ISOLATION AND CHARACTERIZATION OF THE NOVEL HUMAN GENE, MOST-1

JEANNE TAN MAY MAY

(B.Sc. (Hons), NUS)

A THESIS SUBMITTED FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MICROBIOLOGY

NATIONAL UNIVERSITY OF SINGAPORE

2004

ACKNOWLEDGEMENTS

Deepest appreciation to the following:

My supervisor, A/Prof Vincent Chow for this opportunity to pursue research and his constant encouragement.

A/Prof Bay Boon Huat and Prof Edward Tock for providing and help in grading the biopsies and their concern during my study.

Lecturers of the department especially A/P Yap Eu Hian, A/P Mulkit Singh, A/P Poh, A/P Lee Yuan Kun, Dr Mark, A/P Sim and Dr Song for their constant encouragement and guiding me through my chosen path.

A/P Wong Sek Man for letting me have the first encounter with Science.

All the staff of the department especially Mr Wee, Mr Lim, Mrs Phoon, Josephine, Joe and KT, Lip Chuan, Mayling, Mdm Chew, Mr Loh, Boon, Mr Chan, Goek Choo, Lini, Han Chong, Kim Lian, Ishak, Miss Siti, Mary and Geetha.

All my lab members especially William, Kingsley, Calvin, Shuwen and Jessie for their encouragement, friendship and help.

My course mates especially Nasir, Hongxiang, Shuxian, Meiling, Shirley, Justin, Peishan, Kenneth, Janice, Damien, Chew Leng, for being there.

My dearest friends Wee Ming, Del, Siao Yun, Kin Fai, Esther, Kai Soo, Jen Yen, Marieta, Han Liat, Yan Wing, Eng Hoe, Kailing, Sharon, Yen Lee, Jeanette for being there always through the ups and downs.

And most importantly of course, Dr Lim Kah Leong, Dr Soong Tuck Wah and Dr Wong Siew Heng and my NNI lab mates. Thanks for helping me with my presentation and guiding me in my thesis writing. Your concern and friendship really help me through the last few months.

Rocky for being there since I was six and taking all my crankiness.

Dad and Mom for being there for me always and supporting me through these years.
I thank God for you and just want to say I love you!

TABLE OF CONTENTS

CHAPTER 3: MATERIALS AND METHODS

LIST OF TABLES

LIST OF FIGURES

LIST OF GRAPHS

LIST OF ABBREVIATIONS

BrdU	Bromodeozyuridine
CAPS	3-cyclohexylamino-1-porpanesulfonic acid
Cdks	Cyclin-dependent kinases
CFS	Common fragile sites
CGH	Comparative genome hybridization
CK	Creatine Kinase
DMF	Dimethyl formamide
DEPC	Diethyl pyrocarbonate
EST	Expressed sequence tag
FCS	Fetal calf serum
FISH	Fluorescence in situ hybridization
G3DPH	Glyceraldehyde-3-phosphate dehydrogenase
HGP	Human Genome Project
HPV	Human papillomavirus
MPTP	Mitochondrial permeability transition pore
NASBA	Nucleic acid sequence based amplification
NP-40	Nonidet P-40
ORF	Open reading frame
PBR	Peripheral benzodiazepine receptor
PBS	Phosphate buffered saline
PCNA	Proliferating cellular nuclear antigen
PFA	Paraformaldehyde
PI	Propidium iodide
RACE	Rapid amplification of cDNA ends

RNAi	RNA interference
ROS	Reactive oxygen species
TdT	Terminal deoxynucleotidyl transferase
TE	Tris-EDTA
UTR	Untranslated region
V	Volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Y2H	Yeast-two hybrid
YPD	Yeast peptone dextrose

SUMMARY

Using PCR with human papillomavirus E6 gene primers, we amplified an expressed sequence tag from the MOLT-4 T-lymphoblastic leukemia cell line. Via RACE and cycle sequencing, we characterized overlapping cDNAs of 2786 bp and 2054 bp of the corresponding novel human intronless gene designated *MOST-1* (for MOLT-4 Sequence Tag-1) from MOLT-4 and fetal lung cDNA libraries, respectively. Both cDNAs contained a potential ORF of 297bp incorporating a methionine codon with an ideal Kozak consensus sequence for translation initiation, and encoding a putative hydrophilic polypeptide of 99 amino acids. Computational analysis of cDNA showed presence of 3 AUUUA mRNA destabilizing signals at its 3' untranslated region (UTR), suggesting *MOST-1* mRNA to be unstable. Additional computational analysis of putative ORF predicted MOST-1 protein to be unstable and non-globular with a secondary structure mainly of extended sheets.

Although RT-PCR demonstrated *MOST-1* expression in all 19 cancer and 2 normal cell lines tested, only differential expression was observed in 9 out of 16 normal tissues tested (heart, kidney, liver, pancreas, small intestine, ovary, testis, prostate and thymus).

The *MOST-1* gene was mapped by FISH to chromosome 8q24.2, a region amplified in many breast cancers and prostate cancers, and is also the candidate site of potential oncogene(s) other than *c-myc* located at 8q24.1. Analysis of paired biopsies of invasive ductal breast cancer and adjacent normal tissue by semi-quantitative and real-time RT-PCR revealed average tumor: normal ratios

of *MOST-1* expression that were two-fold greater in grade 3 cancers compared with grade 1 and 2 cancers. Quantitative real-time PCR of archival prostatic biopsies displayed *MOST-1* DNA levels that were 9.9, 7.5, 4.2 and 1.4 times higher respectively in high, intermediate, low grade carcinomas and benign hyperplasias than in normal samples.

In an attempt to elucidate MOST-1 function, a polyclonal antibody was raised. Characterization of the polyclonal antibody showed that it only recognizes the aggregated form of MOST-1 protein. Confocal immunofluorescence microscopy showed punctuate pattern of the MOST-1 aggregated protein in human cell lines namely hTERT-HME1 normal human mammary epithelial, MCF7 breast adenocarcinoma, PrEC normal human prostate epithelial and DU145 prostate carcinoma. Aggregation of overexpressed or misfolded proteins has been implicated in neurodegenerative disorder and many cancer types. Knock down of MOST-1 expression levels via RNA interference suggested that MOST-1 is needed for cancer cells proliferation. Yeast two-hybrid screening revealed interactions of MOST-1 with 8 partner proteins namely creatine kinase, ferritin, peripheral benzodiazepine receptor, immunoglobulin C (mu) and C (delta) heavy chain genes, SNC73 protein, Gardner feline sarcoma v-FGR and telethonin. Most of the interactors are reported to be amplified or deregulated in tumors with a majority involved in cell cycle or energy metabolism. Co-immunoprecipitation assays validated the interaction of MOST-1 with 3 of the proteins,

immunoglobulin C (mu) and C (delta) heavy chain, ferritin and peripheral benzodiazepine receptor.

Taken together, MOST-1 appears to be involved in cancer progression and its interaction with interactors involved in energy metabolism and cell cycle suggest a mitogenic function.

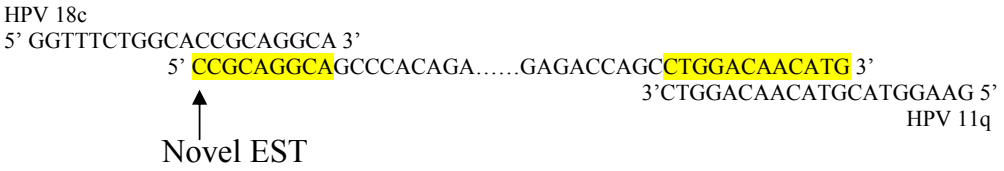
CHAPTER 1. INTRODUCTION

Following the publication of a working draft of the human genome sequence (Venter et al, 2001), the Human Genome Project (HGP) functions as a scaffold for the identification of the estimated 35,000 genes residing within three billion base pairs of DNA, the characterization of their regulatory elements, transcriptional units and translated products (Wright et al, 2001). Deregulation of gene expression result in cancer. Carcinogenesis has been shown to be a multifactor process in which genetic aberrations involving large amplicon containing multiple genes are often implicated (Ethier S, 2003). One of the ways to isolate these numerous expressed genes amidst large tracts of non-coding genomic DNA is the use of expressed sequence tags (ESTs) which represents an efficient and economical "short-cut" route for gene identification. The idea of exploiting ESTs has been established as a practical approach for the discovery of novel human genes (Adams et al, 1991; Sim and Chow, 1999). The search for ESTs and their corresponding genes implicated in the causation of human cancers is intensifying in the quest for better diagnostic markers and therapeutic agents (Strausberg, 2001; Onyango, 2002).

Since viral-induced cancers account for approximately 15% of human cancers, searching for genes deregulated by these viruses allows a directed search for potential genes involved in carcinogenesis. In particular, certain viruses have been shown to contribute significantly to the development of specific cancers such as the association of human papillomavirus (HPV) and carcinomas. Studies have shown that progression of HPV infected cells to

malignant phenotype requires further modifications of host gene expression; however molecular pathways underlying this phenomenon are still poorly understood despite epidemiological evidence (Kaufmann et al, 2002; Fiedler et al, 2004). In 1991, Couturier et al reported integration of HPV in cellular genomes near *myc* gene in genital cancers. This integration was found in most invasive genital carcinomas as compared to intraepithelial neoplasia where HPV DNA is detected most commonly as episomal molecules. This finding suggests a mechanism which may result in alteration of gene structure or overexpression of proto-oncogene. Subsequent work by Thorland et al in 2000 showed integration into genome to be non-random with HPV 16 integration to frequently occur at common fragile sites suggesting presence of chromosome 'Hot Spots' for viral integration. This also suggest that genes at or near the sites of integration may play an important role in tumor development as HPV integration could directly influence gene expression by changing the normal human DNA composition. Since HPV E6 early gene/oncoprotein of high-risk genital HPV types possess transforming abilities and are crucial in genital carcinogenesis (Chow et al, 2000; Stoler, 2000; Mantovani and Banks, 2001), this study was initiated in the view of isolating gene(s) near sites of HPV E6 integration using E6 consensus primers. Isolation and characterization of these genes would allow better elucidation of the underlying processes of carcinogenesis and subsequent therapeutics. MOLT-4 T-lymphoblastic leukemia cell line, a cancer cell line established directly from leukemia patient with relapse, with no viral integration reported (www.atcc.org), was chosen for

mRNA extraction so as to reduce background amplification of E6. RT-PCR of MOLT-4 RNA using primers targeting the E6 genes of HPV types 11 and 18 generated a novel EST of 350bp whose sequence revealed no significant homology to any known gene in the GenBank database and whose homology to HPV E6 primers as depicted below.



Arising from this novel EST which bears no homology to E6 except for the region indicated above, a study of isolation and characterization of a novel human gene was initiated. The objectives of this study were as follow:

1. To isolated full length cDNA;
2. To analyze the genomic structure of MOST-1;
3. To map its chromosome location;
4. To characterize its expression profiles in human tissues, cell lines and clinical biopsies; and
5. To produce polyclonal antibody for protein characterization.

CHAPTER 2. LITERATURE SURVEY

2.1 Human genome project – scaffold for functional genomics

Begun in 1990, the U.S. HGP was a 13-year effort to sequence the complete human genome. Along with it were project goals such as identification of all genes in human DNA, storing the information in databases, improving data analysis tools, transfer of technology to private sectors and to address the ethical, legal and social issues that may arise. The completion of sequencing has open up a new field of functional genomics into human health applications where genetics plays an important role in the diagnosis, monitoring and treatment of diseases. Medical genomics is at best at its infant stage as many genes are still under study as to how they contribute to the disease. The future challenge in genomics would be the elucidation of the function of each human gene. The goal after which would be to use the genetic information to develop new ways for prevention, treatment and cure. The next 20 years plan include the identification of more effective pharmaceuticals in which single base-pair variations in each individual can be used to

- accurately predict responses to drug, and environmental substances;
- anticipate disease susceptibility and aid in prevention;
- aid in organ cloning; and
- solve identity issues.

Of course the major downside of all these would be the ethics issue of social bias and human rights. The next immediate stages now involve the functional genomics technology whereby

- sets of full-length cDNA clones and sequences that represent human genes and model organisms will be generated,
- functional studies on nonprotein-coding sequences and its purpose in gene regulation;
- analysis of gene expression,
- genome-wide mutagenesis methodology development and
- large-scale protein analysis;

And the comparative genomics; which will encompass the complete sequencing of model organisms and appropriate genomic studies (adapted from www.onrl.gov).

With the sequence, the next challenge would be the identification of the various genes, validation of their structure and characterization of their functions. Even after the identification, the next would be to understand how the molecular components of the cells are controlled, interact and function as a system. As the era of molecular biology transcends from genomics to proteomics, progress in methodology in protein characterization reaches a new height with post translational modification becoming the centre stage of molecular biology. Post translation modification has important implications for protein conformation diseases arising from loss of their catalytic activity, structure and stability (Ishimaru et al, 2003). These disease have protein aggregates as hallmarks and the process of aggregation have been shown to be peptide (Milewski et al, 2002) and size specific (Diamant et al, 2000) suggesting that delicate balance is needed for normal cell function.

The discovery and ability to manipulate RNA interference (RNAi) in mammalian cell lines, a process where the introduction of double stranded RNA into a cell inhibits gene expression in a sequence-dependent way for gene silencing effect allow rapid functional studies to be carried out. This in turn accelerate the speed of discovering protein function to the cell in general as well as identification, characterization and development for new molecular targets for cancer in replacement for limited effective conventional treatment presently available (Jansen B. et al, 2002). The development of these methods allows not only individual protein function characterization but also showed an overview of the protein interactors and cellular function. The rampant use of yeast-two hybrid (Y2H) interaction screening allows novel protein-protein interaction to be characterized as well as providing an insight to novel protein function based on the characteristic of the interactors. These tools are timely as cancer research repeatedly and consistently shows that large amplicon that contain multiple genes which together causes a deregulation in cell cycle (Ethier S, 2003).

2.2 Genome research

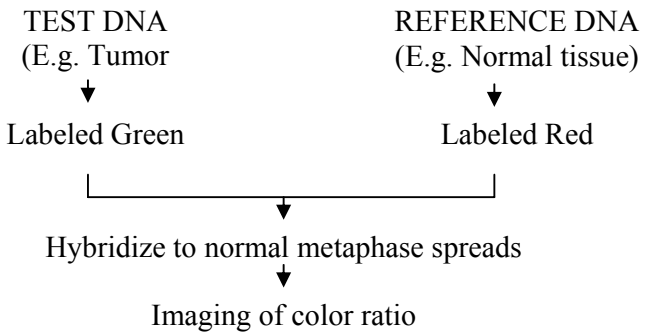
Genome research has taken off in leaps over the last decade with many techniques available for genome wide screening of gene copy number, expression and structure. There are basically 2 groups of techniques, the molecular cytogenetic group such as comparative genome hybridization (CGH) and FISH, and the molecular genetic techniques such as differential display and

microarray. Of these, CGH has become one of the popular genome scanning techniques for cancer research as it allows easy screening for DNA sequence copy number changes (Forozan et al, 1997).

CGH is used to detect amplified or deleted chromosome regions in tumors by mapping their locations on normal metaphase chromosomes and has been used to screen for deletions and amplifications in several types of human neoplastic diseases (Angelis et al, 1999).

