
SSCs_specific TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 4306
Tr-kit_c18-4_2.7kb TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 2309
Tr-kit_c18-4_2.9kb TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 2525
Tr-kit_c18-4_4.0kb TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 3622
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 2711
Tr-kit_CRL2053_3.9kb TTTGGAGGTCCTGTTGTCATGTACGAGACTGTCACCAGTTACCGCGCTCTGTTTGAAACA 3483
NM_001122733.1 TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 4830
NM_021099.3 TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 4814
X65997.1 TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 2842
Full_length TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 4812
Short_3_end_UTR
SSCs_specific TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 4366
Tr-kit_c18-4_2.7kb TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 2369
Tr-kit_c18-4_2.9kb TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 2585
Tr-kit_c18-4_4.0kb TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 3682
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 2771
Tr-kit_CRL2053_3.9kb TGTCTTTGTATTCCTAATGACTTCAGTTAGAGTAAGGAGAATAGCTGTTAATATGGATGT 3543
NM_001122733.1 CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 4890
NM_021099.3 CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 4874


X65997.1 CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 2902
Full_length CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 4872
Short_3_end_UTR
SSCs_specific CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 4426
Tr-kit_c18-4_2.7kb CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 2429
Tr-kit_c18-4_2.9kb CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 2645
Tr-kit_c18-4_4.0kb CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 3742
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 2831
Tr-kit_CRL2053_3.9kb CAGGTACTTAAGGGGCCACACCATTGAGAATTTTGTCTTGGATATTCTTGAAAGTTTATA 3603
NM_001122733.1 TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 4950
NM_021099.3 TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 4934
X65997.1 TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 2962
Full_length TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 4932
Short_3_end_UTR
SSCs_specific TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 4486
Tr-kit_c18-4_2.7kb TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 2489
Tr-kit_c18-4_2.9kb TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 2705
Tr-kit_c18-4_4.0kb TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 3802
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 2891
Tr-kit_CRL2053_3.9kb TTTTTATAATTTTTTTTACATCAGATGTCAGATGTTTCTTTCAGTTGCTTGATGTTTGGA 3663
NM_001122733.1 ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 5010
NM_021099.3 ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 4994
X65997.1 ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 3022
Full_length ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 4992
Short_3_end_UTR
SSCs_specific ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 4546
Tr-kit_c18-4_2.7kb ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 2549
Tr-kit_c18-4_2.9kb ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 2765
Tr-kit_c18-4_4.0kb ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 3862
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 2951
Tr-kit_CRL2053_3.9kb ATTATTATGTGGCTTTTTTTGTAAATATTGAAATGTAGCAATAATGTCTTTTGAATATTC 3723
NM_001122733.1 CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 5070
NM_021099.3 CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 5054
X65997.1 CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 3082
Full_length CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 5052
Short_3_end_UTR
SSCs_specific CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 4606
Tr-kit_c18-4_2.7kb CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 2609
Tr-kit_c18-4_2.9kb CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 2825
Tr-kit_c18-4_4.0kb CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 3922
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 3011
Tr-kit_CRL2053_3.9kb CTGAGCCCATGAGTCCCTGAAAATATTTTTTATATATACAGTAACTTTATGTGTAAATAA 3783
NM_001122733.1 TACGCTGTGCAAGTTTAAACATGTCACGTTACATGTGGGTTTTTTCTGATATGTTGTCCA 5130



Legend

NM_001122733.1: Mus musculus kit oncogene, transcript variant 1 reference sequence;

NM_021099.3: Mus musculus kit oncogene, transcript variant 2 reference sequence;

X65997.1: M.musculus *c-kit* mRNA for truncated tyrosine-kinase, Tr-kit;

Full_length: *c-kit* full length transcript we got from RACE and sequencing;

Short_3_end_UTR: *c-kit* full length transcript with a short 3′ UTR;

SSCs_specific: transcript only expressed in undifferentiated spermatogonia;

Tr-kit_c18-4_2.7kb: truncated *c-kit* transcript found in c18-4, composed by exons 17-21;

Tr-kit_c18-4_2.9kb: truncated *c-kit* transcript found in c18-4, composed by exons 15-21;

Tr-kit_c18-4_4.0kb: truncated *c-kit* transcript found in c18-4, composed by intron 17-exon 21;

Tr-kit_CRL2053_1.9kb: truncated *c-kit* transcript found in CRL-2053, composed by exons 13-21, with a short 3′ UTR;

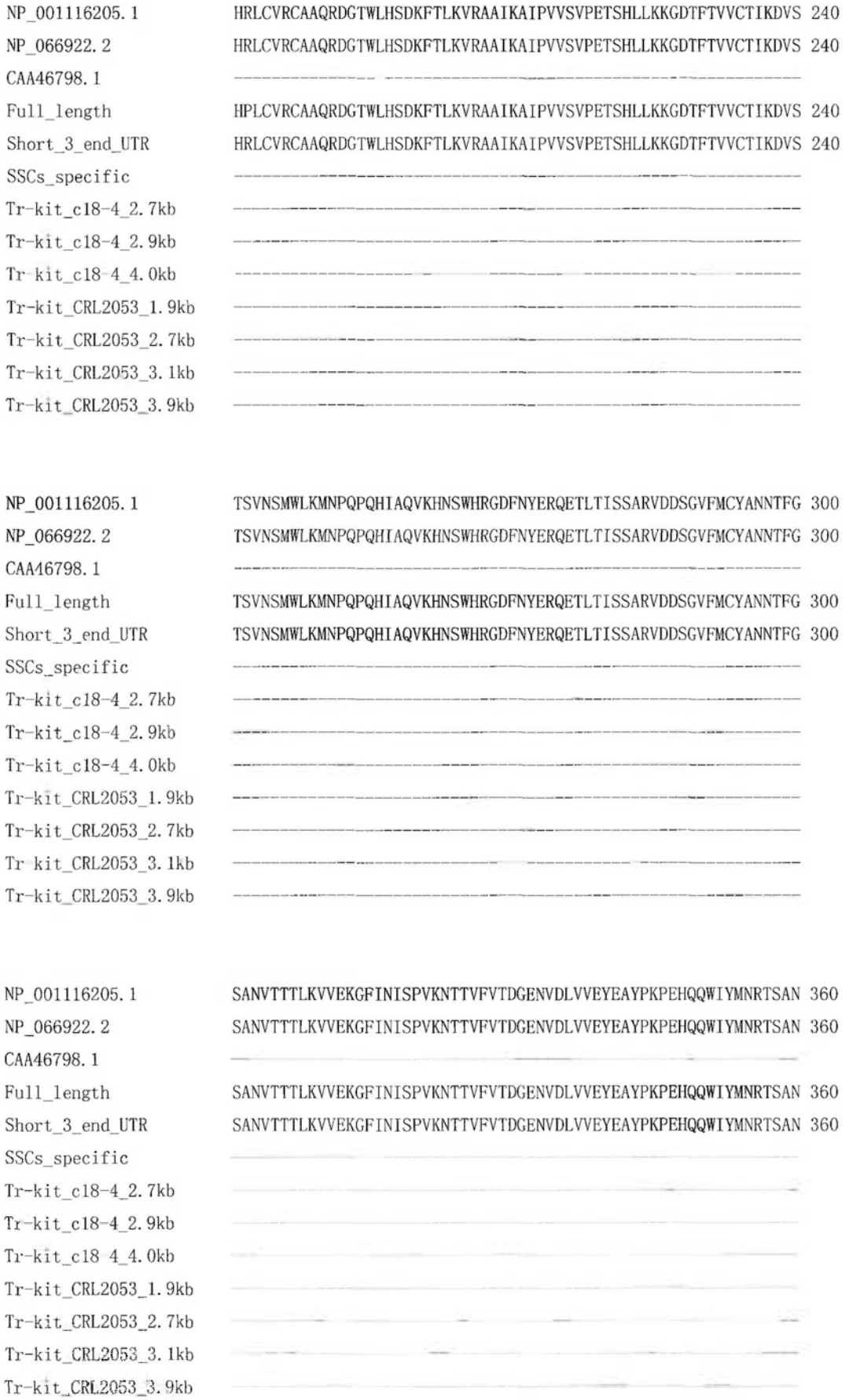
Tr-kit_CRL2053_2.7kb: truncated *c-kit* transcript found in CRL-2053, composed by exons 1-5 and exons 14-21, with a short 3′ UTR;