Figure 1 below shows the principle of CGH. In brief, CGH is a modified *in situ* hybridization which uses differentially labeled test and reference DNA for co-hybridization on normal metaphase chromosomes. Quantitation of test to reference DNA using a digital imager allows gains or losses of test DNA to be seen. Subsequent confirmation of chromosomal location was then done with FISH (Forozan et al, 1997).

Figure 1: Comparative Genome Hybridization technique.



2.2.1. Comparative Genome Hybridization

An overview of CGH studies of selected genital and urological tumors showed chromosome 8q to be most commonly gained in breast, ovarian, prostate, bladder and testicular tumors (Forozan et al, 1997) suggesting that there may be genes which are involved in common pathway for carcinogenesis irregardless of the tissue origin. CGH is also useful in the analysis of the biological basis of tumor progression process in which two cancer specimens from the same patient at different stages of progression can be analyzed. For example in one study, it appeared that CGH showed gain of 1q and 8q in breast cancer, and upon analysis, it was found that 1q appeared early on during tumor progression while 8q was suggested to be associated with subsequent tumor progression (Forozan et al, 1997).

Chromosome 8 has been shown to contain genomic regions which are commonly amplified in a number of cancers as mentioned above. One of the most famous gene, and is also the candidate oncogene, found in this chromosome is c-myc at 8q24 (Garnis et al, 2004). There are also novel regions and genes which are implicated that are distinct from c-myc since c-myc amplification is not always found to be amplified in all cancers *in vivo* (Nupponen et al, 1998). In a recent study, RAD21 and K1AA0196 at 8q24 are found to be amplified and overexpressed in prostate cancer in addition to the common amplification of 8q23-24 in prostate cancer (Porkka et al, 2004). Other note-worthy studies showing 8q gain are the following studies such as CGH of tumor samples from young women ≤ 35 years of age with sporadic

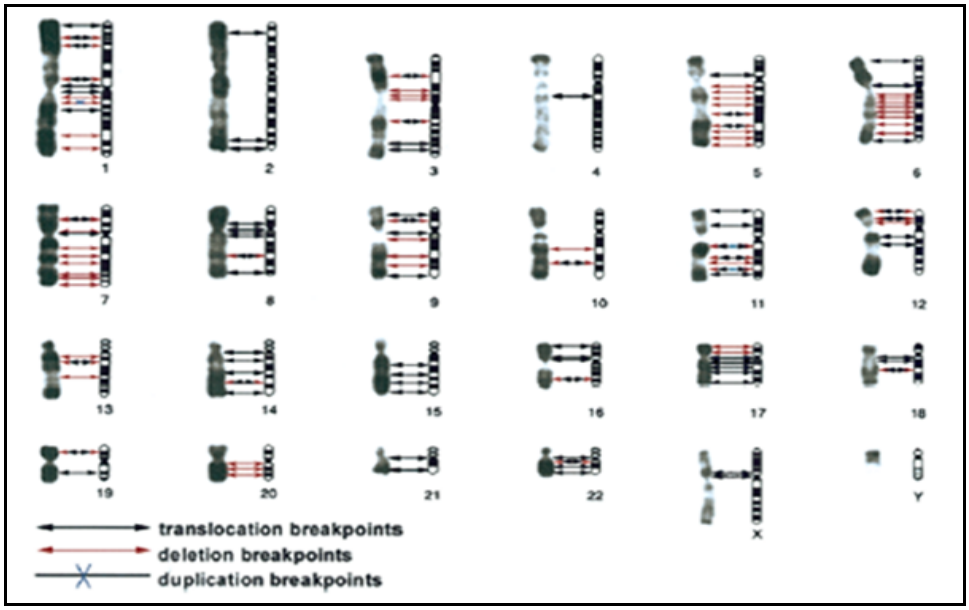
breast cancer revealed genomic gains of 8q in 61.4% of the cases (Weber-Mangal et al, 2003), DNA gains at 8q23.2 serve as a potential early marker in head and neck carcinomas (Da Silva Veiga et al, 2003) and 8q24.12-8q24.13 segment being identified as a common region of over-representation in 10 chronic myeloid leukemia-derived cell lines suggesting that this region could harbor gene (s) driving disease progression (Shigeeda et al, 2003).

2.2.2. Alu repeats and genetic aberrations

With the complete sequence of the human genome, genetic research into database mining for repeat sequences has also intensified. It has been found that more than a third of the human genome consists of repetitive sequences. Almost all of these have arisen by retroposition of an RNA intermediate followed by insertion of the resulting cDNA into the genome. Of these, Alu elements are the most abundant class of interspersed repeats (Smit, 1999). Alu repeats comprise 5 to 10% of the human genome and are shown to hybridize preferentially to reverse bands (R-bands) of metaphase chromosomes (Holmquist, 1992). Cytogenetic studies of tumor cells have shown that recurring chromosomal abnormalities such as translocations, deletions and inversions are present in many tumors. Many of these rearrangements mechanisms proposed are sequence dependent. As shown in figure 2, there is a correlation between chromosomal abnormalities in cancer and presence of Alu repeats. Alu repeats has been shown to increase the recombination frequency between vector DNA and host genome loci (Kato et al, 1986, Wallenburg et al

1987, Kang et al, 1999) suggesting a role as a recombinant "Hot spot" for the Alu element. In addition to this finding, it has been shown that several genetic diseases have both a DNA instability phenotype and a high frequency of carcinogenesis. Several mechanisms leading to genetic instability have been suggested. Examples are mutations in mitotic checkpoint genes and loss of telomere capping function (DePinho and Polyak, 2004). This seems to suggest a correlation between recombination frequency, genetic instability and cancer predisposition (Bishop and Schiestl, 2000 and Hoeijmakers, 2001).

Figure 2: Position of cancer breakpoints of recurrent chromosome aberrations mapped to Alu repeats within R-bands.



Adapted from Kolomietz et al, 2002.

2.3 Cancer research

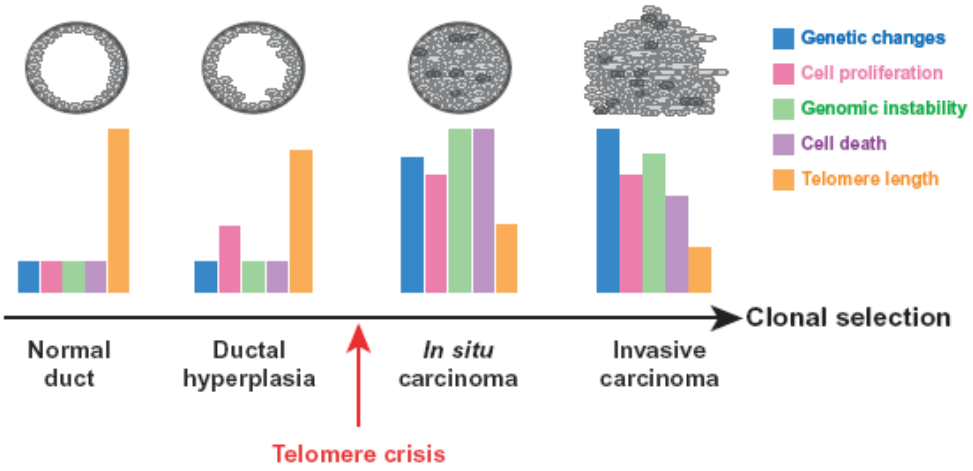
90% of all human malignancies are carcinomas (derived from epithelial cells), many of which are heterogeneous in their biological and clinical behavior thus warranting a greater understanding of their development and progress for better diagnosis and therapy (Alaiya et al, 2000). Correlative studies of genes and clinical outcome allows identification of biomarkers while understanding the mechanism of genetic basis of cancer progression would allow development of new therapies. One example of mechanistic approach in therapy development is the use of small molecule tyrosine kinase inhibitors as anti-tumor and anti-angiogenic agents upon understanding the role of tyrosine kinases in tumor progression (Morin 2000).

2.3.1. Carcinogenesis – changes in the cell

Cancer cells undergo a multistep process which lead to genetic changes that favors deregulated cell cycle, decreased apoptosis, increased invasion and metastasize properties. The initial transformation requires the cell to immortalize. Malignant transformation is generally associated with telomere maintenance which results in cell immortality. Normally, telomeres shorted every cell generation and once they reach a critical length, the cells die. Telomerase is the enzyme responsible for the maintenance of telomere length. Regulation of telomerase activity in turn, is via regulation of transcriptional control of telomerase catalytic subunit gene, human telomerase reverse transcriptase, hTERT. Termination of the life span through a pathway leading

to cellular senescence has been triggered by activation of p53 and/or pRb in response to critically shortened telomere DNA. To propagate indefinitely, the cell has to overcome this senescence checkpoint and activates telomerase. This is known as the telomere-induced crisis. This phenomenon has been shown to implicate in breast tumor progression (DePinho and Polyak, 2004) and oral keratinocytes immortalization (Kang and Park, 2001). Figure 2 shows a schematic model for breast tumor progression where there is a marked increase in genomic instability, genetic changes and cell death after telomere-induced crisis (DePinho and Polyak, 2004).

Figure 3: Changes in cells during carcinogenesis



Adapted from DePinho and Polyak, 2004.

Cellular immortality alone is not enough to warrant tumorigenicity. Environmental cofactors which include chemical carcinogens and viral

infections are needed to effect additional genetic aberrations which results in the ability to create a microenvironment for the transformed cells.

2.3.2. Genes and cancer

Carcinogenesis involves the deregulation of cell cycle. The 2 groups of genes which regulate cell cycle include the proto-oncogenes and tumor suppressors. The former increases proliferation rate while the latter inhibits it. As such, most of which produces short-lived transcript which suggest a tight regulation of their expression (Mercola and Welsh, 2004).

One mechanism proposed for activation of proto-oncogenes is gene amplification. This results in increased gene expression which results in large increase in message and protein levels (Ethier S, 2003). Oncogenes are altered versions of normal proto-oncogenes which regulate normal cell growth and differentiation. Many of these proto-oncogenes proteins are involved in signal transduction processes in the cell (Kang and Park, 2001). One of the most studied genes is the c-myc amplification in a number of cancers. However, there is evidence that suggest that gene amplification alone may not be the only explanation for overexpression of some genes at protein level. Example would be a recent report on TPD52 in which studies suggest that a combination of gene amplification and androgen stimulation likely contributes to the up regulation of the gene (Rubin et al, 2004). This reinforces the argument that there are multiple factors affecting cancer progression which result in their heterogeneous behavior.

Tumor suppressors proteins are frequently involved in the inhibition of cellular proliferation. Deregulation of these would enable extended proliferation. Perturbation of p53 protein function is a common finding in human cancer. p53 is a transcription factor which is phosphorylated following DNA damage. p53 function to arrest cells in G1-phase in order to allow DNA repair and activation of apoptosis to eliminate cells with damaged DNA. One of its target, promoter of p21Cip1, leads to induction of p21Cip1 which leads to a stop in cell cycle progression (Kang and Park, 2001). There is interplay of various genes during cancer progression. One such example is that several transcription factors such as c-Myc and tumor suppressor gene e.g., p53 are able to control hTERT transcription when overexpressed suggesting a tight control of hTERT is needed to prevent deregulation of cell cycle. hTERT deregulation has been shown to implicate in cellular immortalization. Indeed, both oncogenes and tumor suppressors are frequent targets of viral proteins to allow increase proliferation for viral particle propagation. In some cases, integration of viral genome may occur which predisposes the host to genetic aberrations which in turn may lead to viral induced cancers.

2.4 Viral induced cancers

Viruses results in transformed host cells which occur following insertion of viral nucleic acid into host's DNA resulting in mutations in the two classes of human genes mentioned before; the proto-oncogenes and tumor-suppressor

genes. Table 1 shows a few examples of viruses causing different types of cancers.

Table 1: Types of virus-induced cancers

Virus	Types of cancer
HTLV-1	Adult T-cell leukemia
HPV	Squamous cell and genital carcinomas
Hepatitis B and C virus	Hepatocellular carcinoma
Epstein-Barr virus	Burkitt's lymphoma; B cell lymphoma; Nasopharyngeal carcinoma
Human neurotrophic polyomavirus, JCV	Medulloblastoma
Human herpes virus B	Kaposi's sarcoma
Simian Virus 40	Broad range including mammary and salivary glands, pancreas, prostate, liver, lung, kidney, intestine, brain, choroids plexus, lens of the eye, bone, smooth muscle, cartilage and lymphomas

2.5 HPV carcinogenesis

Of the three causative molecular mechanisms of cervical cancer, two are associated with HPV. One is the effect of the viral oncogenes E6 and E7; another would be the integration of the viral DNA into chromosomal regions of tumor phenotype (Ledwaba et al, 2004). HPV are small double-stranded DNA viruses which consist of a circular genome. 90% of all cervical cancers contain

HPV DNA. The genome contain 8 early genes and 2 late genes of which E6 and E7 early genes has been demonstrated to be important for viral-induced carcinomas of genital tract as they cause the inactivation of p53 and retinoblastoma respectively. Inactivation of p53 by E6 has been shown to precede the development of tumors with a fully malignant and invasive phenotype. E7 binding to pRB results in inactivation of pRB which results in hyperproliferation, but not immortalization of cells and tumorigenesis. (Chow et al, 2000; Stoler, 2000; Mantovani and Banks, 2001). E7 also binds and activates cyclin complexes such as p33-cyclin dependent kinase 2, which control progression through the cell cycle. HPV replicates only within the host cell's nucleus, but the mechanism by which HPV transform cells is unclear. Most studies focus on HPV-16 and HPV-18, the viruses most frequently associated with anogenital carcinomas. Variations in carcinogenic potential results from the capacities of E6 and E7 proteins to interact with and alter or destroy key cell cycle regulatory molecules (Bosch et al, 2002). As mentioned earlier, HPV alone is not sufficient to progression to cervical cancer. Additional environmental, behavioral, immunological and genetic factors must be implicated in the pathogenesis and progression of cervical carcinogenesis. Recent evidence that HPV E6 and E7 oncoproteins induce chromosomal abnormalities resulting in genomic instability suggesting downstream effect of infection to carcinogenesis (Duensing and Munger, 2004). Cellular immune surveillance has been shown to be important in the control of HPV infection. Presentation to T cells of target viral peptides is though to influence the host

response and clinical outcome of HPV infection (Beutner KR, and Tyring S. 1997). It has been suggested that the viral persistence, viral load and genetic predisposition are criteria involved in the pathogenesis of cervical cancer development.