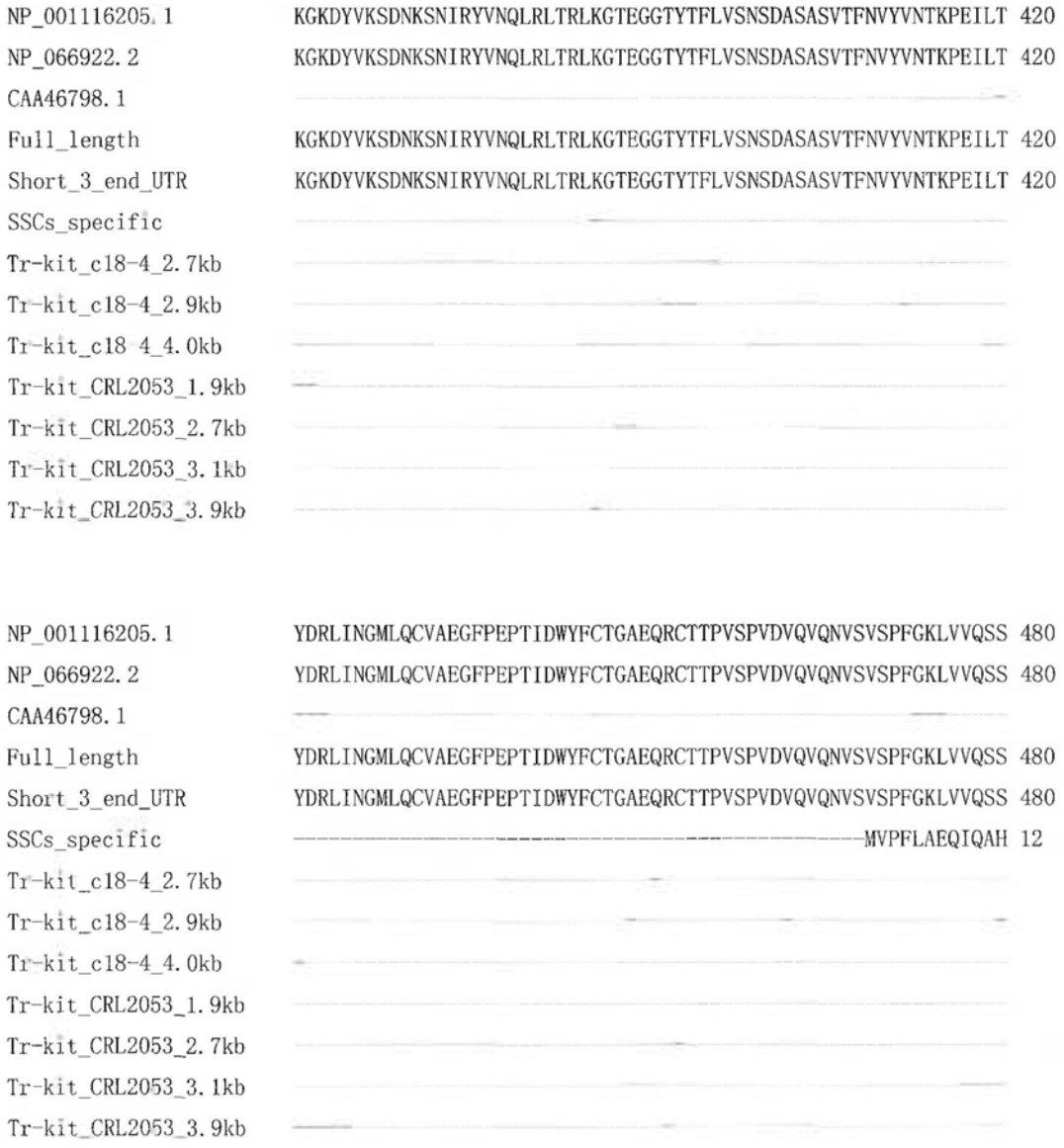
Tr-kit_CRL2053_3.1kb: truncated *c-kit* transcript found in CRL-2053, composed by exons 13-21;

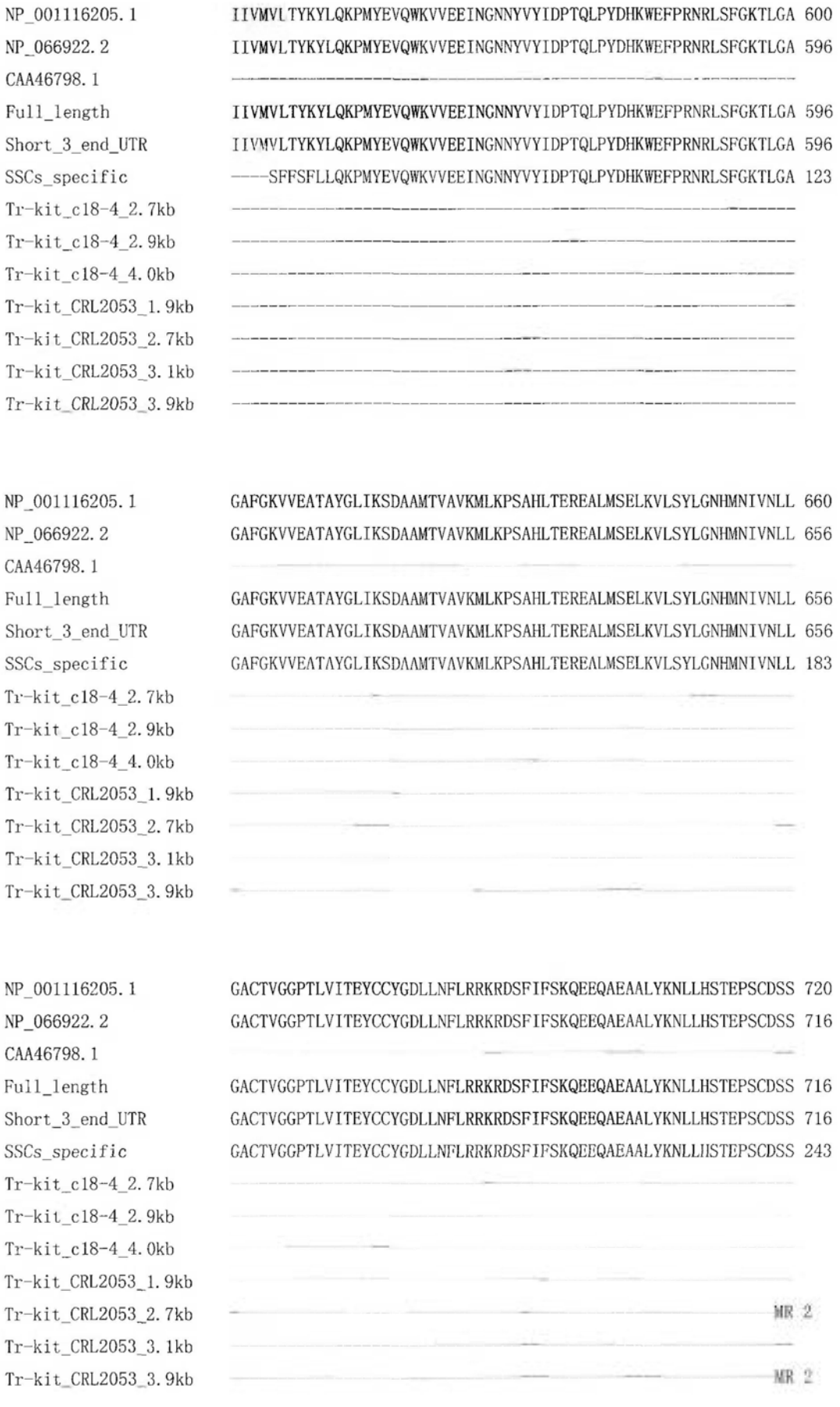
Tr-kit_CRL2053_3.9kb: truncated *c-kit* transcript found in CRL-2053, composed by exons 1-5 and exons 14-21.