Table 2: HPV Gene Products and Their Function

Gene	Function
E1	Initiation of viral DNA replication
E2	Transcriptional regulation/DNA replication
E3	Unknown
E4	Alteration of mitotic signals
E5	Transforming protein, interacts with growth factor receptors
E6	Oncogene, binds to p53, leading to degradation
E7	Oncogene, inhibits pRb, p107 and p130
E8	Unknown
L1	Major capsid protein
L2	Minor capsid protein

Adapted from http://www.mf.uni-lj.si/acta-apa/acta-apa-02-3/derma3-2cl.html

2.5.1. HPV integration into human genome

The HPV genome replicates as an extra chromosomal episome or plasmid in benign HPV-associated lesions. However, the viral DNA is often integrated into the host's chromosome in malignant HPV associated lesions. Viral integration results in lifetime persistence of certain viral genes and

increases the risk of cancer. It has been shown that the viral integration into host genome is the necessary event for the keratinocytes immortality (McGlennen 2000). The 1991 Couturier et al report shows integration of HPV in cellular genomes near *myc* gene in genital cancers. This was found to be the case in most invasive genital carcinomas as compared to intraepithelial neoplasia where HPV DNA is detected most commonly as episomal molecules. This finding suggests a mechanism which may result in alteration of gene structure or overexpression of proto-oncogene.. Recently, human telomerase reverse transcriptase, hTERT, gene has been shown to be another target site for viral integration (Ferber et al, 2003). Viral integration occurs throughout the host genome, leading to the presumption that there are no preferred sites of integration (Ferber et al, 2003b). However studies by Thorland et al in 2000, showing integration into genome to be non-random with HPV 16 integration to frequently occur at common fragile sites suggesting presence of chromosome 'Hot Spots' for viral integration. Ferber et al supported this finding by showing a preferential integration of HPV 18 near the c-myc locus in cervical carcinomas suggesting a nonrandom integration process. Another study by Ferber et al in 2003b involving the hTERT gene cited above supports the hypothesis of a nonrandom integration of viral genome and that the sites of integration may play an important role in carcinogenesis. HPV genomes attached to the host chromatin via E2 protein and replicate at a steady state, once for every cell division. During this process of integration, the viral genome breaks at E1 and E2 regions, never at the E6 and E7 region. The loss of

E2 results in the loss of E6 and E7 regulation. This allows the overexpression of E6 and E7 (zur Hausen, 1996). These proteins would continue to stimulate cells to ignore the DNA damage which have been accumulating and produce clones with extended lifespan. In addition, Pett et al 2004 showed high level of chromosomal instability upon HPV16 integration in cervical keratinocytes.

2.5.2. Chromosome "hotspots" for integration and their implications

As suggested above, the sites of integration may constitute genes involved in the regulation of cell cycle. Recent implication of HPV integration in hTERT genes further support the resulting immortalization of these cell clones indicating a directed integration into chromosomes "hotspots" which are probe to chromosomal breakage. The mechanism on the number of additional independent mutations or deregulation of genes and cell cycle needed for these cells to transform to full malignancy is still under investigation. Similarly, the apparent "recognition" of chromosomal hotspots has not been investigated. Perhaps isolation of genes involved in integration sites would provide a clearer understanding on the process of carcinogenesis following viral infection and integration.

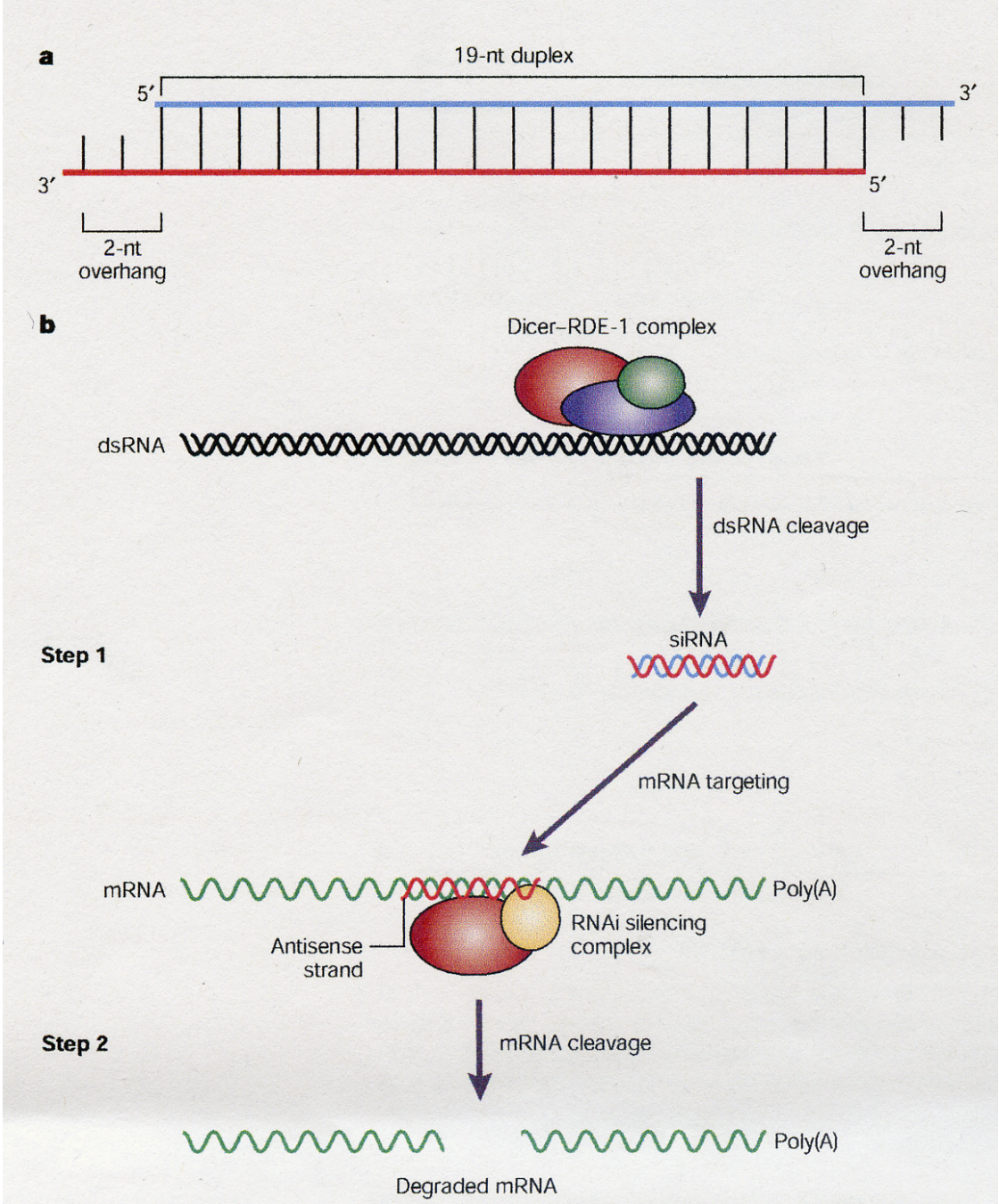
2.6 RNA interference as a tool for cancer research

RNAi is a process in which short 21-mers double-stranded (ds RNA) homologous to a specific gene post-transcriptionally silence or knock down the gene expression (McManu et al, 2002). Figure 3, adapted from McManus et al, 2002 shows how short interfering RNAs (siRNA) which consist of 19bp duplex which is gene specific with 2nt 3' overhang is recognize by a RNA interference silencing complex to guide mRNA cleavage and hence mRNA degradation. The siRNA mimics cellular products of Dicer-RDE-1 complex which processes long dsRNA to siRNA. RNAi has been shown to be able to induce degradation of target mRNA in the cytoplasm and has been suggested to also induce degradation during the process of nuclear mRNA export. True nucleoplasmic mRNAs or pre-mRNAs are however resistant to RNAi degradation (Zeng et al, 2002). The natural function of RNAi and co suppression is thought to be protection of the genome against invasion by mobile genetic elements such as transposons and viruses. It is regarded as a posttranscriptional gene silencing regulatory process (Tuschl T, 2001). RNAi knockout and knockdown experiments in mammalian cells to elucidate gene function have been shown in gene targets such as CD4, CD8, CP110 etc (McManus MT et al, 2002). What determines the success of silencing would be the choice of oligo types, such as plasmid or chemically synthesized (Leirdal et al, 2002), oligo sequence (Brown et al) and the type of transfection reagent (McManus MT et al, 2002). RNAi has also been used in HeLa cells with sequence targeted to the E7 region of bicistronic E6 and E7 mRNA reducing expression of E6 and E7 in HeLa cells

which has been transformed with HPV18. This study shows that RNAi inhibited cellular DNA synthesis and induced morphological and biochemical changes characteristic of cellular senescence rather than apoptosis. The advantage of RNAi specificity without provoking a nonspecific interferon response, suggest a possible therapeutic use for RNAi in HPV related diseases. (Hall and Alexander, 2003). However, the molecular mechanism is still under investigation. Further studies are necessary before the use of RNAi for therapy could be considered. With the association of HPV and cervical carcinoma development, RNAi would serve as a valuable tool in elucidating the possible mechanisms for HPV induced carcinogenesis. The challenge here would be to target genes involved in this process. RNAi libraries which are increasingly available would enable investigations in this aspect.

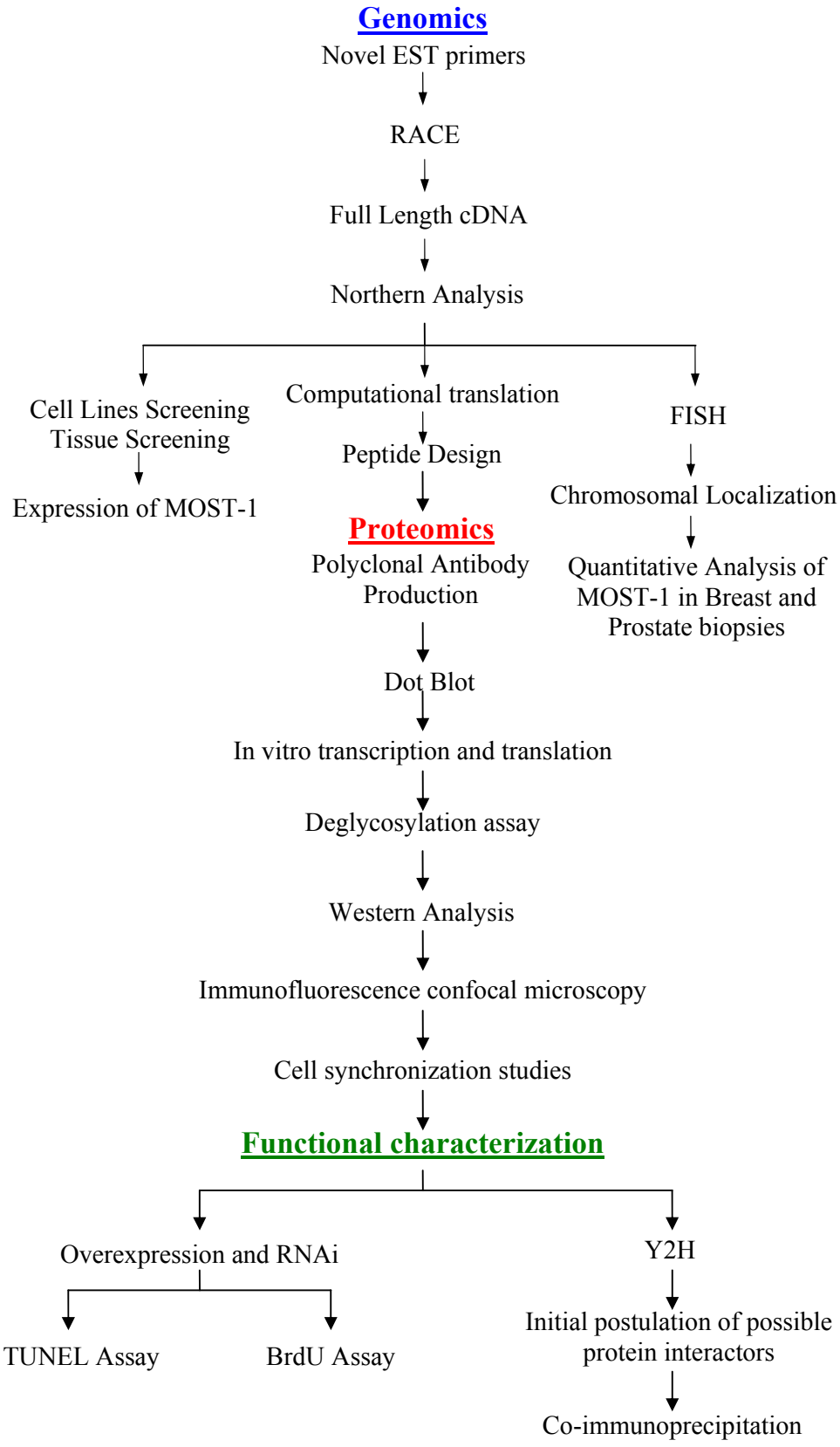
The limitation of this technique would be the non-specific antiviral interferon (IFN) pathway activation which affects cellular behavior independent of targeted gene. In addition, it has been reported that RNAi can also induce non-specific effects on untargeted protein levels independent of the IFN pathway (Scacheri et al, 2004). Hence careful interpretation and necessary controls are needed for conflicting results generated via this technique.

Figure 4: RNA interference mechanism



CHAPTER 3. MATERIALS AND METHODS

Figure 5: Flow chart of gene characterization



3.1. Mammalian cell tissue culture

Cells were thawed in 37ºC water bath by gentle agitation. Vial was decontaminated by dipping in 70% ethanol before transferring vial contents to centrifuge tube with 10ml media. Cells were pelleted at 125 g for 5 min. Supernatant was discarded and resuspended in 5 ml media and transferred to 25 cm^3 tissue culture flask.

For monolayer, culture media were removed and discarded. Cells were then rinsed with 1X PBS pH 7.4 (Appendix 1) before adding 0.5 ml trypsin. Trypsin was removed when cells started to lift off and appropriate media was added to cells and transferred to a new flask. Suspension cells were collected at 125 g for 5 min before rinsing with 1X PBS pH 7.4. Appropriate media was then added to cells and transferred to a new flask.

Cells must be 90%-100% confluent and not more than 2 days old before cell collection. Monolayer cells were scrapped followed by centrifugation while suspension cells were directly collected by centrifugation. Supernatant was discarded and resupended in 1.5 ml of appropriate freezing media (growth media with 10% DMSO) before being transferred to a 2 ml cryovial.

TABLE 3: List of cells with respective growth media used

Cell Line	Type*	Media
1. MOLT-4 T-lymphoblastic leukemia	S	RPMI
2. HeLa cervical carcinoma	M	DMEM
3. PLC/PRF/5 hepatocellular carcinoma	M	DMEM
4. MRC-5 fetal lung	M	MEM
5. MCF7 breast adenocarcinoma	M	DMEM
6. Prostate Normal	M	PrEGM
7. Mammary Normal	M	MEGM
8. DU-145 Prostate carcinoma	M	MEM

* S denotes suspension cell line while M denotes monolayer.

3.2. Gene isolation

Initial genomic gene isolation approaches such as positional cloning involved the comparative use of physical maps, which are mainly with genetic markers of known distances in base pairs apart, with linkage map, which is made up of distances measured by recombinant frequency (Lewin, 1997). Other methods include the detection of the presence of CpG islands, which are believed to mark 5' end of transcriptionally active DNA sequences (Larsen et al, 1992), cDNA selection of YAC clones library to obtain genomic region of interest (Lovett, 1994) and "inverse genetics" whereby the protein of interest has been purified, but the gene that encodes that protein has not been identified. Thus using the protein sequence, one can then synthesized all possible coding

sequence and use Southern blot for gene isolation (Robert, 2003). However, these methods are laborious and time consuming.

For ease of gene identification, ESTs provide the answer for rapid gene scanning amidst the huge genome information. ESTs are expressed sequence of a gene. ESTs otherwise known as cDNA fragments are also useful for the annotation of physical maps, and function as handles for the classification of multigene families (Adams et al, 1991; Sim and Chow, 1999). They are isolated via use of mRNAs reversed transcribed into cDNAs. ESTs have the advantage of stability over mRNAs and they give an indication on the gene expression level and nature of the transcript such as splice variants (Tan et al, 2003). Once an EST is isolated, nucleic acid hybridizations such as northern hybridization or RNase protection assay are used to detect the specific mRNA. But the sensitivity may not be sufficient to detect low levels of the expression of certain genes. Hence the use of RT-PCR is frequently used as a replacement. As more and more genes are isolated and studied, the usual belief that mammalian genes contain introns is also slowly changing. The only drawback of RT-PCR is the presence of false positive results due to possible contamination from genomic DNA. The problem of genomic DNA contamination becomes a big issue for intronless gene. How does one differentiate intronless genes with genomic DNA? To overcome the problem of false positive and to detect gene expression of an intronless gene, a highly sensitive detection system known as NASBA can be used. NASBA is an isothermal nucleic acid amplification reaction which uses a specific primer pair one of which contains a T7 promoter, AMV reverse transcriptase, *E.coli*.

RNase H and T7 RNA polymerase. NASBA achieve selectivity for ssRNA by use of initial denaturation of nucleic at 65°C which prevents denaturation of dsDNA followed by reverse-transcription at 41°C (Heim et al, 1998). Another method known as rapid amplification of cDNA ends (RACE) in which mature mRNAs are used as priming sites for elucidation of full-length cDNA.

3.2.1. Genomic DNA Isolation

Cells were grown to confluency with their respective growth media before being scrapped and harvested by centrifugation at 1500 rpm for 5 min at 4°C. Cells were then washed thrice with 1X PBS before incubation with 4 ml of SSE (Sodium acetate, SDS, EDTA) extraction buffer (Appendix 2) for 5 min. After which 1 volume (V) of SSE equilibrated phenol (Appendix 2) was added and shake gently for 5 min. Mixture was then spun at 3500 rpm for 5 min at 4°C. The upper phase was recovered and added to 1 V of chloroform/isoamyl alcohol (25:1) and shake gently for 5 min. Mixture was then spun at 3500 rpm for 5 min at 4°C. The upper phase was recovered and DNA was precipitated overnight with 2 V absolute ethanol and 1 V of 4 M LiCl at -20°C before centrifugation at 14 000 rpm for 20 min at 4°C. Supernatant was discarded and pellet was washed with 1 ml 70% ethanol prior to drying before resuspension in 100 ul 1X TE (Appendix 2).

3.2.2. Total mRNA Preparation

Cells were grown to confluency with their respective growth media before being scrapped and harvested by centrifugation at 1500 rpm for 5 min at

4°C. Extraction of total cytoplasmic RNA was carried out with modifications using the method described by Gough (1988). Briefly, cells were washed thrice with 1X PBS and resuspended in 200 ul RNA extraction buffer A (Appendix 2). After which, vortexing and spinning at 1000 rpm for 5 min at 4°C was carried out to separate cytoplasmic lysate from nuclear pellet. Cytoplasmic lysate was transferred to a new tube with 200 ul RNA extraction buffer B (Appendix 2). After vigorous vortexing, 200 ul of TE-equilibrated phenol followed by 200 ul of chloroform/isoamyl alcohol (49:1) was added. Tube was vortex then spun at 14,000 rpm for 2 min at 4°C. Supernatant was then transferred to new tube and total cytoplasmic RNA precipitated using 1 V 100% ethanol and incubated 1 hour at -20°C before centrifugation at 14,000 rpm for 20 min at 4°C. Pellet was washed with 1 ml 70% ethanol prior to drying before resuspension in 100 ul double distilled H_2O (treated with diethyl pyrocarbonate, DEPC). Spectrophotometric readings were taken at 260 nm and 280 nm using a Shimadzu UV160 spectrophotometer. The concentration of RNA was subsequently standardized to 100 ng/ul. RNAsin (Gibco #15518-012) was added to a final concentration of 1 U/ul. The RNA was stored at -80°C until further use.

3.3. Primers location and use

Primers were synthesized by either GENSET Singapore Biotech. Pte Ltd, Operon Technologies Inc., or 1st Base. All primers were made to 100 uM and the working solution for PCR and cycle sequencing reactions were 10 uM and 1 uM respectively.

Table 4: List of primers and their respective cDNA positions

Primer	Sequence	Position	Use of primers
TR4.1	5’ ACT CCT GAG CTC AGG CAA TC 3’	1368-1387	Intron/Exon mapping, expression studies, RACE library screening
TR4.2	5’ ATT TCT ACC CCA GGA GTC AGG TC3’	1277-1299	Intron/Exon mapping, expression studies (nested)
TR1	5’ CAT TCT CCA TGG TGG CTA CCC TG 3’	1102-1125	RACE, Intron/Exon mapping
TR2	5’ CTG TGG CCA AGC CTG ACA TCC AC 3’	1136-1157	RACE, Intron/Exon mapping
TR3	5’ TCT GAC AGG CCA TCG TCA GCT GC 3’	714- 736	Sequencing of full length cDNA
TR4	5’ TCT GAT CGA GGG CTC AAG AC 3’	2350-2369	Sequencing of full length cDNA
TR5	5’ TCT CCT CAA GGC AGC TGG AG 3’	1968-1987	Sequencing of full length cDNA
TR6	5’ AGA GGC TCA GGT CCA AAG C 3’	394- 412	Sequencing of full length cDNA
TR7	5’ AAA ACG CGG TAC GGA ACT TCT C 3’	1806-1827	Sequencing of full length cDNA
TR8	5’ TTT ACC GGA GCT CAC ACG GCT AC 3’	2648-2670	Generation of FISH probe
TR9	5’ TGA AAC ACA CCC TTC TCG 3’	2546-2563	Generation of Northern probe
TF1	5’ AGC CAC CAT GGG AGA ATG TC 3’	1108-1127	RACE, Intron/Exon mapping, RACE library screening
TF2	5’ TGG ATG TCA GGC TTG GCC ACA GG 3’	1136-1158	RACE, Intron/Exon mapping
TF3	5’ ACA TGC CTG TAG TCC CAG CTA CG 3’	1462-1480	Sequencing of full length cDNA
TF4	5’ TCC ACC TTA CGT AGC CGT GTG AG 3’	2638-2660	Sequencing of full length cDNA
TF5	5’ TTG TCC AGC ATT ACA AGA GCA CC 3’	860-882	Intron/Exon mapping, expression studies
TF6	5’ TCT TGA GCC CTC GAT CAG AG 3’	2351-2370	Sequencing of full length cDNA
TF7	5’ TAC AGG GAG GAA CAA GGA GAA AC 3’	1000-1022	Intron/Exon mapping, expression studies (nested)
TF8	5’ TCC GCT GTC TAC CCA TG 3’	1901-1917	Sequencing of full length cDNA
TF9	5’ TGC AAC CCA AAC TCG AC 3’	2462-2478	Sequencing of full length cDNA
TF10	5’ ACC ATA GAC AAC TGC CAC TCA CTG 3’	34- 57	Generation of FISH probe
TF11	5’ ACA TCC CCA TCT CAC AG 3’	124- 140	Generation of Northern probe
TF12	5’ AAG TCT GAG GCT CTG 3’	565- 579	Sequencing of full length cDNA

3.4. Rapid amplification of cDNA ends (RACE)

Marathon-Ready MOLT-4 and human fetal lung (Clontech) cDNA libraries ligated with adaptor oligonucleotides were used as templates for RACE reactions. Based on the original MOST-1 EST, gene-specific primers TF1 (nt 1108-1127) and TR4.2 (nt 1299-1277) were designed. Before RACE, PCR was performed with these primers to confirm that the cDNA libraries contained the cDNA of interest. Primers TR4.2 and TF1 were individually paired with the adaptor primer AP1 in initial RACE experiments via long-distance PCR using Advantage 2 polymerase (Clontech). Nested primers TR1 (nt 1125-1102) and TF2 (nt 1136-1158) were then paired with the adaptor primer AP2 in subsequent RACE to obtain the 5′ and 3′ RACE fragments, respectively. Thermal cycling conditions were 94°C for 30 sec, followed by 25 cycles each of 94°C for 5 sec and 68°C for 4 min. The RACE products were cycle-sequenced as described below.

3.5. Cycle Sequencing

ABI PRISM BigDye terminator cycle sequencing ready reaction kit was used with 30-90 ng DNA and 3.5 uM primer. Reaction mix was topped to 20 ul with water. Thermal cycling conditions were 25 cycles each of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min. After which, ethanol precipitation was carried out. Contents were dried and sent for sequence analysis with ABI PRISM 377 DNA sequencer (PE Applied Biosystems).

3.6. Bioinformatics Analysis of MOST-1 gene

Bioinformatics provide prediction softwares and faster database servers in user friendly-environment. Various programs for different purpose have been produced for different studies such as genomic analysis, computational translation and putative protein analysis. These have accelerated information sharing and technology transfer. Listed below are several programs used in this study.

Table 5: Computation programs for gene structure analysis

Program	**Purpose**
Blast www.ncbi.nlm.nih.gov	Sequence alignment for DNA and protein for human and model organisms
Bioportal www.bic.nus.edu.sg	Sequence analysis program for DNA analysis and protein analysis
PubMed www.ncbi.nlm.nih.gov	Latest update on periodicals even before press
ProtParam www.expasy.ch	Analysis of computed/experimental protein sequence for prediction such as protein solubility, stability and pI
Predict protein http://maple.bioc.columbia.edu/predictprotein/	Motif and domain search, globularity prediction, protein structure prediction
SUMOplot prediction http://www.abgent.com/sum	Attachment sites of SUMO protein (11kDa)

oplot.html	
NetOGlyc 2.0 Server http://www.cbs.dtu.dk/services/NetOGlyc/	Mucin type GalNAc O-glycosylation sites in mammalian proteins.
DNASIS and DnaSTAR	Software analysis packages for sequence analysis, protein analysis, sequence alignment and chromatogram overlay

Nucleotide analysis was carried out using GCG analysis (bioportal) and BLAST program while putative amino acid sequence analysis and motif searches were performed using bioinformatics softwares such as ProtParam, Predictprotein, SUMOplot and NetOGlyc 2.0 Server. These programs allows the prediction of post-translational modification. The type of modification allows one to further understand the role of the protein in the cells. Post-translational modifications include the following:

- proteolytic cleavage of inactive proproteins, eg pancreatic enzymes, and signal peptide from preproteins
- glycoproteins, in which oligosaccharides such as glucose, galactose, fructose, GalNAc, GlcNAc and NANA are covalently attached to proteins, either via O-glycosidic or N-glycosidic bonds. N-glycosidic linkage being the predominant attachment to glycoproteins of mammalian through the amide group of asparagines within a consensus sequence N-X-S (T) where X is any amino acid except praline. O-glycosidic linkage is to the hydroxyl of serine, threonine or

hydroxylysine (generally only found in collagens). Most secretory or plasma membrane bound proteins are modified by carbohydrate attachment and it is normally the extracellular portion of the protein that is modified. Attachment of carbohydrate to intracellular proteins is rare but confers unique functional activities and occurs via O-linkage and involves the attachment of GlcNAc to serine or threonine residues. Several transcription factors and RNA polymerase II have been shown to be modified by O-GlcNAc linkage,

- lysosomal targeting of enzymes are directed by a specific carbohydrate modification in which GlcNAc-1-phosphate is added to the carbon-6 hydroxyl group of one or more specific mannose residues that have been added to these enzymes. Removal of GlcNAc leaves mannose residues phosphorylated as Man-6-P which is then able to bind to a specific Man-6-P receptor that is present in the membranes of the Golgi apparatus which upon binding will target proteins to the lysozyme,
- acylation occurs when N-termini methionine is hydrolyzed and an acetyl group is added. Some protein have 14 C myristoyl group added which allows association with membranes,
- methylation occurs at lysine residues of some proteins for eg, calmodulin,
- phosphorylation is one of the most common protein modification that occurs in animal cells and is normally regulatory in nature. As such, these modification is transient and occurs at serine, threonine and tyrosine residues ratio being 1000:100:1 respectively,

- sulfation is the permanent modification of proteins at tyrosine residues that is normally necessary for biological activity,
- prenylation is the addition of 15 C farnesyl group to cysteine residues at C-terminus of protein in a thioether linkage (C-S-C) at consensus sequence of CAAX where C is cysteine, A is any aliphatic amino acid except alanine and X is the C-terminal amino acid. Numerous G-proteins of signal transduction cascades have γ-subunits with this modification,
- vitamin C and K dependent modifications occur when enzymes carrying out the modification require either as a cofactor. Proline and lysine hydroxylation and carboxy terminal amidation requires vitamin C as cofactor while carboxylation of glutamine residues requires vitamin K. (http://www.indstate.edu/thcme/mwking/proteinmodifications.html)

Once the protein modification is predicted, experiments can then be designed to verify the prediction. As more proteins are found to have different active states, function and localization depending on their post-translational modifications, there are also systems to study some of the more common modifications such as

- signal peptide cleavage which can be observed as a shift to a lower molecular weight upon SDS-PAGE analysis when the template is expressed in rabbit reticulocyte lysate in the presence of canine pancreatic microsomal membranes (Walter et al, 1983),
- glycosylation can be observed with the use of enzymatic deglycosylation assays or antibody detection of oligosaccharides,

- acetylation inhibition with the addition of S-acetonyl CoA (Rubenstein. 1981) and 3 mM EDTA to block phosphorylation (Nakamura. 1993).

3.7. Organization of MOST-1 gene

RNA is first modified in the nucleus after transcription by having a cap added on the 5' end and a poly-A tail at the 3' end. Splicing then occurs after which the now matured mRNA is transported through the nuclear pore to the cytoplasm where it is translated. Splicing occurs with introns having a GU-AG rule flanking its ends (Lewin, 1997). To determine the presence of exon or intron, primer pairs flanking 200-400 bp target fragments were designed based on the full length MOST-1 cDNA sequence derived from RACE. With these MOST-1 primer pairs, PCR products amplified from MOLT-4 genomic DNA and cDNA were sequenced, compared and subjected to Grail computational analysis. (http://compbio.ornl.gov/Grail-1.3).

3.8. Chromosomal Localization of MOST-1 gene

After obtaining the full length cDNA, chromosomal localization should be done to further understand the gene function. With the completion of the human genome project, chromosome mapping allows us to have an idea if the gene is present at a particular disease locus. Chromosomal position of MOST-1 was mapped by fluorescence in situ hybridization (FISH), using a genomic probe of ~2.6 kb generated from MOLT-4 genomic DNA and primers TF10 (nt 34-57) and TR8 (nt 2670-2648). The probe was then sent to SeeDNA (www.seedna.com) for hybridization. Chromosomal slides were prepared using

cultured lymphocytes isolated from human blood. Cells were synchronized and harvested. Cells were then mount onto slides using standard procedures of hypotonic treatment, fixation and air-dry. DNA probe was biotinylated with dATP and procedure for FISH detection were performed according to Heng et al., 1992 and Heng et Tsui, 1993. Assignment of FISH mapping data with chromosomal bands was achieved by superimposing FISH signals with DAPI banded chromosomes. The MOST-1 sequence was also subjected to computational analysis with the UniGene program (http://www.ncbi.nlm.nih.gov/UniGene)

3.9. MOST-1 Expression

Since the chromosomal location of MOST-1 is mapped to 8q 24.2, a locus gained in many cancers, expression studies such as screening of cells lines, tissues and biopsies, was done to further correlate the gene with the disease. Multiple tissue cDNA panels generated from poly(A)$^+$ RNA of 16 human adult tissues (0.8 ng each) (Clontech), and total mRNA (200 ng each) extraction were reverse transcribed to cDNAs from 19 human cancer and 2 normal cell lines extraction as described above, were screened for *MOST-1* expression with PCR using primers TF5 (nt 860-882) and TR4.2. This primer pair flank a 440 bp fragment was performed using a profile of 95°C for 1 min, followed by 25 cycles each of 95°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec. A subsequent amplification was repeated using a 50× dilution of the first PCR product as template. To serve as controls for the human adult tissues and cell lines, housekeeping genes were also subjected to PCR using

glyceraldehyde-3-phosphate dehydrogenase (*G3PDH*)-specific primers (Abel-Mageed et al., 1997), and/or *HUEL* gene-specific primers 5R2 (5′-AAGTATGTAATGGAAAGTCGTG-3′) and R3 (5′-AAGCATAAAGGTCTCTAGTTCTTCAGG-3′) (Sim and Chow, 1999). All reactions were performed in duplicates.