Appendix 2 Multiple sequence alignment of ORF finder predicted *c-kit* proteins

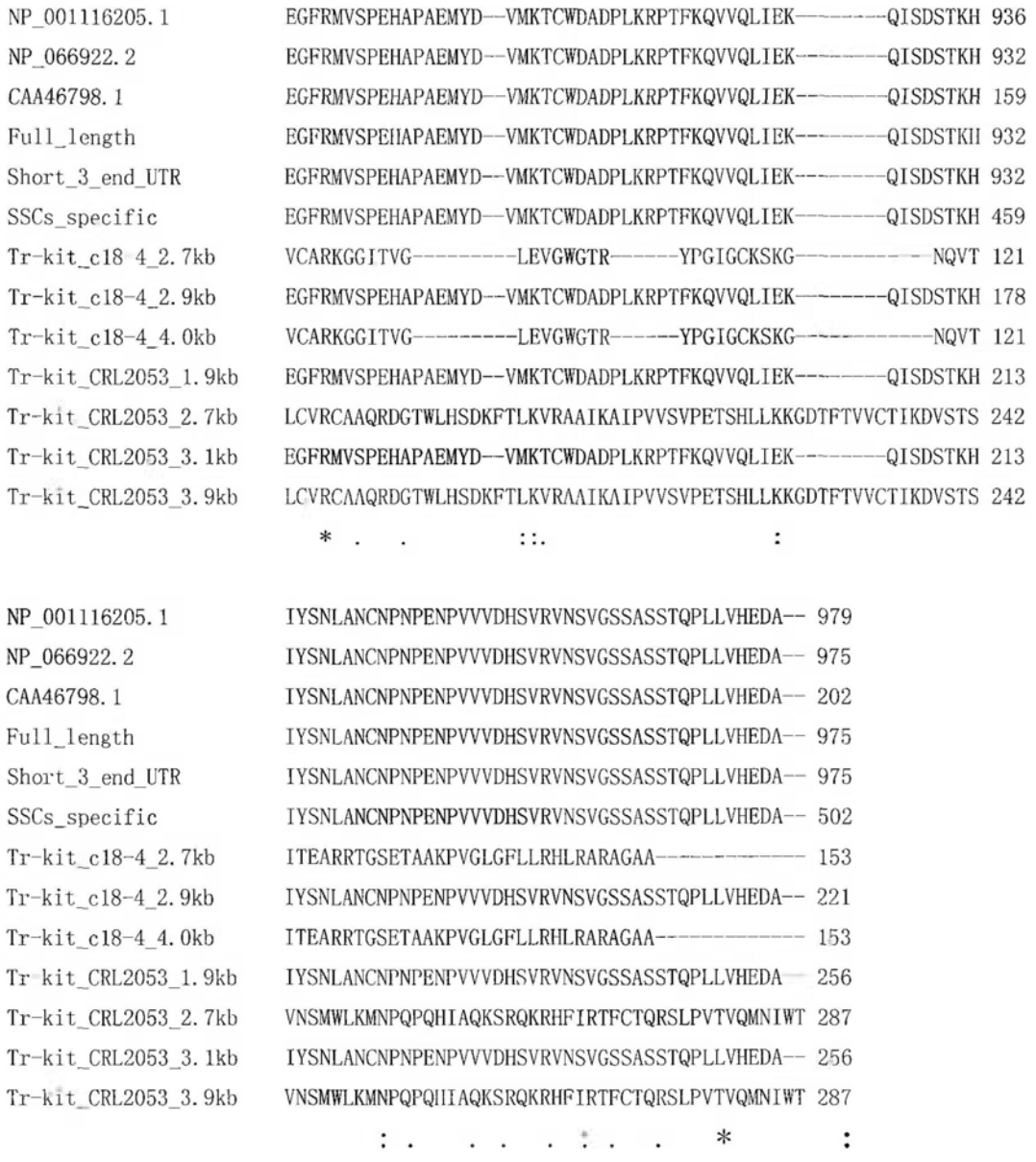
```
NP_001116205.1          MRGARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCID 60
NP_066922.2             MRGARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCID 60
CAA46798.1              ------------------------------------------------------------
Full_length             MRGARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCID 60
Short_3_end_UTR         MRGARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCID 60
SSCs_specific           ------------------------------------------------------------
Tr kit_c18-4_2.7kb      ------------------------------------------------------------
Tr-kit_c18-4_2.9kb      ------------------------------------------------------------
Tr-kit_c18-4_4.0kb      ------------------------------------------------------------
Tr-kit_CRL2053_1.9kb    ------------------------------------------------------------
Tr-kit_CRL2053_2.7kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.1kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb    ------------------------------------------------------------

NP_001116205.1          PDFVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFL 120
NP_066922.2             PDFVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFL 120
CAA46798.1              ------------------------------------------------------------
Full_length             PDFVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFL 120
Short_3_end_UTR         PDFVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFL 120
SSCs_specific           ------------------------------------------------------------
Tr-kit_c18-4_2.7kb      ------------------------------------------------------------
Tr-kit_c18-4_2.9kb      ------------------------------------------------------------
Tr-kit_c18-4_4.0kb      ------------------------------------------------------------
Tr-kit_CRL2053_1.9kb    ------------------------------------------------------------
Tr-kit_CRL2053_2.7kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.1kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb    ------------------------------------------------------------

NP_001116205.1          VGLPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAY 180
NP_066922.2             VGLPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAY 180
CAA46798.1              ------------------------------------------------------------
Full_length             VGLPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAY 180
Short_3_end_UTR         VGLPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAY 180
SSCs_specific           ------------------------------------------------------------
Tr-kit_c18-4_2.7kb      ------------------------------------------------------------
Tr-kit_c18-4_2.9kb      ------------------------------------------------------------
Tr kit_c18-4_4.0kb      ------------------------------------------------------------
Tr-kit_CRL2053_1.9kb    ------------------------------------------------------------
Tr-kit_CRL2053_2.7kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.1kb    ------------------------------------------------------------
Tr-kit_CRL2053_3.9kb    ------------------------------------------------------------
```


NP_001116205.1 HRLCVRCAAQRDGTWLHSDKFTLKVRAAIKAIPVVSVPETSHLLKKGDTFTVVCTIKDVS 240
NP_066922.2 HRLCVRCAAQRDGTWLHSDKFTLKVRAAIKAIPVVSVPETSHLLKKGDTFTVVCTIKDVS 240
CAA46798.1
Full_length HPLCVRCAAQRDGTWLHSDKFTLKVRAAIKAIPVVSVPETSHLLKKGDTFTVVCTIKDVS 240
Short_3_end_UTR HRLCVRCAAQRDGTWLHSDKFTLKVRAAIKAIPVVSVPETSHLLKKGDTFTVVCTIKDVS 240
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 TSVNSMWLKMNPQPQHIAQVKHNSWHRGDFNYERQETLTISSARVDDSGVFMCYANNTFG 300
NP_066922.2 TSVNSMWLKMNPQPQHIAQVKHNSWHRGDFNYERQETLTISSARVDDSGVFMCYANNTFG 300
CAA46798.1
Full_length TSVNSMWLKMNPQPQHIAQVKHNSWHRGDFNYERQETLTISSARVDDSGVFMCYANNTFG 300
Short_3_end_UTR TSVNSMWLKMNPQPQHIAQVKHNSWHRGDFNYERQETLTISSARVDDSGVFMCYANNTFG 300
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 SANVTTTLKVVEKGFINISPVKNTTVFVTDGENVDLVVEYEAYPKPEHQQWIYMNRTSAN 360
NP_066922.2 SANVTTTLKVVEKGFINISPVKNTTVFVTDGENVDLVVEYEAYPKPEHQQWIYMNRTSAN 360
CAA46798.1
Full_length SANVTTTLKVVEKGFINISPVKNTTVFVTDGENVDLVVEYEAYPKPEHQQWIYMNRTSAN 360
Short_3_end_UTR SANVTTTLKVVEKGFINISPVKNTTVFVTDGENVDLVVEYEAYPKPEHQQWIYMNRTSAN 360
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb


NP_001116205.1 KGKDYVKSDNKSNIRYVNQLRLTRLKGTEGGTYTFLVSNSDASASVTFNVYVNTKPEILT 420
NP_066922.2 KGKDYVKSDNKSNIRYVNQLRLTRLKGTEGGTYTFLVSNSDASASVTFNVYVNTKPEILT 420
CAA46798.1
Full_length KGKDYVKSDNKSNIRYVNQLRLTRLKGTEGGTYTFLVSNSDASASVTFNVYVNTKPEILT 420
Short_3_end_UTR KGKDYVKSDNKSNIRYVNQLRLTRLKGTEGGTYTFLVSNSDASASVTFNVYVNTKPEILT 420
SSCs_specific
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 YDRLINGMLQCVAEGFPEPTIDWYFCTGAEQRCTTPVSPVDVQVQNVSVSPFGKLVVQSS 480
NP_066922.2 YDRLINGMLQCVAEGFPEPTIDWYFCTGAEQRCTTPVSPVDVQVQNVSVSPFGKLVVQSS 480
CAA46798.1
Full_length YDRLINGMLQCVAEGFPEPTIDWYFCTGAEQRCTTPVSPVDVQVQNVSVSPFGKLVVQSS 480
Short_3_end_UTR YDRLINGMLQCVAEGFPEPTIDWYFCTGAEQRCTTPVSPVDVQVQNVSVSPFGKLVVQSS 480
SSCs_specific MVPFLAEQIQAH 12
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 IDSSVFRHNGTVECKASNDVGKSSAFFNFAFKGNNKEQIQAHTLFTPLLIGFVVAAGAMG 540
NP_066922.2 IDSSVFRHNGTVECKASNDVGKSSAFFNFAFK----EQIQAHTLFTPLLIGFVVAAGAMG 536
CAA46798.1
Full_length IDSSVFRHNGTVECKASNDVGKSSAFFNFAFK----EQIQAHTLFTPLLIGFVVAAGAMG 536
Short_3_end_UTR IDSSVFRHNGTVECKASNDVGKSSAFFNFAFK----EQIQAHTLFTPLLIGFVVAAGAMG 536
SSCs_specific TLFTPLLIGFVVAAGAMGIIVMVLTYKYLQVS----IELFSSWGRQGGRAGTDCSAGDT- 67
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb


NP_001116205.1 IIVMVLTYKYLQKPMYEVQWKVVEEINGNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGA 600
NP_066922.2 IIVMVLTYKYLQKPMYEVQWKVVEEINGNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGA 596
CAA46798.1
Full_length IIVMVLTYKYLQKPMYEVQWKVVEEINGNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGA 596
Short_3_end_UTR IIVMVLTYKYLQKPMYEVQWKVVEEINGNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGA 596
SSCs_specific SFFSFLLQKPMYEVQWKVVEEINGNNYVYIDPTQLPYDHKWEFPRNRLSFGKTLGA 123
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 GAFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMNIVNLL 660
NP_066922.2 GAFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMNIVNLL 656
CAA46798.1
Full_length GAFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMNIVNLL 656
Short_3_end_UTR GAFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMNIVNLL 656
SSCs_specific GAFGKVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSYLGNHMNIVNLL 183
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb
NP_001116205.1 GACTVGGPTLVITEYCCYGDLLNFLRRKRDSFIFSKQEEQAEAALYKNLLHSTEPSCDSS 720
NP_066922.2 GACTVGGPTLVITEYCCYGDLLNFLRRKRDSFIFSKQEEQAEAALYKNLLHSTEPSCDSS 716
CAA46798.1
Full_length GACTVGGPTLVITEYCCYGDLLNFLRRKRDSFIFSKQEEQAEAALYKNLLHSTEPSCDSS 716
Short_3_end_UTR GACTVGGPTLVITEYCCYGDLLNFLRRKRDSFIFSKQEEQAEAALYKNLLHSTEPSCDSS 716
SSCs_specific GACTVGGPTLVITEYCCYGDLLNFLRRKRDSFIFSKQEEQAEAALYKNLLHSTEPSCDSS 243
Tr-kit_c18-4_2.7kb
Tr-kit_c18-4_2.9kb
Tr-kit_c18-4_4.0kb
Tr-kit_CRL2053_1.9kb
Tr-kit_CRL2053_2.7kb MR 2
Tr-kit_CRL2053_3.1kb
Tr-kit_CRL2053_3.9kb MR 2

```
NP_001116205.1          NEYMDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 780
NP_066922.2             NEYMDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 776
CAA46798.1              ------------------------------------------------------MAVA 4
Full_length             NEYMDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 776
Short_3_end_UTR         NEYMDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 776
SSCs_specific           NEYMDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 303
Tr-kit_c18-4_2.7kb      ------------------------------------------MALRAKWDYIYSSSE 15
Tr-kit_c18-4_2.9kb      ------------------------------------MEDDELALDLDDLLSFSYQVAK 22
Tr kit_c18-4_4.0kb      ------------------------------------MALRAKWDYIYSSSE 15
Tr-kit_CRL2053_1.9kb    ---MDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 57
Tr-kit_CRL2053_2.7kb    GARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCIDPD 62
Tr-kit_CRL2053_3.1kb    ---MDMKPGVSYVVPTKTDKRRSARIDSYIERDVTPAIMEDDELALDLDDLLSFSYQVAK 57
Tr-kit_CRL2053_3.9kb    GARGAWDLLCVLLVLLRGQTATSQPSASPGEPSPPSIHPAQSELIVEAGDTLSLTCIDPD 62

NP_001116205.1          GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 832
NP_066922.2             GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 828
CAA46798.1              VFPFLP-QQCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 55
Full_length             GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 828
Short_3_end_UTR         GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 828
SSCs_specific           GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 355
Tr-kit_c18-4_2.7kb      LLAILFKYTQGSTG----------GRTLGIQFHPALPFSSKAGGVLLFTVG--------- 56
Tr-kit_c18-4_2.9kb      GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 74
Tr-kit_c18-4_4.0kb      LLAILFKYTQGSTG----------GRTLGIQFHPALPFSSKAGGVLLFTVG--------- 56
Tr-kit_CRL2053_1.9kb    GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 109
Tr-kit_CRL2053_2.7kb    FVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFLVG 122
Tr-kit_CRL2053_3.1kb    GMAFLASKNCIHRDLAARN-ILLTHGRITKICDFGLAR-DIRNDSNYVVKGN-------AR 109
Tr-kit_CRL2053_3.9kb    FVRWTFKTYFNEMVENKKNEWIQEKAEATRTGTYTCSNSNGLTSSIYVFVRDPAKLFLVG 122

NP_001116205.1          LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 886
NP_066922.2             LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 882
CAA46798.1              LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 109
Full_length             LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 882
Short_3_end_UTR         LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 882
SSCs_specific           LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 409
Tr-kit_c18-4_2.7kb      ---------ATLLLGKYIHTVRTFAAGRWLMAEKKRPS-------------------ATK 88
Tr-kit_c18-4_2.9kb      LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 128
Tr kit_c18-4_4.0kb      -------ATLLLGKYIHTVRTFAAGRWLMAEKKRPS--------------ATK 88
Tr-kit_CRL2053_1.9kb    LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 163
Tr-kit_CRL2053_2.7kb    LPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAYHR 182
Tr-kit_CRL2053_3.1kb    LPVKWMAPESIFSCVYTFESDVWSYGIFLWELFSLGS-------SPYPGMPVDSKFYKMIK 163
Tr-kit_CRL2053_3.9kb    LPLFGKEDSDALVRCPLTDPQVSNYSLIECDGKSLPTDLTFVPNPKAGITIKNVKRAYHR 182
```



Legend

NP_001116205.1: mast/stem cell growth factor receptor isoform 1 [Mus musculus];

NP_066922.2: mast/stem cell growth factor receptor isoform 2 [Mus musculus];

CAA46798.1: truncated tyrosine kinase receptor [Mus musculus];

Full_length: protein predicted by ORF finder with the *c-kit* full length transcript we got from RACE and sequencing;

Short_3_end_UTR: protein predicted by ORF finder with the *c-kit* full length transcript with a short 3′ UTR;

SSCs_specific: protein predicted by ORF finder with the SSCs specific transcript.;

Tr-kit_c18-4_2.7kb: protein predicted by ORF finder with the 2.7 kb truncated *c-kit* transcript found in c18-4;

Tr-kit_c18-4_2.9kb: protein predicted by ORF finder with the 2.9 kb truncated *c-kit* transcript found in c18-4;

Tr-kit_c18-4_4.0kb: protein predicted by ORF finder with the 4.0 kb truncated *c-kit* transcript found in c18-4;

Tr-kit_Tr-kit_CRL2053_1.9kb: protein predicted by ORF finder with the 1.9 kb truncated *c-kit* transcript with a short 3′ UTR found in CRL-2053;

Tr-kit_Tr-kit_CRL2053_2.7kb: protein predicted by ORF finder with the 2.7 kb truncated *c-kit* transcript with a short 3′ UTR found in CRL-2053;

Tr-kit_CRL2053_3.1kb: protein predicted by ORF finder with the 3.1 kb truncated *c-kit* transcript found in CRL-2053;

Tr-kit_CRL2053_3.9kb: protein predicted by ORF finder with the 3.9 kb truncated *c-kit* transcript found in CRL-2053.

Transmembrane domain protein sequence was labeled with red color. Before the transmembrane domain is the extracellular domain; after the transmembrane domain is the intracellular domain.