3.10 Northern Blot analysis

To verify the full length of gene and screening of tissues DNA, Northern analysis was carried out. Probe was generated by PCR and purified with Qiagen gel extraction kit. The DNA was then labeled with Redivue γ^{32}P dCTP (Amersham pharmacia biotech) with Megaprime^{TM} DNA labeling system (Amersham pharmacia biotech). Briefly, 25 ng of DNA was incubated with 50 uM primers and heat at 95°C for 5 mins before spun down and cooled to room temperature. 10 ul of labeling buffer, 2 ul of enzyme and water was added to a final volume of 45 ul. 5 ul of Redivue γ^{32}P dCTP was then added and incubated at 37°C for 60 mins. 5 ul of 0.2 M EDTA was then added to stop the reaction prior to purification via ChromaSpin columns (Clontech). Activity should be in the range of 1-2 x 10^7 cpm/ul. MTN blot (Clontech) was then prehybridize in ExpressHyb solution at 37°C for 30 mins. Probe was then added and incubated with the blot at 60°C for 60 mins. Blot was then washed with wash solution 1 (See Appendix 2) thrice with 10 min incubation at room temp each and then with wash solution 2 (See Appendix 2) thrice with 10 min incubation at 50°C. Blot is then expose to X-ray film O/N in -70°C.

3.11 Semi-quantitative RT-PCR analysis

Since chromosomal location mapped MOST-1 to 8q24.2, a region found by comparative genomics to be amplified in ~70% of prostate and 50% of breast cancer tissues, a screen of breast and prostate biopsies was carried out. Total RNAs, previously isolated via needle biopsy and graded according to the Nottingham modified WHO criteria, from 27 pairs of snap-frozen invasive ductal breast cancer and adjacent normal tissue samples were used (Jin et al., 2000, 2001, 2002) (Appendix 5). After reverse transcription, 80ng of cDNA sample was subjected to PCR amplification using primers TF5 and TR4.2 at 95°C for 30 sec, followed by 35 cycles each of 95°C for 30 sec, 65°C for 30 sec and 72°C for 1 min. Semi-quantitative analysis of *MOST-1* mRNA was facilitated by concurrent RT-PCR of the *G3PDH* housekeeping gene (Abel-Mageed et al, 1997). A prior PCR cycle optimization was conducted to ensure that all reactions remained in a linear region (by removal and analysis of samples at 15, 20, 25, 30, 35 and 38 cycles). The resultant RT-PCR products were electrophoresed along with DNA markers in agarose gels stained with ethidium bromide, visualized under UV, and analyzed by densitometric scanning using a GS-700 Imaging Densitometer and Quantity One software (Bio-Rad, Hercules, CA). The level of *MOST-1* expression was quantified by calculating the ratio of the intensity of each RT-PCR product of *MOST-1* to that of *G3PDH* (Jin et al, 2000, 2001, 2002). Relative *MOST-1* expression was further presented as the differential expression between corresponding tumor (T) and normal (N) tissues as follows: overexpression in tumor tissue compared with matched normal tissue (T>N); decreased expression in tumor tissue

compared with matched normal tissue (T<N); and no difference (T=N) (Xu et al, 2000). All reactions were performed in triplicate or duplicate, and mean values derived.

3.12 Real time PCR analysis

DNA were extracted from randomly selected, paraffin-embedded tissue sections of normal, hyperplastic and malignant prostatic biopsies (Appendix 5) using buffer comprising 0.5 M Tris, 10 mM NaCl, 20 mM EDTA, 1% SDS (pH 9.0) and proteinase K. This was accompanied by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. DNA samples were reconstituted to the same concentration of 100 ng/μl, and 100 ng of each sample was used for amplification. Real-time PCR was performed with primers TF1 and TR4.2 using the iCycler iQ SG-1 quantitative PCR reaction buffer kit (Bio-Rad) and LightCycler - DNA Master SYBR Green I (Roche Diagnostics), at 95°C for 4 min followed by 40 cycles each of 95°C for 30 sec, 65°C for 30 sec and 72°C for 1 min. Also included was the amplification of the *G3PDH* housekeeping gene (Abdel-Mageed et al, 1997). All reactions were performed in triplicate or duplicate, and mean values derived. Real-time PCR conditions were optimized by varying $MgCl_2$ concentrations between 1.5 mM and 5 mM. Different starting amounts of template were also used for validation of amplification efficiency. For each sample, the threshold cycle (C_T) values for both *MOST-1* and *G3PDH* were determined, and the difference between their C_T values (ΔC_T) was calculated as recommended by the manufacturer of the real-time

PCR kit (Bio-Rad). The mean ΔC_T values of each category of normal, hyperplastic, low, intermediate and high grade carcinoma samples were calculated using SPSS for Windows (release 10.0.5). These mean ΔC_T values were then subjected to analysis by multiple comparisons, and the p values of significance between their differences were computed. Acting as the reference, the mean ΔC_T of normal tissue samples was compared against the mean ΔC_T values of hyperplastic samples and carcinomas of varying grades in order to obtain the relative DNA levels presented as a numerical-fold change for each category of prostatic neoplasia (Xu et al, 2000).

3.13 Raising of polyclonal antibody

3.13.1 Design of synthetic peptide

Rabbit polyclonal anti-MOST-1 was made by using synthetic peptide Ac- TGPEQSDICHTGSEAR-NH2 (aa 19-34), corresponding to the predicted hydrophilic and antigenic regions of MOST-1, conjugated to Diphtheria toxoid (Mimotopes Pty. Ltd, Australia, http://www.mimotopes.com).

3.13.2 Generation of antibody

Polyclonal antisera were raised in three rabbits by immunizing with synthetic peptides. Peptide conjugate were mixed with Freud's incomplete adjuvant 1:1 ratio until a thick, white emulsion. The peptide conjugate was injected at week 3, 5 and 8 after acclimation (Week 1). Rabbits were bled at weeks 2 (pre-immune), 7 and 10 (hyper-immune bleed). Bleeds were then incubated at 37°C for 1 hour followed by 4°C O/N. Serum was drawn and residual red blood cells were spun down at 200 g for 5 min. Serum was then

subjected to dot blot assays. Antibody characterization was done using dot blot analysis to select for positive serum before affinity purification.

Batchwise affinity purification procedure was carried out according to manufacturer's recommendation. Briefly, specific peptide covalently coupled to thiopropyl-sepharose 6B gel syringe column was washed 5 times with 8 ml PBS pH 7.4 each time. 5 ml of serum followed by 3 ml of PBS were drawn into the syringe before sealing nozzle. Syringe was mixed O/N at 4°C. Solution was then expelled from the syringe and washed with PBS 5 times as above followed by 3 washes with saline. 7 ml of 0.1 M glycine/HCl buffer pH 2.5 were mixed with gel slurry for 1 min before expulsion from syringe. Elution was repeated 2 more times followed by 2 washes with saline. Procedure was repeated with 0.1 M glycine/NaOH buffer pH 11.5. All solutions were pulled and pH of antibody solution was adjusted to 7.0 using 5% phosphoric acid. Antibody solution was then concentrated using AMICON Ultrafiltration apparatus model M3 to 5 ml. Antibodies was then aliquot and stored at -20°C.

3.13.3 Dot Blot analysis

Conjugated and unconjugated peptides (500 mM) were dot on nitrocellulose membrane 5 times each with 1ul of peptide mixture. Membrane was then block O/N at 4°C with blocking solution (Appendix 3) before wash with TBST (Appendix 3) 10 min twice. The collected serum was then added in 1:1000 dilution and incubated at room temperature for 1h before washed twice with TBST buffer, followed by incubation with biotinylated goat anti-rabbit IgG (diluted 1:3000 with TBST) (Biorad) 1 hour at room temperature, washed

twice with TBST buffer and reacted with streptavidin conjugated with alkaline phosphatase (diluted 1:10, 000 with alkaline phosphatase buffer) (Appendix 3). Immunoreactive dots were visualized by color development with NBT/BCIP substrates (Appendix 3).

3.14 Polyclonal antibody verification

For verification of affinity purified antibody, Western analysis was done on in vitro translated recombinant MOST-1 protein.

3.14.1 In vitro translation

MOST-1 was cloned into p-Bluescript SK II vector (Stratagene) and used in TNT® T7 quick coupled transcription/translation system (Promega). In vitro TNT experiments were carried out with and without Redivue™ L-[S^{35}] labeled methionine (Amersham Pharmacia Biotech). Briefly, TNT mix with 1 ug plasmid containing MOST-1 insert and 1 mM labeled or unlabelled methionine was incubated at 30°C for 90 min. 5 ul of the translated products were then ran on a 10% SDS-PAGE gel. Western detection as described below was done on unlabeled reaction mix while labeled reaction mix lane was exposed to X-ray film O/N and developed.

3.14.2 Differential treatment for aggregates

As antibody only detects high molecular weight species of recombinant MOST-1, differential treatment of MOST-1 aggregates was done. TNT experiment with unlabelled methionine was carried out as above and products

were subjected to either 95°C for 5min, incubation with 2% SDS or both treatments. 5 ul of products or untreated rabbit lysate only were then electrophoresed in 10% SDS-PAGE gel before subjected to western analysis described below. No sample buffer was used to maintain non-reducing conditions.

3.15 Protein Characterization

Protein can be membrane associated or free in cytosol in the cells. Proteins synthesized in the cytosol can remain there and function as individual catalytic centers, or as macromolecular structures constructed from cytosolic proteins, or as nuclear proteins in which after synthesis in cytosol they are transported through the nuclear envelope, or in cytoplasmic organelles such as mitochondria. The cytoplasm also contains the reticuloendothelial system which includes the endoplasmic Golgi apparatus, endosomes and lysosomes. Proteins that localize to these compartments are inserted into the ER and transported to their respective location via the transport system of Golgi apparatus. Secretory proteins are transported to the plasma membrane via the reticuloendothelial system before passing into the exterior of the cell. Short signal motifs are present on the proteins for protein trafficking. A summary of the signal types are illustrated in Table 6.

Table 6: Cell signaling motifs

Organelle	Signal location	Type	Length
Mitochondria	N-terminal	Amphipathic helix	12-30
Nucleus	Internal	Basic or bipartite	7-9
Peroxisome	C-terminal	SKL	3

(Adapted from Genes VI)

Chaperones also play a role in protein localization as they maintain protein in an unfolded state before membrane passage so that it will fit into the available channel. Once in the organelles, a hierarchy of sequences will determine the location within the organelle (Lewin, 1997). Protein location also gives a hint of the possible changes that occur during functional vs. non-functional state or disease and non-disease state. Such is in the case of *src* kinases, which have to be myristolyated and membrane bound before being able to induce cellular transformation. An example of this phenomenon is the study on *v-fgr* oncogene in which its subcellular localization is a factor for its ability to induce transformation (Baker et al, 1998).

3.15.1 Total protein extraction

Total protein extraction was prepared by culturing cells to 90% confluence before washing trice with cold PBS (pH 7.4). Lysis buffer I (Appendix 3) with complete protease inhibitor cocktail (Roche) was added and incubated on ice for 30 min with occasional mixing before centrifugation at 14,000 rpm, 15 min, 4°C. The supernatant was harvested.

3.15.2 Fractionated protein extraction

Subcellular fractions of cultured cells were prepared as previously described with modifications (4). Cells were cultured to 90% confluence before washing trice with cold PBS. Lysis buffer II (Appendix 3) with a complete protease inhibitor cocktail (Roche) and incubated on ice for 15 min with occasional mixing before centrifugation at 1000 rpm, 5 min, 4°C. The pellet representing the nuclear fraction was processed with Lysis buffer I as above. The supernatant was harvested and ultracentrifuged at 100,000 g, 1 hour, 4°C. The resulting pellet representing the microsomal fraction was processed in the same way as the nuclear pellet. The supernatant containing the cytosol fraction was brought to final concentration of 1% sodium deoxycholate and 1% Triton X-100.

3.15.3 Western Blot Analysis

Fifteen ug of lysate obtained above was loaded per lane and subjected to 12% polyacrylamide-SDS gel electrophoresis in Laemmli running buffer (Laemmli, 1970) and electroblotted onto polyvinylidene difluoride (PVDF) (Roche) membrane in CAPS transfer buffer. Membranes were blocked for 1 hour at room temperature and washed twice with TBST buffer. Each blot was incubated with primary antibody (diluted 1:1000 with TBST) (Biorad) O/N 4°C, washed twice with TBST buffer, followed by incubation with biotinylated goat anti-rabbit IgG (diluted 1:3000 with TBST) (Biorad) 1 hour at room temperature, washed twice with TBST buffer and reacted with streptavidin conjugated with alkaline phosphatase (diluted 1:10, 000 with alkaline

phosphatase buffer) (Biorad). Immunoreactive bands were visualized by color development with NBT/BCIP substrates. Benchmark™ protein ladder (Invitrogen Life Technologies) was concurrently run, transferred, cut, and stained with Coomassie Blue. Immunoreactive molecular masses were then estimated from logarithms graph obtained.

3.15.4 Indirect Immunofluorescence

Cells were cultured on 13 mm glass coverslips in 24-well microplates until 70% to 80% confluency, washed trice with cold PBS, fixed with 3% paraformaldehyde (PFA) for 10 min at room temperature, washed with PBS, permeabilize with 0.1% Triton-X 100 for 1 min, washed twice with PBS before quenching aldehyde groups with 5 mM NH_4Cl in PBS. Cells were then blocked with 0.1% FCS in PBS for 30 min at room temperature before incubation with primary antibody (diluted 1:100 with blocking solution) O/N at 4°C in a humidified chamber. Cells were then washed with 3 changes of blocking solution over 30 min before incubation with FITC-conjugated goat anti-rabbit IgG antibody for 1 hour at room temperature (diluted 1:10 with blocking solution) (Pharmingen). After which, 3 washes with blocking solution and 1 wash with PBS were done before mounting coverslips on ethanol-cleaned slides and viewed using a Zeiss LSM 510 laser scanning conformal inverted microscope and a Zeiss fluorescence light microscope.

3.16 Cloning

3.16.1. Preparation of competent cells

Chemical competent cells were prepared by picking colony to grow in 5 ml LB O/N at 230 rpm at 37°C. After which, 1:100 dilution was incubated at 230 rpm at 37°C until OD_{600} = 0.5 to 0.6. Culture was then left on ice for 20 min, spun down at 2 500 rpm for 15 min and resuspended in ½ volume cold, sterile 50 mM $CaCl_2$ before incubation on ice for 20 min. Cells were spun down again and resuspended in 1/10 volume of 20% glycerol in 50 mM $CaCl_2$. Cells were then aliquoted and stored at -70°C.

3.16.2. Transformation

Transformation was carried out by incubating 2 ul of ligation mix with 40 ul of competent cells prepared above. Incubation on ice for 20 mins before heat shock at 42°C for 45 s and ice for 2 min. One ml LB (See Appendix 3) was added and culture was incubated at 230 rpm at 37°C for 1 h. Cells were then spun down and resuspended in 150 ul LB and plated out on LB selection plates.

3.17 Cell synchronization studies

Cell cycle is the universal process for cell reproduction. It underlies the growth and development of all living organisms. Hence the precision at which the events are executed will ensure the survival of the living organism. It is usually the loss of this precision which results in genomic instability which in turn leads to cancer formation. Cell cycle is defined as the period between two

mitotic divisions. The start being the interphase with mitosis being the actual division. Mitosis results in two identical daughter cells failing which will result in either duplication or loss of chromosomes. The cell cycle consists of 4 phases, G1, S, G2 and M phase. There is a double of DNA content from G1 to G2 phase. Mitosis consists of 4 stages namely prophase, where chromosomes becomes visible as extended structures; metaphase, where bivalent chromosomes aligned on equator; anaphase where homologue pairs move to opposite poles and telophase where chromosomes reaches the poles and decondense into chromatin. All these are collectively known as mitotic figures.

There are basically 2 cycle controls, the commitment to chromosome replication at G1 and the commitment to mitotic division at G2 phase. In addition, the cell cycle has many checkpoints which ensure correct duplication of cellular components. There are check points for division in G1 phase, DNA integrity at S phase and unreplicated DNA checks at G2 phase. The control of cell cycle proteins usually occurs via protein phosphorylation and dephosphorylation. Kinases which are responsible for phosphorylation, consists of a catalytic subunit called cyclin-dependent kinases (cdks) and a regulatory cyclin partner. It is the timing of activity of the various forms of cdk/cyclins that regulate cell cycle progression (Lewin Benjamin, 1997). CDKs mentioned above regulate the cell cycles by fluctuation during the cycle. CDKs have a catalytic protein kinase subunit and a regulatory cyclin subunit. These were found universally from yeast to mammals. CDKs are regulated through several mechanisms such as the availability of the cyclin regulatory subunit, the phosphorylation of the catalytic subunit and by the presence of inhibitory

subunits. Expression of these proteins are also regulated at both transcription and translation level. Lastly, degradation of these proteins involved the ubiquitin-mediated proteolysis. This shows a large number of proteins orchestrating in cell cycle progression and hence in cancer formation, a large number of protein have to be likewise deregulated. Cell synchronization is used to allow precise study of regulatory mechanisms of cell cycle. Common ways include the use of serum starvation and reversible inhibitors such as mimosine. Mimosine, a non-protein plant amino acid, is an effective inhibitor of DNA replication and is able to block cell cycle progression in late G1 phase prior to onset of DNA synthesis (Vackova et al, 2003 and Krude T, 1999). Cells can then be released and test product or treatment can be done. The cell cycle can then be analyzed with flow cytometry of propidium iodine stained cells or fluorescent conjugated intracellular antibodies which detects intracellular cell cycle proteins such as cyclins.

Since previous results show an increased in MOST-1 expression in cancer cells, experiment to investigate MOST-1 association with cell cycle was carried out using mimosine cell synchronization studies. Cells were grown to 50% confluency and medium containing 0.5mM mimosine (Sigma) was then added and incubated at 37°C for 24 h. Medium was then removed and cells were washed with PBS and normal grown medium was added. Cells were then harvested at 0, 2, 4, 6, 8, 10, 12, 36, 48 h and subjected to immunofluorescent microscopy as described above or flow cytometry.

3.18 Overexpression and RNA interference studies

Since MOST-1 expression levels has been associated with high grade breast and prostate biopsies, experiments were design to investigate the effect of MOST-1 expression on normal-like and cancer cell lines proliferation and apoptosis levels.

3.18.1. Overexpression

MOST-1 was amplified from HA-MOST1 using primers pcDNAF (5'CACCATGGGAGAATGTCCCCACC3') and pcDNARev (5'CCATGTTGTCCAGGCTGGTCTCG 3') to give a product size encompassing the MOST-1 ORF of 297bp and cloned into pcDNA 3.1 expression vector with V5 tag (Invitrogen). Clones were sequenced to ensure that the protein was in frame with tag and was checked for expression via in vitro translation as described by manufacturers (Promega) followed by Western analysis. Cells were then transfected with lipofectamine 2000 (Invitrogen) according the recommendation. Timed extraction (12, 24, 36 and 48 h) of cell lysates followed by Western analysis using anti-V5-AP (Invitrogen) were done to optimized time of expression. After determination of time, cells were harvested for RT-PCR analysis using MOST-1 gene specific primers TR4.2 and TF2 and housekeeping primers G3DPH. The cells were also subjected to cell proliferation assay via BrdU labeling (Clontech) and TUNEL assay (Clontech) described below.

3.18.2. RNA interference

Duplex RNA from Dharmacon Inc, targeting the ORF start region was constructed with the following sequences: 5'-UGUCCCCACCUCGUGGAUGdTdT-3' and 5'-CAUCCACGAGGUGGGGACAdTdT-5'. The oligonucleotides were annealed according to the Dharmacon protocol. 3 ul of 20 mM duplex RNA were used per 24 wells as recommended by manufacturer. As a control for the specificity, we used a scrambled ribo-oligonucleotide pair with the following sequences 5'-GCGCGCUUUGUAGGAUUCGdTdT-3' and 5'-CGAAUCCUACAAAGCGCGCdTdT-3'. To test the cell lines for RNAi silencing, luciferase GL3-duplex with the following sequences 5'-CUUAGCCUGAGUACUUCGAdTdT-3' and 5'-UCGAAGUACUCAGCGUAAGdTdT-3' to co-transfect the cells using lipofectamine 2000 together with pGL3-control vector (Promega). Again, cells were subjected to cell proliferation assay via BrdU labeling (Roche) and TUNEL assay (Clontech) described below.

3.18.3. Cell Proliferation assay

To detect cell proliferation, there are several methods which use detection of proliferating antigens, measurement of DNA synthesis and alamar Blue^{TM} reduction. Example of proliferating antigens are proliferating cellular nuclear antigen (PCNA) which is essential for DNA synthesis and cyclins such as cyclin E which is expressed as cells move from G1 to S phase. DNA synthesis can be measured using ^{3}H-thymidine or bromodeozyuridine (BrdU)

incorporation into nascently synthesized DNA. The latter is preferred as it is non-radioactive and can be easily detected with an antibody (Vackova et al, 2003). This makes it suitable for immunohistochemistry. Alamar Blue™ reduction occurs when metabolites intermediates of the internal environment of proliferating cells (which is more reduced compared to non-proliferating cells). The reduction is then monitored spectrophotometrically. For our experiments, BrdU assay was chosen for its ease in detection via flow cytometry.

Thymidine analogue, BrdU was added for last 6 h of culture as recommended. 1 x 10^6 cells/ml concentration was then fixed in 70% ethanol for 4 h prior to acid treatment of 15 min. Cells were then labeled with anti BrdU-FITC. Before flow analysis via Coulter imaging system, PI was added as a counter stain.

3.18.4. Apoptosis assay

One million cells/ml concentration was fixed in 4% paraformaldehyde for 20 min prior to 0.1% Triton X-100 permeabilization of 5 min on ice. Cells were then labeled with terminal deoxynucleotidyl transferase (TdT)-mediated dUTP nick-end-labeling with FITC-dUTP. Before flow analysis via Coulter imaging system, PI was added as a counter stain.

3.19 Yeast two hybrid

Protein-protein interactions are critical to biological processes of the cells. Y2H screening provides a fast and rapid method for identification on putative protein-protein interactions. Prior to the development of Y2H,

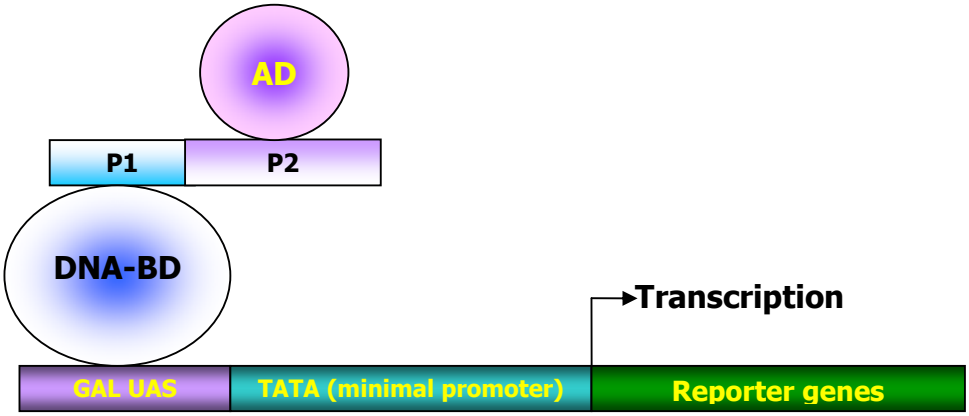
identification of physical protein-protein interactions and their subsequent functional characterization are laborious and time consuming. The main advantage of Y2H is the ease of use of yeast as a genetic selection system. In addition, it relies on an assay performed *in vivo* and is thus not limited by artificial conditions of *in vitro* assays. Using a common protocol, a wide variety of protein-protein interaction can be detected (Vidal et al, 1999). The main disadvantage of the system is the presence of high percentage of false positives which have to be "weed" out after the initial screening.

The Y2H system using 2 reporter yeast strains of different mating types. By itself, the yeast is unable to grow on dropout medias which do not contain Ade, His, Leu and Trp. Upon mating, if the 2 proteins (which are attached to respective activation and DNA-binding domain) in each strain interacts, the DNA-binding domain would bind to an Gal promoter fused with the reporter genes and the activation domain if in close proximity will activate transcription of the reporter gene which will allow the hybrids to grow on dropout media as depicted in figure 4.

Yeast two hybrid screening was carried out using MATCHMAKER GAL4 Two-Hybrid system (Clontech). MOST-1 cDNA contain inherent NcoI restriction sites. Entire putative ORF of MOST-1 was generated using high-fidelity Pfu DNA polymerase (Promega). The amplicon and pAS2-1 vector were then digested with the above restriction enzymes and ligated O/N at 4°C and transformed into *E. coli* competent cells, DH5α. In frame cloning of the bait construct was confirmed by sequencing prior to transformation into PJ69-2A

yeast cells using lithium acetate-based method. Fusion proteins from yeast lysates were checked with MOST-1 specific antibodies via Western blot.

Figure 6: Schematic diagram on the mechanism of Y2H screen



DNA-binding domain (DNA-BD)

- **target to specific promoter sequences such as UAS**

Activation domain (AD)

- **facilitate assembly of transcription complex allowing for initiation of transcription**

(Neither of these alone is able to activate transcription.)

MATCHMAKER 2 pretransformed cDNA libraries (Clontech) of human bone marrow were screened via yeast mating. Mating was carried out by incubating 1 ml of the pretransformed library culture with the bait culture constructed above in YPDA/kan O/N at 30°C shaking at 30 rpm. Mating cultures were checked under phase-contrast microscope after 20 h for presence of zygotes and allowed to mate for a further 4 hours after which it is collected by centrifuging at 1000 g for 10 min. Mating flask was rinsed twice with YPDA/kan and the first pellet was resuspended before collecting cells by repeating centrifugation. Cell pellet was then resuspended in 10 ml of

YPDA/kan. 100 µl of 1:10,000, 1: 1000, 1:100 and 1:10 dilution of the mating mixture were plated on SD/-leu, SD/-trp and SD/-leu/-trp plates to assay for mating efficiency while the remaining mating mixture were spread onto SD/-Ade/-His/-Leu/-Trp (QDO) plates at 200 µl each. Plates were then incubated at 30°C for 21 days. Colony appearance was checked every 3 days and subcultured on fresh QDO plates. B-gal colony lift filter assay and direct colony PCR screening of positive clones for the presence of both the bait and the library were done using FAD and RADO (primers previously described in Sim et al, 2002) and MOST-1 specific primers. Sequencing of clones was done and analyzed using BLASTN and BLASTP algorithms.

3.20 Transfection of mammalian cell lines

Transfections was done with lipofectamine 2000 (Invitrogen) for overexpression and RNAi studies and cytofectin (Biorad) for Co-immunoprecipitation studies.

3.21 Co-immunoprecipitation

The bait construct in yeast two hybrid experiments was subjected to *Sal*1 and *Xho*1 restriction enzyme reactions and cloned into pCMV-HA vector (Clontech) giving rise to HA-MOST1 while library interactors were cloned into pCMV-myc vector (Clontech) using restriction enzymes *Eco*R1 and *Xho*1. In frame cloning was confirmed via sequencing.

A total of 1 ug of plasmid DNA were used in separate transfections using cytofectene (Biorad) according to manufacturer's instruction. RPMI-

1640 supplemented with 10% FCS and opti-MEM 1 (Gibco BRL). BHK cells at ~80% confluency were used for transfection and incubated in CO_2 incubator for a period of 24 to 48 hours. Cells were washed trice with cold PBS prior to incubation on ice for 30 min with mild lysis buffer containing 40 mM Tris (pH 7.4), 300 mM NaCl, 20 mM EDTA (pH 8), 20% glycerol, 2% NP 40 and complete protease inhibitors (Roche). Cell debris was removed by centrifugation at 14,000 rpm for 15 min at 4°C. Lysates were then mixed in vitro and expressed proteins were co-immunoprecipitated overnight at 4°C with the respective antibodies and subsequently pulled down with protein A/G Plus agarose (Oncogene) for 2 hours before washing twice with PBS. SDS loading buffer was then added and the mixture was heated at 95°C for 5 min before loading onto SDS-PAGE gel and subjected to Western analysis.