Appendix 3 Absent region of short 3′UTR *c-kit* transcripts

ORIGIN

```
   1 TACTTGCTCT GTCTACACTT CGGAACAAGC CTTCCATGTC AGAGTGGCTT
  51 TGCAGGCAGG AGAACTGAGG CTGTTTGAAA AGGTTACCAC AGGATGGAGA
 101 AAACAGTGCA GTCCTGGTTT GGATTCTCAC ATAGCAGGGA GCACAAGTTA
 151 AACTCAGCCT TTTATAGGCA CGTCCCGGAC ATCGGGCCAG TATCTATTCA
 201 AGTGTGTATG TGTGTGCATG CGTGTGTCTA TGCGTGTGGG TGAGTTGTGT
 251 TGGGAAACTT GCCCTGCATC CCTGAGGGTC CTCCTTCAGG ACCCAAGACG
 301 TAACAGCTTC TGTCACCGCT CCTGTCTCTC CAGTTTCCCT GCATGTCGCT
 351 CACTGTCTAG AATTTACTCA AAGCCGCCAC AGAGGCTTAG CGGAGTGAAG
 401 TGCCGAAGGA CCTCTTTATT TGGAGTCCTC CTGTATTTAA CAACACTCTT
 451 ATCGTAGACC CATTCATTAG ACCTTATGTA ATGCTGCCAA TCCAGGGAAA
 501 CAGATTTAAA GTGTACCCCG TAGACAGGGC CCAGAGGTTC CCTTGTCCTT
 551 GCCCTCCCCC ACACCACCCA TGATCACTGT CCAACATAAA GGGTTCAGTG
 601 TGTACGTGGT CATGTGTTGT CCTTACAGGA TTCAGGTATG TTGCCTTCAC
 651 GGTTTTCCCC ACCCCCTCCT GCCCTTTATC CTTTAGGCCG TGTGGCCATG
 701 AACCTGGAAG AAGTGATCGT TTGCACTTGA GTGCTACACT CTTGCACCTT
 751 TCCAAAGTAA GCTGGTTTGG AGGTCCTGTT GTCATGTACG AGACTGTCAC
 801 CAGTTACCGC GCTCTGTTTG AAACATGTCT TTGTATTCCT AATGACTTCA
 851 GTTAGAGTAA GGAGAATAGC TGTTAATATG GATGTCAGGT ACTTAAGGGG
 901 CCACACCATT GAGAATTTTG TCTTGGATAT TCTTGAAAGT TTATATTTTT
 951 ATAATTTTTT TTACATCAGA TGTCAGATGT TTCTTTCAGT TGCTTGATGT
1001 TTGGAATTAT TATGTGGCTT TTTTTGTAAA TATTGAAATG TAGCAATAAT
1051 GTCTTTTGAA TATTCCTGAG CCCATGAGTC CCTGAAAATA TTTTTTATAT
1101 ATACAGTAAC TTTATGTGTA AATAATACGC TGTGCAAGTT TAAACATGTC
1151 ACGTTACATG TGGGTTTTTT CTGATATGTT GTCCAACTGT TGACAGTTCT
1201 GAAGAATTCT AATAAAAATG TAAATATATA AATC
```

//

References

Abel MH, Wootton AN, Wilkins V, Huhtaniemi I, Knight PG, Charlton HM. The effect of a null mutation in the follicle-stimulating hormone receptor gene on mouse reproduction. Endocrinology.2000;141:1795-1803.

Albanesi C, Geremia R, Giorgio M, Dolci S, Sette C, Rossi P. A cell- and developmental stage-specific promoter drives the expression of a truncated c-kit protein during mouse spermatid elongation. Development.1996;122:1291-1302.

Anderson DM, Lyman SD, Baird A, Wignall JM, Eisenman J, Rauch C, March CJ, Boswell HS, Gimpel SD, Cosman D, et a. Molecular cloning of mast cell growth factor, a hematopoietin that is active in both membrane bound and soluble forms. Cell.1990;63:235-243.

Andrews PW. Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. Dev Biol.1984;103:285-293.

Aponte PM, van BMP, de Rooij DG, van PAM. Spermatogonial stem cells: characteristics and experimental possibilities. APMIS.2005;113:727-742.

Asaoka-Taguchi M, Yamada M, Nakamura A, Hanyu K, Kobayashi S. Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. Nat Cell Biol.1999;1:431-437.

Baleato RM, Aitken RJ, Roman SD. Vitamin A regulation of BMP4 expression in the male germ line. Dev Biol.2005;286:78-90.

Ballow D, Meistrich ML, Matzuk M, Rajkovic A. Sohlh1 is essential for spermatogonial differentiation. Dev Biol.2006;294:161-167.

Baltus AE, Menke DB, Hu YC, Goodheart ML, Carpenter AE, de Rooij DG, Page DC. In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nat Genet.2006;38:1430-1434.

Barroca V, Lassalle B, Coureuil M, Louis JP, Le PF, Testart J, Allemand I, Riou L, Fouchet P. Mouse differentiating spermatogonia can generate germinal stem cells in vivo. Nat Cell Biol.2009;11:190-196.

Bashamboo A, Taylor AH, Samuel K, Panthier JJ, Whetton AD, Forrester LM. The survival of differentiating embryonic stem cells is dependent on the SCF-KIT pathway. J Cell Sci.2006;119:3039-3046.

Biermann K, Goke F, Nettersheim D, Eckert D, Zhou H, Kahl P, Gashaw I, Schorle H, Buttner R. c-KIT is frequently mutated in bilateral germ cell tumours and down-regulated during progression from intratubular germ cell neoplasia to

seminoma. J Pathol.2007;213:311-318.

Blechman JM, Lev S, Barg J, Eisenstein M, Vaks B, Vogel Z, Givol D, Yarden Y. The fourth immunoglobulin domain of the stem cell factor receptor couples ligand binding to signal transduction. Cell.1995;80:103-113.

Blechman JM, Lev S, Givol D, Yarden Y. Structure-function analyses of the kit receptor for the steel factor. Stem Cells.1993;11 Suppl 2:12-21.

Blume-Jensen P, Jiang G, Hyman R, Lee KF, O'Gorman S, Hunter T. Kit/stem cell factor receptor-induced activation of phosphatidylinositol 3'-kinase is essential for male fertility. Nat Genet.2000;24:157-162.

Bowles J, Knight D, Smith C, Wilhelm D, Richman J, Mamiya S, Yashiro K, Chawengsaksophak K, Wilson MJ, Rossant J, Hamada H, Koopman P. Retinoid signaling determines germ cell fate in mice. Science.2006;312:596-600.

Brinster RL. Male germline stem cells: from mice to men. Science.2007;316:404-405.

Buaas FW, Kirsh AL, Sharma M, McLean DJ, Morris JL, Griswold MD, de Rooij DG, Braun RE. Plzf is required in adult male germ cells for stem cell self-renewal. Nat Genet.2004;36:647-652.

Buehr M, Gu S, McLaren A. Mesonephric contribution to testis differentiation in the fetal mouse. Development.1993;117:273-281.

Carlomagno G, van BMP, Korver CM, Repping S, De Rooij DG, van PAM. BMP4-Induced Differentiation of a Rat Spermatogonial Stem Cell Line Causes Changes in Its Cell Adhesion Properties. Biol Reprod.2010.

Chabot B, Stephenson DA, Chapman VM, Besmer P, Bernstein A. The proto-oncogene c-kit encoding a transmembrane tyrosine kinase receptor maps to the mouse W locus. Nature.1988;335:88-89.

Chambon P. A decade of molecular biology of retinoic acid receptors. FASEB J.1996;10:940-954.

Chen C, Ouyang W, Grigura V, Zhou Q, Carnes K, Lim H, Zhao GQ, Arber S, Kurpios N, Murphy TL, Cheng AM, Hassell JA, Chandrashekar V, Hofmann MC, Hess RA, Murphy KM. ERM is required for transcriptional control of the spermatogonial stem cell niche. Nature.2005;436:1030-1034.

Chiba H, Clifford J, Metzger D, Chambon P. Distinct retinoid X receptor-retinoic acid receptor heterodimers are differentially involved in the control of expression of retinoid target genes in F9 embryonal carcinoma cells. Mol Cell Biol.1997;17:3013-3020.

Cook MS, Munger SC, Nadeau JH, Capel B. Regulation of male germ cell cycle arrest and differentiation by DND1 is modulated by genetic background. Development.2011;138:23-32.

Costoya JA, Hobbs RM, Barna M, Cattoretti G, Manova K, Sukhwani M, Orwig KE, Wolgemuth DJ, Pandolfi PP. Essential role of Plzf in maintenance of spermatogonial stem cells. Nat Genet.2004;36:653-659.

De Felici M, Di CA, Pesce M. Role of stem cell factor in somatic-germ cell interactions during prenatal oogenesis. Zygote.1996;4:349-351.

de Rooij DG. Stem cells in the testis. Int J Exp Pathol.1998;79:67-80.

de Rooij DG. Proliferation and differentiation of spermatogonial stem cells. Reproduction.2001;121:347-354.

Di CAD, Travia G, De Felici M. The meiotic specific synaptonemal complex protein SCP3 is expressed by female and male primordial germ cells of the mouse embryo. Int J Dev Biol.2000;44:241-244.

Dierich A, Sairam MR, Monaco L, Fimia GM, Gansmuller A, LeMeur M, Sassone-Corsi P. Impairing follicle-stimulating hormone (FSH) signaling in vivo: targeted disruption of the FSH receptor leads to aberrant gametogenesis and hormonal imbalance. Proc Natl Acad Sci U S A.1998;95:13612-13617.

Dolci S, Pellegrini M, Di AS, Geremia R, Rossi P. Signaling through extracellular signal-regulated kinase is required for spermatogonial proliferative response to stem cell factor. J Biol Chem.2001;276:40225-40233.

Durcova-Hills G, Capel B. Development of germ cells in the mouse. Curr Top Dev Biol.2008;83:185-212.

Dym M, Jia MC, Dirami G, Price JM, Rabin SJ, Mocchetti I, Ravindranath N. Expression of c-kit receptor and its autophosphorylation in immature rat type A spermatogonia. Biol Reprod.1995;52:8-19.

Edwards MK, McBurney MW. The concentration of retinoic acid determines the differentiated cell types formed by a teratocarcinoma cell line. Dev Biol.1983;98:187-191.

Eguizabal C, Shovlin TC, Durcova-Hills G, Surani A, McLaren A. Generation of primordial germ cells from pluripotent stem cells. Differentiation.2009;78:116-123.

Falender AE, Freiman RN, Geles KG, Lo KC, Hwang K, Lamb DJ, Morris PL, Tjian R, Richards JS. Maintenance of spermatogenesis requires TAF4b, a gonad-specific subunit of TFIID. Genes Dev.2005;19:794-803.

Farini D, La Sala G, Tedesco M, De Felici M. Chemoattractant action and molecular signaling pathways of Kit ligand on mouse primordial germ cells. Dev Biol.2007;306:572-583.

Feng LX, Ravindranath N, Dym M. Stem cell factor/c-kit up-regulates cyclin D3 and promotes cell cycle progression via the phosphoinositide 3-kinase/p70 S6 kinase pathway in spermatogonia. J Biol Chem.2000;275:25572-25576.

Filipponi D, Hobbs RM, Ottolenghi S, Rossi P, Jannini EA, Pandolfi PP, Dolci S. Repression of kit expression by Plzf in germ cells. Mol Cell Biol.2007;27:6770-6781.

Flanagan JG, Chan DC, Leder P. Transmembrane form of the kit ligand growth factor is determined by alternative splicing and is missing in the Sld mutant. Cell.1991;64:1025-1035.

Francavilla S, Zamboni L. Differentiation of mouse ectopic germinal cells in intra- and perigonadal locations. J Exp Zool.1985;233:101-109.

Ginsburg M, Snow MH, McLaren A. Primordial germ cells in the mouse embryo during gastrulation. Development.1990;110:521-528.

Godin I, Deed R, Cooke J, Zsebo K, Dexter M, Wylie CC. Effects of the steel gene product on mouse primordial germ cells in culture. Nature.1991;352:807-809.

Gu Y, Runyan C, Shoemaker A, Surani A, Wylie C. Steel factor controls primordial germ cell survival and motility from the time of their specification in the allantois, and provides a continuous niche throughout their migration. Development.2009;136:1295-1303.

Guerif F, Cadoret V, Rahal-Perola V, Lansac J, Bernex F, Jacques PJ, Hochereau-de RMT, Royere D. Apoptosis, onset and maintenance of spermatogenesis: evidence for the involvement of Kit in Kit-haplodeficient mice. Biol Reprod.2002;67:70-79.

Hamra FK, Chapman KM, Nguyen DM, Williams-Stephens AA, Hammer RE, Garbers DL. Self renewal, expansion, and transfection of rat spermatogonial stem cells in culture. Proc Natl Acad Sci U S A.2005;102:17430-17435.

Hofmann MC, Abramian D, Millan JL. A haploid and a diploid cell coexist in an in vitro immortalized spermatogenic cell line. Dev Genet.1995;16:119-127.

Hofmann MC, Braydich-Stolle L, Dettin L, Johnson E, Dym M. Immortalization of mouse germ line stem cells. Stem Cells.2005;23:200-210.

Hofmann MC, Hess RA, Goldberg E, Millan JL. Immortalized germ cells undergo meiosis in vitro. Proc Natl Acad Sci U S A.1994;91:5533-5537.

Hofmann MC, Narisawa S, Hess RA, Millan JL. Immortalization of germ cells and somatic testicular cells using the SV40 large T antigen. Exp Cell Res.1992;201:417-435.

Hogarth CA, Griswold MD. The key role of vitamin A in spermatogenesis. J Clin Invest.2010;120:956-962.

Izadyar F, Pau F, Marh J, Slepko N, Wang T, Gonzalez R, Ramos T, Howerton K, Sayre C, Silva F. Generation of multipotent cell lines from a distinct population of male germ line stem cells. Reproduction.2008;135:771-784.

Jones-Villeneuve EM, McBurney MW, Rogers KA, Kalnins VI. Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. J Cell Biol.1982;94:253-262.

Kanatsu-Shinohara M, Ogonuki N, Inoue K, Miki H, Ogura A, Toyokuni S, Shinohara T. Long-term proliferation in culture and germline transmission of mouse male germline stem cells. Biol Reprod.2003;69:612-616.

Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L, Boiani M, Lomeli H, Nagy A, McLaughlin KJ, Scholer HR, Tomilin A. Oct4 is required for primordial germ cell survival. EMBO Rep.2004;5:1078-1083.

Kokkinaki M, Lee TL, He Z, Jiang J, Golestaneh N, Hofmann MC, Chan WY, Dym M. The molecular signature of spermatogonial stem/progenitor cells in the 6-day-old mouse testis. Biol Reprod.2009;80:707-717.

Koubova J, Menke DB, Zhou Q, Capel B, Griswold MD, Page DC. Retinoic acid regulates sex-specific timing of meiotic initiation in mice. Proc Natl Acad Sci U S A.2006;103:2474-2479.

Krishnamurthy H, Danilovich N, Morales CR, Sairam MR. Qualitative and quantitative decline in spermatogenesis of the follicle-stimulating hormone receptor knockout (FORKO) mouse. Biol Reprod.2000;62:1146-1159.

Kubota H, Avarbock MR, Brinster RL. Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. Proc Natl Acad Sci U S A.2004;101:16489-16494.

Kubota H, Avarbock MR, Schmidt JA, Brinster RL. Spermatogonial stem cells derived from infertile Wv/Wv mice self-renew in vitro and generate progeny following transplantation. Biol Reprod.2009;81:293-301.

Kumar TR, Palapattu G, Wang P, Woodruff TK, Boime I, Byrne MC, Matzuk MM. Transgenic models to study gonadotropin function: the role of follicle-stimulating hormone in gonadal growth and tumorigenesis. Mol Endocrinol.1999;13:851-865.

Kumar TR, Wang Y, Lu N, Matzuk MM. Follicle stimulating hormone is required for ovarian follicle maturation but not male fertility. Nat Genet.1997;15:201-204.

Lev S, Yarden Y, Givol D. Dimerization and activation of the kit receptor by monovalent and bivalent binding of the stem cell factor. J Biol Chem.1992;267:15970-15977.

Li H, MacLean G, Cameron D, Clagett-Dame M, Petkovich M. Cyp26b1 expression in murine Sertoli cells is required to maintain male germ cells in an undifferentiated state during embryogenesis. PLoS One.2009;4:e7501.

Livera G, Rouiller-Fabre V, Pairault C, Levacher C, Habert R. Regulation and perturbation of testicular functions by vitamin A. Reproduction.2002;124:173-180.

Lolicato F, Marino R, Paronetto MP, Pellegrini M, Dolci S, Geremia R, Grimaldi P. Potential role of Nanos3 in maintaining the undifferentiated spermatogonia population. Dev Biol.2008;313:725-738.

Lufkin T, Lohnes D, Mark M, Dierich A, Gorry P, Gaub MP, LeMeur M, Chambon P. High postnatal lethality and testis degeneration in retinoic acid receptor alpha mutant mice. Proc Natl Acad Sci U S A.1993;90:7225-7229.

MacLean G, Li H, Metzger D, Chambon P, Petkovich M. Apoptotic extinction of germ cells in testes of Cyp26b1 knockout mice. Endocrinology.2007;148:4560-4567.

Manova K, Bachvarova RF. Expression of c-kit encoded at the W locus of mice in developing embryonic germ cells and presumptive melanoblasts. Dev Biol.1991;146:312-324.

Mark M, Ghyselinck NB, Chambon P. Function of retinoid nuclear receptors: lessons from genetic and pharmacological dissections of the retinoic acid signaling pathway during mouse embryogenesis. Annu Rev Pharmacol Toxicol.2006;46:451-480.