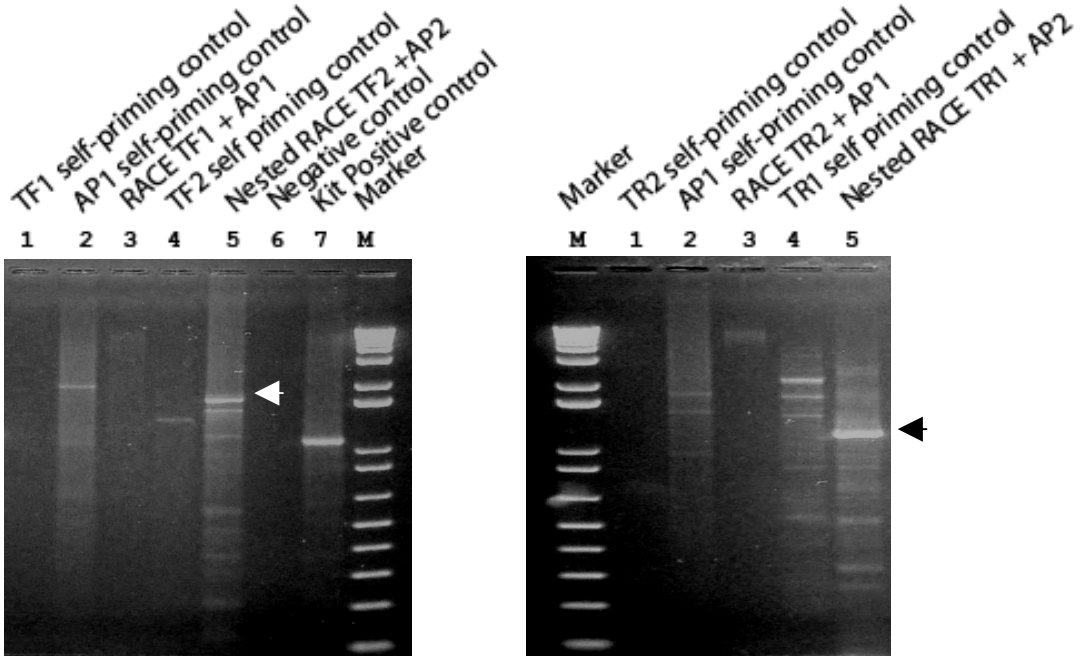
CHAPTER 4. RESULTS

4.1. Elucidation of MOST-1 full length sequence

A novel EST was obtained using HPV E6 consensus primers on MOLT-4 cDNA and named MOST-1 (MOLT-4 sequence tag 1). Sequence analysis of this EST showed no homology to any known gene in GenBank database (www.ncbi.nlm.nih.gov) or to HPV E6 gene. To obtain the full-length cDNA, RACE was performed on 2 libraries. MOLT-4 T-lymphoblastic leukemia cDNA of which MOST-1 was initially derived from and human fetal lung (MRC-5) cDNA. MRC-5 was chosen to demonstrate if MOST-1 could have any developmental function. Screening for this novel EST in the two RACE libraries authenticated its presence in both (Fig 7A). Isolation of the full-length cDNA sequence corresponding to this EST from which specific primer pairs TF1, TR1, TF2, TR2 were designed for use in RACE experiments. All DNA bands obtained were excised and sequenced, of which products of ~1.6 kb and ~1.2 kb were generated via 3′ and 5′ MOLT-4 RACE, respectively. (Figure 7B) The MRC-5 library yielded bands of ~1.6kb and ~0.5kb from 3′ and 5′ RACE reactions, respectively. See Figure 7C. All other bands were either truncated sequences or self-priming products of primers. Cycle sequencing of the overlapping 3′ and 5′ RACE fragments revealed composite cDNA sequences of 2786 bp and 2054 bp in MOLT-4 and fetal lung, respectively (Fig. 8). The latter cDNA sequence in fetal lung was identical to nt 664-2705 of its MOLT-4 counterpart. The resultant cDNA sequence of 2786 bp was found to contain a continuous ORF of 297 bp with an in-frame nonsense codon 168 bp upstream of the putative start codon. The first methionine codon lies in an extremely

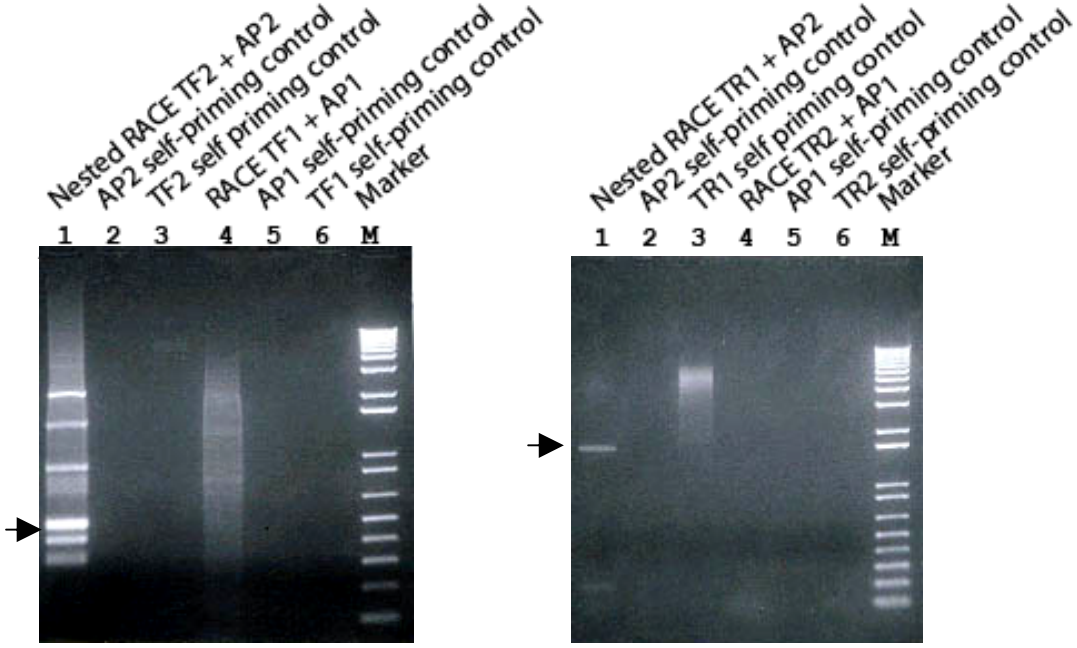
favorable context for translation initiation, i.e. GGCACC*ATG*G, which conforms exactly to the ideal sequence for a eukaryotic translation start site (Kozak, 1991). Northern analysis results in a high background which may be due to the Alu repeat present in *MOST-1*. Hence screening of normal tissue cDNA was done subsequently. See Section 4.4.

Figure 7A: RACE screen of MRC-5 (human fetal lung) and MOLT-4 (T-lymphoblastic leukemia) cDNA library. MOST-1 EST was found in both MOLT-4 and MRC-5 RACE cDNA library. Screening was done with primers TR4.1 and TF1 resulting in a 279 bp PCR product prior to RACE experiments.



3' RACE 5' RACE

Figure 7B: RACE products of MOLT-4 cDNA RACE library. As MOST-1 cDNA was present in low amounts, nested RACE experiments was done. 3' RACE showed a ~1.6 kb fragment, lower bands were truncated versions. 5' RACE obtained a 1.3 kb fragment. Self priming controls were done to eliminate non-specific bands in the reaction. Positive controls were done with G3DPH primers provided to verify RACE library integrity.



3' RACE 5' RACE

Figure 7C: RACE products of MRC-5 cDNA RACE library. As MOST-1 cDNA was present in low amounts, nested RACE experiments was done. 3' RACE obtained a 0.5 kb fragment all other fragments were truncated versions or self priming products. 5' RACE obtained a ~ 1.6 kb fragment. Self priming controls were done to eliminate non-specific bands in the reaction.

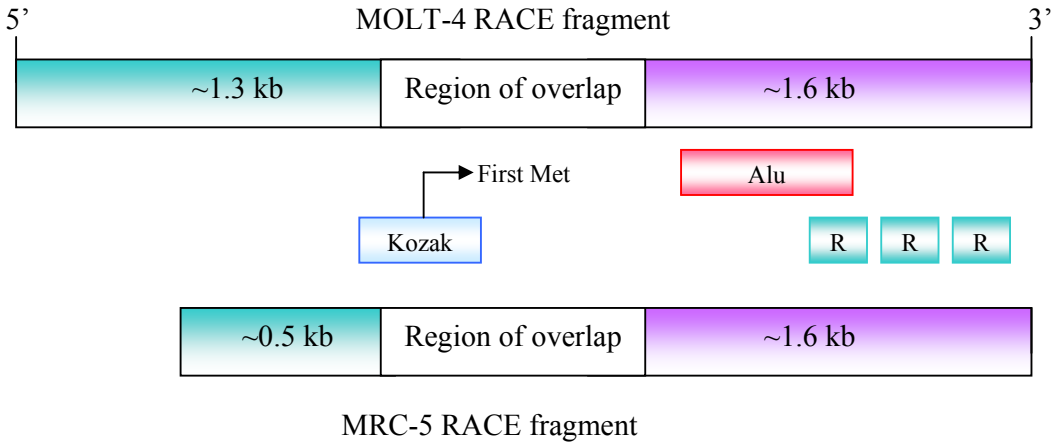


Figure 8: Schematic diagram of *MOST-1* full length cDNA upon sequence analysis. There is a region of overlap of the 3’ and 5’ RACE fragments as indicated in both RACE libraries. Red box showing Alu repeat region. Alu repeat has been implicated for gene amplification and jumping phenomenon. Blue box illustrates Kozak sequence location comprising of the first methionine for translation of ORF. Ideal Kozak sequence in MOST-1 indicates high translation efficiency. Green box (R) represents AUUUA RNA degradation motifs. Presence of 3 R signals suggests an unstable RNA, which indicates a tight control on mRNA expression level.

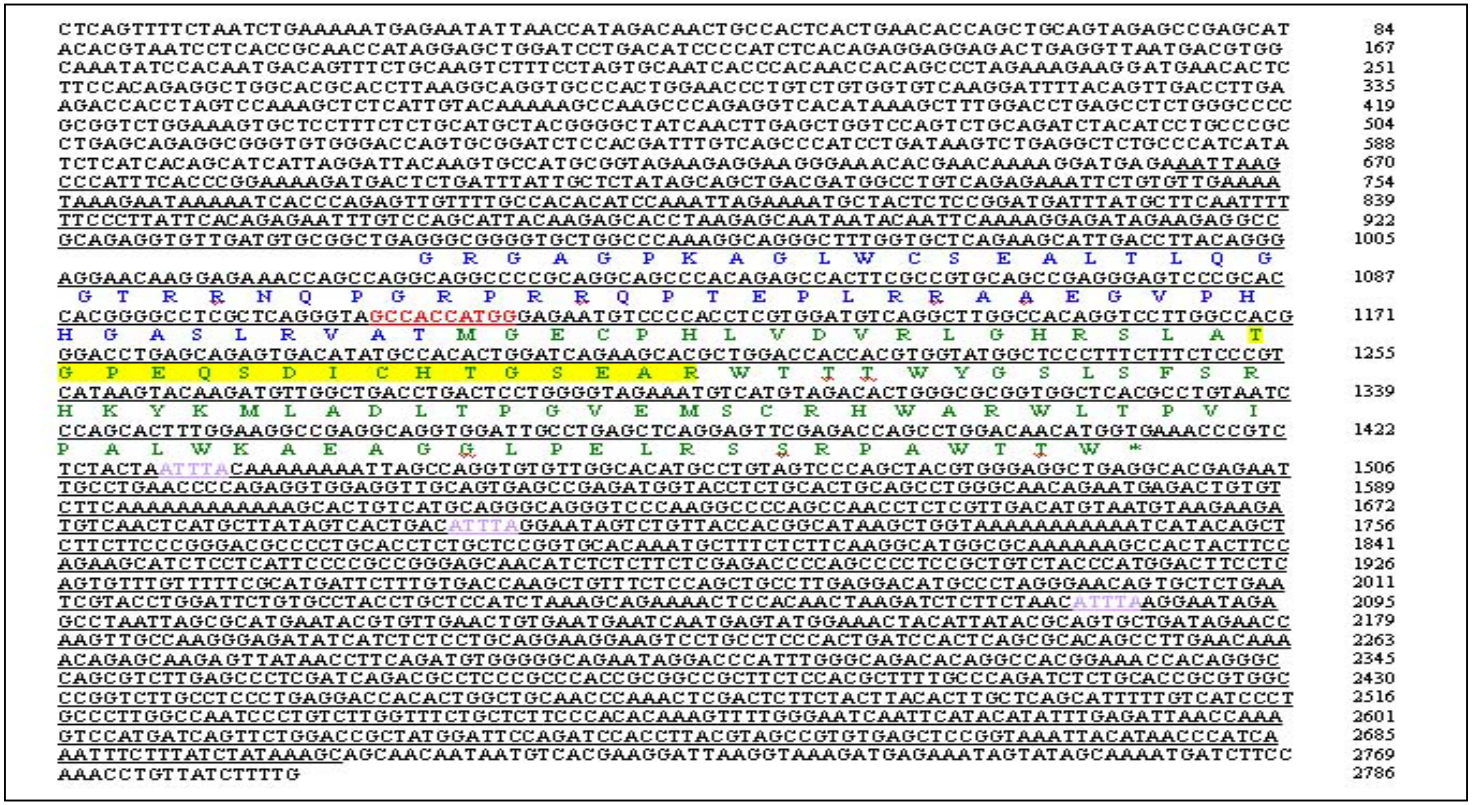


Figure 9: Nucleotide sequence of full length MOST-1 sequence from MOLT-4 RACE. Putative ORF sequence in green; Highlighted region in yellow is sequence used for synthetic peptide generation; Kozak sequence in red; RNA destabilizing signals in purple; Underlined sequence is the MRC-5 full length cDNA sequence which overlaps with MOLT-4 derived sequence.

4.2. Bioinformatics analysis of MOST-1 putative ORF

A computational translation of the methionine at the Kozak sequence predicted a potential ORF of 99 amino acids (illustrated in Figure 9 as green) forming a basic, hydrophilic protein (grand average of hydrophobicity of –0.417) of 11,219 Da. This putative protein had an isoelectric point of 8.59, and was classified as unstable in nature with an instability index computed to be 76.92 using ProtParam tool (www.expasy.ch). In this program, instability index above 40 indicated an unstable protein based on previous statistical analysis of 12 unstable and 32 stable proteins (Guruprasad et al, 1990). No notable amino acid sequence motifs were evident but three ATTTA mRNA destabilizing signals (Shaw and Kamen, 1986) were present in the 3′ untranslated region (UTR) of the gene. *MOST-1* cDNA sequence was deposited in GenBank under accession number AF220264 with a symbol C8orf17 by the HUGO Nomenclature Committee.

Analysis of MOST-1 putative ORF

Figure 10 shows glycine at aa2 was predicted to be a N-myristoylation site while threonine at aa19 and 29 were predicted to be casein kinase II phosphorylation sites and serine at aa69 and 91 were predicted to be protein kinase C phosphorylation sites. Serine at aa91 is also predicted to be a site for O-glycosylation while lysine80 is a potential sumoylation site. Protein predict program predicted MOST-1

to be non-globular in nature which correlates with a predicted secondary structure with predominant extended sheet structure from aa7-11, aa34-56, aa60-64, aa67-71 and aa80-81. At the same time, ProtParam predicts MOST-1 to be an unstable protein.

4.3. MOST-1 genomic structure analysis

To search for introns, concurrent PCR reactions were carried out with primer pairs flanking approximately 200bp products with overlapping sequences. The PCR products were then excised and the sequences of PCR products amplified from cDNA and genomic templates respectively using identical primer pairs were compared but showed no differences. Primer pairs indicated in Figure 11 with templates cDNA and genomic respectively. Most of the primer pairs gave identical size fragments with the exception of those primers in the 5'UTR of *MOST-1* when using MRC-5 as a template due to the shorter gene fragment as found by RACE. *MOST-1* is also present in lower amounts in MRC-5 cDNA preparations resulting in low intensity products which are apparent only upon high exposure time to UV during photography. Since the products were the same size, it indicates that *MOST-1* do not contain introns and hence intronless.

Subjecting the *MOST-1* cDNA sequence to computational analysis by exon-intron search engines such as Grail also arrived at the same conclusion that this gene was intronless.