Matsui Y, Zsebo KM, Hogan BL. Embryonic expression of a haematopoietic growth factor encoded by the Sl locus and the ligand for c-kit. Nature.1990;347:667-669.

McBurney MW. P19 embryonal carcinoma cells. Int J Dev Biol.1993;37:135-140.

McLaren A. Meiosis and differentiation of mouse germ cells. Symp Soc Exp Biol.1984;38:7-23.

Meng X, Lindahl M, Hyvonen ME, Parvinen M, de Rooij DG, Hess MW, Raatikainen-Ahokas A, Sainio K, Rauvala H, Lakso M, Pichel JG, Westphal H, Saarma M, Sariola H. Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. Science.2000;287:1489-1493.

Mithraprabhu S, Loveland KL. Control of KIT signalling in male germ cells: what

can we learn from other systems. Reproduction.2009;138:743-757.

Mol CD, Lim KB, Sridhar V, Zou H, Chien EY, Sang BC, Nowakowski J, Kassel DB, Cronin CN, McRee DE. Structure of a c-kit product complex reveals the basis for kinase transactivation. J Biol Chem.2003;278:31461-31464.

Morales C, Griswold MD. Retinol-induced stage synchronization in seminiferous tubules of the rat. Endocrinology.1987;121:432-434.

Morimoto H, Kanatsu-Shinohara M, Takashima S, Chuma S, Nakatsuji N, Takehashi M, Shinohara T. Phenotypic plasticity of mouse spermatogonial stem cells. PLoS One.2009;4:e7909.

Muciaccia B, Sette C, Paronetto MP, Barchi M, Pensini S, D'Agostino A, Gandini L, Geremia R, Stefanini M, Rossi P. Expression of a truncated form of KIT tyrosine kinase in human spermatozoa correlates with sperm DNA integrity. Hum Reprod.2010;25:2188-2202.

Nayernia K, Nolte J, Michelmann HW, Lee JH, Rathsack K, Drusenheimer N, Dev A, Wulf G, Ehrmann IE, Elliott DJ, Okpanyi V, Zechner U, Haaf T, Meinhardt A, Engel W. In vitro-differentiated embryonic stem cells give rise to male gametes that can generate offspring mice. Dev Cell.2006;11:125-132.

Niederreither K, Abu-Abed S, Schuhbaur B, Petkovich M, Chambon P, Dolle P. Genetic evidence that oxidative derivatives of retinoic acid are not involved in retinoid signaling during mouse development. Nat Genet.2002;31:84-88.

Niederreither K, Dolle P. Retinoic acid in development: towards an integrated view. Nat Rev Genet.2008;9:541-553.

OAKBERG EF. A description of spermiogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. Am J Anat.1956;99:391-413.

Oatley JM, Avarbock MR, Brinster RL. Glial cell line-derived neurotrophic factor regulation of genes essential for self-renewal of mouse spermatogonial stem cells is dependent on Src family kinase signaling. J Biol Chem.2007;282:25842-25851.

Oatley JM, Avarbock MR, Telaranta AI, Fearon DT, Brinster RL. Identifying genes important for spermatogonial stem cell self-renewal and survival. Proc Natl Acad Sci U S A.2006;103:9524-9529.

Oatley JM, Kaucher AV, Avarbock MR, Brinster RL. Regulation of mouse spermatogonial stem cell differentiation by STAT3 signaling. Biol Reprod.2010;83:427-433.

Oatley JM, Oatley MJ, Avarbock MR, Tobias JW, Brinster RL. Colony stimulating

factor 1 is an extrinsic stimulator of mouse spermatogonial stem cell self-renewal. Development.2009;136:1191-1199.

Ohbo K, Yoshida S, Ohmura M, Ohneda O, Ogawa T, Tsuchiya H, Kuwana T, Kehler J, Abe K, Scholer HR, Suda T. Identification and characterization of stem cells in prepubertal spermatogenesis in mice small star, filled. Dev Biol.2003;258:209-225.

Ohta H, Tohda A, Nishimune Y. Proliferation and differentiation of spermatogonial stem cells in the w/wv mutant mouse testis. Biol Reprod.2003;69:1815-1821.

Ohta H, Yomogida K, Dohmae K, Nishimune Y. Regulation of proliferation and differentiation in spermatogonial stem cells: the role of c-kit and its ligand SCF. Development.2000;127:2125-2131.

Onishi T, Machida T, Masuda F. [Histological evaluation of renal cell carcinoma. II. Histopathological features affecting prognosis]. Nippon Hinyokika Gakkai Zasshi.1983;74:2097-2106.

Orr-Urtreger A, Avivi A, Zimmer Y, Givol D, Yarden Y, Lonai P. Developmental expression of c-kit, a proto-oncogene encoded by the W locus. Development.1990;109:911-923.

Orth JM, Qiu J, Jester WF Jr, Pilder S. Expression of the c-kit gene is critical for migration of neonatal rat gonocytes in vitro. Biol Reprod.1997;57:676-683.

Packer AI, Besmer P, Bachvarova RF. Kit ligand mediates survival of type A spermatogonia and dividing spermatocytes in postnatal mouse testes. Mol Reprod Dev.1995;42:303-310.

Pellegrini M, Di SS, Claps G, Di CS, Dolci S, Rossi P, Geremia R, Grimaldi P. Microgravity promotes differentiation and meiotic entry of postnatal mouse male germ cells. PLoS One.2010;5:e9064.

Pellegrini M, Filipponi D, Gori M, Barrios F, Lolicato F, Grimaldi P, Rossi P, Jannini EA, Geremia R, Dolci S. ATRA and KL promote differentiation toward the meiotic program of male germ cells. Cell Cycle.2008;7:3878-3888.

Pellegrini M, Grimaldi P, Rossi P, Geremia R, Dolci S. Developmental expression of BMP4/ALK3/SMAD5 signaling pathway in the mouse testis: a potential role of BMP4 in spermatogonia differentiation. J Cell Sci.2003;116:3363-3372.

Pesce M, Gioia KF, De Felici M. Derivation in culture of primordial germ cells from cells of the mouse epiblast: phenotypic induction and growth control by Bmp4 signalling. Mech Dev.2002;112:15-24.

Prabhu SM, Meistrich ML, McLaughlin EA, Roman SD, Warne S, Mendis S, Itman C, Loveland KL. Expression of c-Kit receptor mRNA and protein in the developing,

adult and irradiated rodent testis. Reproduction.2006;131:489-499.

Roskoski R Jr. Signaling by Kit protein-tyrosine kinase--the stem cell factor receptor. Biochem Biophys Res Commun.2005;337:1-13.

Rossi P, Albanesi C, Grimaldi P, Geremia R. Expression of the mRNA for the ligand of c-kit in mouse Sertoli cells. Biochem Biophys Res Commun.1991;176:910-914.

Rossi P, Dolci S, Sette C, Geremia R. Molecular mechanisms utilized by alternative c-kit gene products in the control of spermatogonial proliferation and sperm-mediated egg activation. Andrologia.2003;35:71-78.

Rossi P, Lolicato F, Grimaldi P, Dolci S, Di SA, Filipponi D, Geremia R. Transcriptome analysis of differentiating spermatogonia stimulated with kit ligand. Gene Expr Patterns.2008;8:58-70.

Rossi P, Marziali G, Albanesi C, Charlesworth A, Geremia R, Sorrentino V. A novel c-kit transcript, potentially encoding a truncated receptor, originates within a kit gene intron in mouse spermatids. Dev Biol.1992;152:203-207.

Runyan C, Schaible K, Molyneaux K, Wang Z, Levin L, Wylie C. Steel factor controls midline cell death of primordial germ cells and is essential for their normal proliferation and migration. Development.2006;133:4861-4869.

Ryu BY, Kubota H, Avarbock MR, Brinster RL. Conservation of spermatogonial stem cell self-renewal signaling between mouse and rat. Proc Natl Acad Sci U S A.2005;102:14302-14307.

Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual, 3rd Edition. 2001.

Santos NC, Kim KH. Activity of retinoic acid receptor-alpha is directly regulated at its protein kinase A sites in response to follicle-stimulating hormone signaling. Endocrinology.2010;151:2361-2372.

Schmidt JA, Avarbock MR, Tobias JW, Brinster RL. Identification of glial cell line-derived neurotrophic factor-regulated genes important for spermatogonial stem cell self-renewal in the rat. Biol Reprod.2009;81:56-66.

Schnabel D, Ramirez L, Gertsenstein M, Nagy A, Lomeli H. Ectopic expression of KitD814Y in spermatids of transgenic mice, interferes with sperm morphogenesis. Dev Dyn.2005;233:29-40.

Scholer HR, Ciesiolka T, Gruss P. A nexus between Oct-4 and E1A: implications for gene regulation in embryonic stem cells. Cell.1991;66:291-304.

Schrans-Stassen BH, de Kant HJ v, de Rooij DG, van PAM. Differential expression

of c-kit in mouse undifferentiated and differentiating type A spermatogonia. Endocrinology.1999;140:5894-5900.

Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol.2005;278:440-458.

Sette C, Bevilacqua A, Bianchini A, Mangia F, Geremia R, Rossi P. Parthenogenetic activation of mouse eggs by microinjection of a truncated c-kit tyrosine kinase present in spermatozoa. Development.1997;124:2267-2274.

Sette C, Dolci S, Geremia R, Rossi P. The role of stem cell factor and of alternative c-kit gene products in the establishment, maintenance and function of germ cells. Int J Dev Biol.2000;44:599-608.

Sette C, Paronetto MP, Barchi M, Bevilacqua A, Geremia R, Rossi P. Tr-kit-induced resumption of the cell cycle in mouse eggs requires activation of a Src-like kinase. EMBO J.2002;21:5386-5395.

Shinohara T, Avarbock MR, Brinster RL. beta1- and alpha6-integrin are surface markers on mouse spermatogonial stem cells. Proc Natl Acad Sci U S A.1999;96:5504-5509.

Shinohara T, Orwig KE, Avarbock MR, Brinster RL. Spermatogonial stem cell enrichment by multiparameter selection of mouse testis cells. Proc Natl Acad Sci U S A.2000;97:8346-8351.

Sikarwar AP, Reddy KV. siRNA-mediated silencing of c-kit in mouse primary spermatogonial cells induces cell cycle arrest. Oligonucleotides.2008;18:145-160.

Snyder EM, Davis JC, Zhou Q, Evanoff R, Griswold MD. Exposure to Retinoic Acid in the Neonatal but Not Adult Mouse Results in Synchronous Spermatogenesis. Biol Reprod.2011.

Snyder EM, Small C, Griswold MD. Retinoic acid availability drives the asynchronous initiation of spermatogonial differentiation in the mouse. Biol Reprod.2010;83:783-790.

Sonoda J, Wharton RP. Recruitment of Nanos to hunchback mRNA by Pumilio. Genes Dev.1999;13:2704-2712.

Sorrentino V, Giorgi M, Geremia R, Besmer P, Rossi P. Expression of the c-kit proto-oncogene in the murine male germ cells. Oncogene.1991;6:149-151.

Suzuki A, Tsuda M, Saga Y. Functional redundancy among Nanos proteins and a distinct role of Nanos2 during male germ cell development.

Development.2007;134:77-83.

Suzuki H, Tsuda M, Kiso M, Saga Y. Nanos3 maintains the germ cell lineage in the mouse by suppressing both Bax-dependent and -independent apoptotic pathways. Dev Biol.2008;318:133-142.

Tadokoro Y, Yomogida K, Ohta H, Tohda A, Nishimune Y. Homeostatic regulation of germinal stem cell proliferation by the GDNF/FSH pathway. Mech Dev.2002;113:29-39.

Tajima Y, Onoue H, Kitamura Y, Nishimune Y. Biologically active kit ligand growth factor is produced by mouse Sertoli cells and is defective in SId mutant mice. Development.1991;113:1031-1035.

Tajima Y, Sawada K, Morimoto T, Nishimune Y. Switching of mouse spermatogonial proliferation from the c-kit receptor-independent type to the receptor-dependent type during differentiation. J Reprod Fertil.1994;102:117-122.

Tegelenbosch RA, de Rooij DG. A quantitative study of spermatogonial multiplication and stem cell renewal in the C3H/101 F1 hybrid mouse. Mutat Res.1993;290:193-200.

Toksoz D, Zsebo KM, Smith KA, Hu S, Brankow D, Suggs SV, Martin FH, Williams DA. Support of human hematopoiesis in long-term bone marrow cultures by murine stromal cells selectively expressing the membrane-bound and secreted forms of the human homolog of the steel gene product, stem cell factor. Proc Natl Acad Sci U S A.1992;89:7350-7354.

Toyoda S, Miyazaki T, Miyazaki S, Yoshimura T, Yamamoto M, Tashiro F, Yamato E, Miyazaki J. Sohlh2 affects differentiation of KIT positive oocytes and spermatogonia. Dev Biol.2009;325:238-248.

Trautmann E, Guerquin MJ, Duquenne C, Lahaye JB, Habert R, Livera G. Retinoic acid prevents germ cell mitotic arrest in mouse fetal testes. Cell Cycle.2008;7:656-664.

Trefil P, Bakst MR, Yan H, Hejnar J, Kalina J, Mucksova J. Restoration of spermatogenesis after transplantation of c-Kit positive testicular cells in the fowl. Theriogenology.2010;74:1670-1676.

Tsuda M, Sasaoka Y, Kiso M, Abe K, Haraguchi S, Kobayashi S, Saga Y. Conserved role of nanos proteins in germ cell development. Science.2003;301:1239-1241.

van PAM, de Rooij DG. Retinoic acid is able to reinitiate spermatogenesis in vitamin A-deficient rats and high replicate doses support the full development of spermatogenic cells. Endocrinology.1991;128:697-704.

Vaskivuo TE, Aittomaki K, Anttonen M, Ruokonen A, Herva R, Osawa Y, Heikinheimo M, Huhtaniemi I, Tapanainen JS. Effects of follicle-stimulating hormone (FSH) and human chorionic gonadotropin in individuals with an inactivating mutation of the FSH receptor. Fertil Steril.2002;78:108-113.

Vernet N, Dennefeld C, Guillou F, Chambon P, Ghyselinck NB, Mark M. Prepubertal testis development relies on retinoic acid but not rexinoid receptors in Sertoli cells. EMBO J.2006;25:5816-5825.

Vernet N, Dennefeld C, Rochette-Egly C, Oulad-Abdelghani M, Chambon P, Ghyselinck NB, Mark M. Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. Endocrinology.2006;147:96-110.

Vincent S, Segretain D, Nishikawa S, Nishikawa SI, Sage J, Cuzin F, Rassoulzadegan M. Stage-specific expression of the Kit receptor and its ligand (KL) during male gametogenesis in the mouse: a Kit-KL interaction critical for meiosis. Development.1998;125:4585-4593.

Vogel G. Embryonic stem cells. Scientists make sperm in a dish. Science.2003;302:1875.

Wang Y, Culty M. Identification and distribution of a novel platelet-derived growth factor receptor beta variant: effect of retinoic acid and involvement in cell differentiation. Endocrinology.2007;148:2233-2250.

Wong MD, Jin Z, Xie T. Molecular mechanisms of germline stem cell regulation. Annu Rev Genet.2005;39:173-195.

Wu JW, Wang RY, Guo QS, Xu C. Expression of the retinoic acid-metabolizing enzymes RALDH2 and CYP26b1 during mouse postnatal testis development. Asian J Androl.2008;10:569-576.

Yan W, Kero J, Huhtaniemi I, Toppari J. Stem cell factor functions as a survival factor for mature Leydig cells and a growth factor for precursor Leydig cells after ethylene dimethane sulfonate treatment: implication of a role of the stem cell factor/c-Kit system in Leydig cell development. Dev Biol.2000;227:169-182.

Yan W, Suominen J, Toppari J. Stem cell factor protects germ cells from apoptosis in vitro. J Cell Sci.2000;113 (Pt 1):161-168.

Yarden Y, Kuang WJ, Yang-Feng T, Coussens L, Munemitsu S, Dull TJ, Chen E, Schlessinger J, Francke U, Ullrich A. Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. EMBO J.1987;6:3341-3351.

Ying Y, Liu XM, Marble A, Lawson KA, Zhao GQ. Requirement of Bmp8b for the generation of primordial germ cells in the mouse. Mol

Endocrinol.2000;14:1053-1063.

Yoshinaga K, Nishikawa S, Ogawa M, Hayashi S, Kunisada T, Fujimoto T, Nishikawa S. Role of c-kit in mouse spermatogenesis: identification of spermatogonia as a specific site of c-kit expression and function. Development.1991;113:689-699.

Zhao GQ, Deng K, Labosky PA, Liaw L, Hogan BL. The gene encoding bone morphogenetic protein 8B is required for the initiation and maintenance of spermatogenesis in the mouse. Genes Dev.1996;10:1657-1669.

Zhou Q, Li Y, Nie R, Friel P, Mitchell D, Evanoff RM, Pouchnik D, Banasik B, McCarrey JR, Small C, Griswold MD. Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. Biol Reprod.2008;78:537-545.