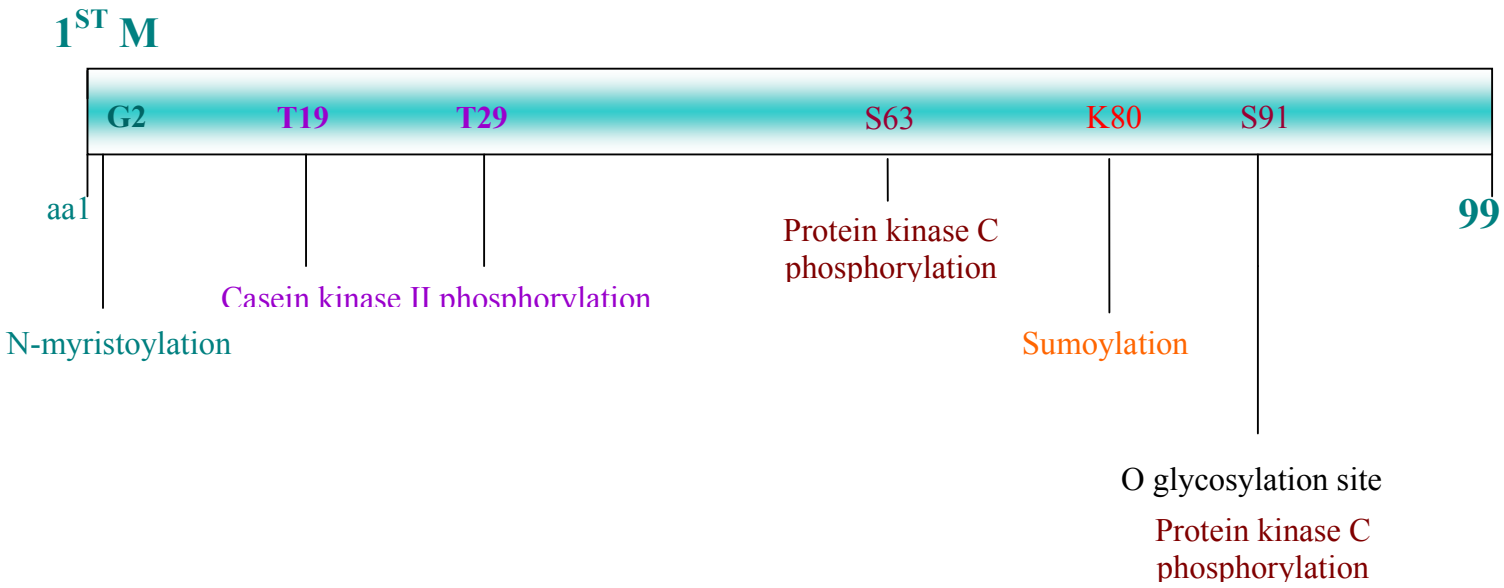


FIGURE 10: Summary of computational analysis of *MOST-1* putative ORF. 5 motifs were elucidated using various bioinformatics analysis indicated in methods. N-myristoylation is a protein modification which promotes binding of protein to membrane. CKII phosphorylates at S/T and tend to be elevated in neoplastic and proliferating tissues. Sumoylation is an enzymatic process analogous to ubiquitination and plays an important role in intracellular protein localization. O glycosylation is found mainly on membrane proteins. Protparam predicts putative ORF to be highly unstable, with a predicted size of 11.2kDa.

FIGURE 11A: Intron/Exon Mapping of MOLT4 cDNA and gDNA

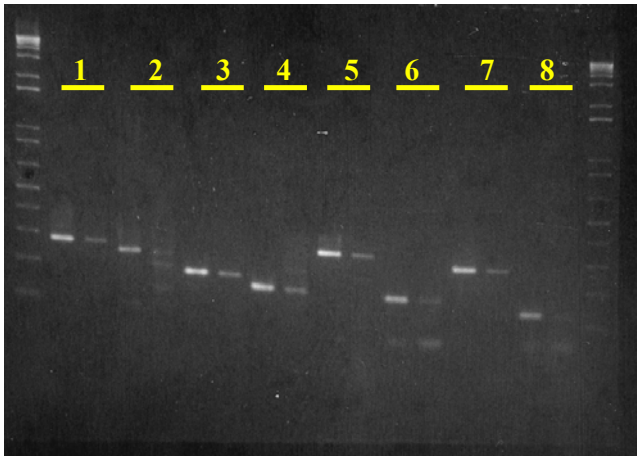


FIGURE 11B: Intron/Exon Mapping of MRC-5 cDNA and gDNA

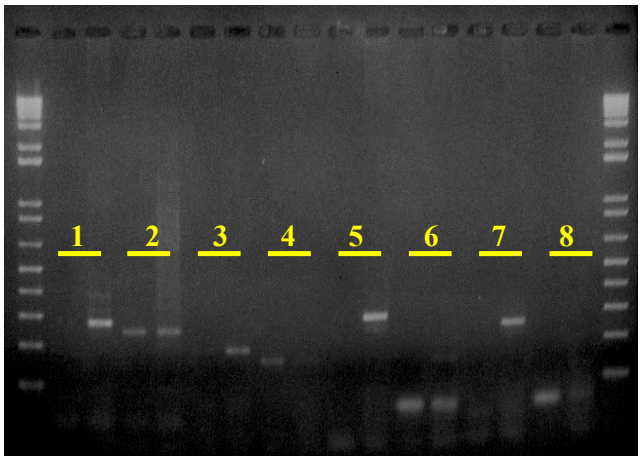


Figure 11: Genomic Structure analysis of *MOST-1.* Mapping of intron/exon junction was done using primers pairs indicated in Table 6. PCR of paired cDNA and gDNA shows no difference in cDNA and gDNA PCR products from predicted size suggesting MOST-1 to be intronless.

Table 7: Primer Pairs and product size used in mapping for Figure 11

Primer Pair	Primer	Location	Product Size
1	TR4.1	nt 1387	279bp
	TF1	nt 1108	
2	TR4.1	nt 1387	250bp
	TF2	nt 1137	
3	TR4.2	nt 1299	191bp
	TF1	nt 1108	
4	TR4.2	nt 1299	162bp
	TF2	nt 1137	
5	TR2	nt 1157	297bp
	TF5	nt 860	
6	TR2	nt 1157	157bp
	TF7	nt 1000	
7	TR1	nt 1125	265bp
	TF5	nt 860	
8	TR1	nt 1125	125bp
	TF7	nt 1000	

4.4. Expression Profile of MOST-1

Cell Line Screening shows that *MOST-1* is ubiquitously expressed in cancer cell lines but differentially expressed in normal human tissues. RT-PCR with *MOST-1* specific primers TF5 and TR4.2 yielded expected 440-bp target fragments of detectable but varying intensities from cDNAs of 19 human cancer cell lines and two normal cell lines tested. However, out of 16 normal human adult tissues tested, *MOST-1* RT-PCR products were detectable in only 9 tissues (heart, kidney, liver, pancreas, small intestine, ovary, testis, prostate, thymus), but not in others (brain, lung, skeletal muscle, colon, placenta, spleen, peripheral blood leukocyte). The *G3PDH* and/or *HUEL* housekeeping genes served as cDNA quality controls, and were successfully amplified from the tested human cell lines and adult tissues. To authenticate the PCR reaction, human total cDNA with or without G3DPH primers were carried out in the positive and negative controls respectively (Figure 12B top panel, lanes 17 and 18). Figure 12 illustrated the reamplified products, while Table 7 illustrated the initial products band intensity reading followed by normalizing with G3DPH product band intensity. See Appendix 4 for densitometric readings. Both reactions were stopped at the log phase of amplification for semi-quantitation.

Table 8: Comparative MOST-1 Expression in Human Tissues, Normal and Cancer Cell Lines

Normal tissues	*MOST-1* : *G3DPH* ratio[a]	Cell lines	*MOST-1* : *G3DPH* ratio[a]
Pancreas	0.61	MRC-5 fetal lung cells	1.35
Prostate	0.57	Breast myoepithelial cells	1.15
Ovary	0.55	MDA-MB-231 breast adenocarcinoma	2.11
Testis	0.33	Hs578T breast ductal carcinoma	1.14
Liver	0.33	MCF7 breast adenocarcinoma	1.13
Heart	0.31	ZR-75-1 breast ductal carcinoma	0.44
Kidney	0.31	PC-3 prostate adenocarinoma	1.20
Thymus	0.28	DU145 prostate carcinoma	1.13
Small intestine	0.26	CaSki cervical carcinoma	1.18
Lung*	0.17	HeLa cervical carcinoma	1.06
Peripheral leukocyte*	0.17	SiHa cervical carcinoma	0.90
Skeletal muscle*	0.14	T24 bladder carcinoma	0.86
Placenta*	0.08	HepG2 hepatocellular carcinoma	1.08
Colon*	0.06	Hep3B hepatocellular carcinoma	0.95
Brain*	0.02	PLC/PRF/5 hepatocellular carcinoma	0.69
Spleen*	0.00	Mahlavu hepatocellular carcinoma	0.64
		KATO III gastric carcinoma	0.76
		U-937 histolytic lymphoma	1.59
		Raji Burkitt's lymphoma	0.96
		HL-60 promyelocytic leukemia	0.95
		MOLT-4 T-lymphoblastic leukemia	0.94

[a]All cDNA samples were subjected to PCR with *MOST-1* and *G3DPH* primers, and the individual ratios of the densitometric readings of their amplified products were calculated.

*Denotes samples for which no *MOST-1* amplified products were observed by agarose gel electrophoresis.

FIGURE 12A: Cell Line Screening of MOST-1.

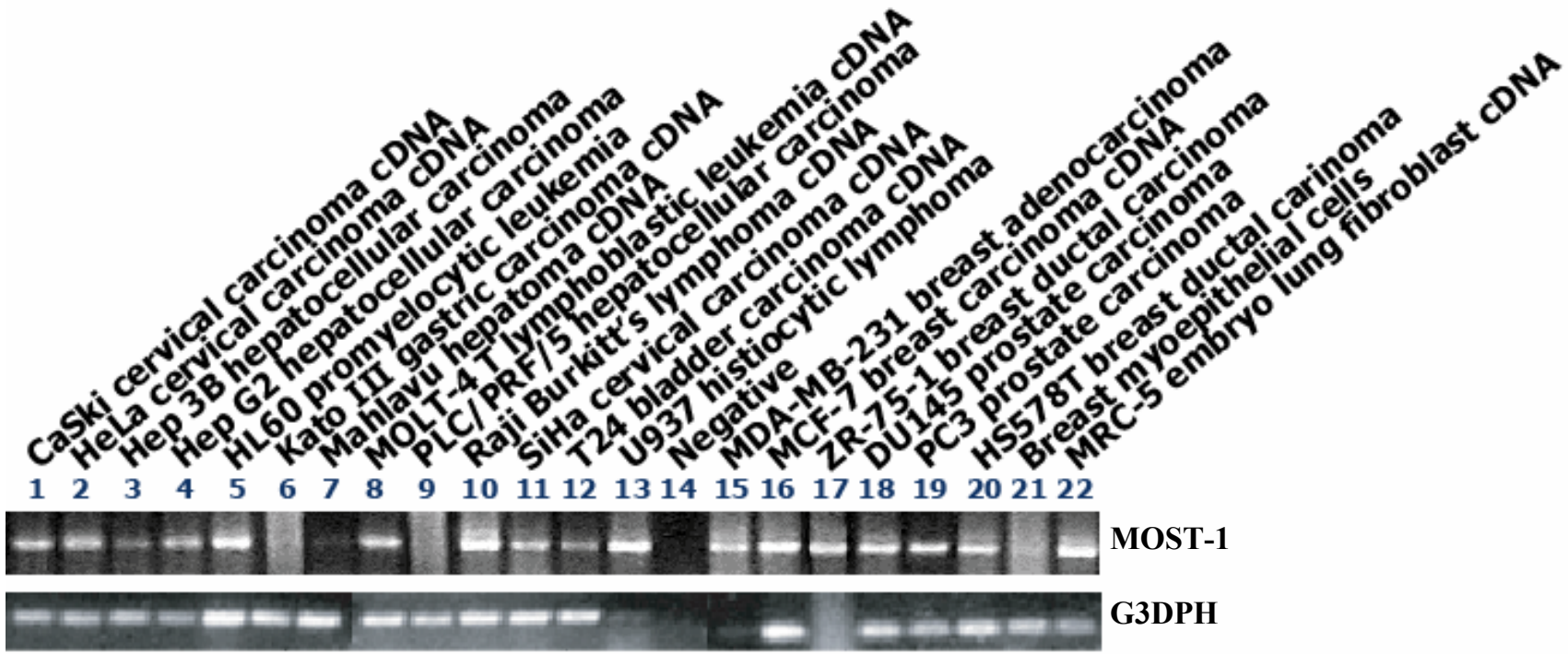


FIGURE 12B: Normal tissue cDNA Screen

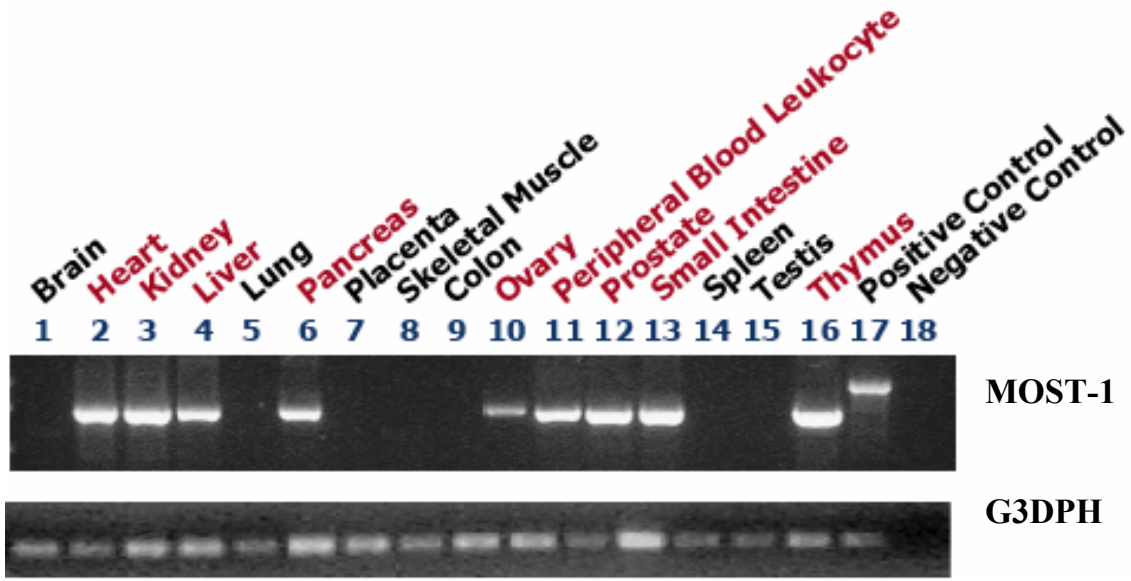


Figure 12: MOST-1 expression profile. Figure 12A shows MOST-1 found in all cancer cell lines tested, one fetal (lung) cell line and one normal epithelial cell line. This suggest a role for MOST-1 in cancer progression. Figure 12B shows MOST-1 found to be differentially expressed in normal tissues suggesting a tissue specific role which is deregulated in cancer cells. Tissues expressing MOST-1 are highlight in red.

4.5. Genomic Localization of MOST-1

Fluorescent in situ hybridization (FISH) Map *MOST-1* gene to chromosome 8q24.2 with a detection efficiency of ~50% was attained with the FISH probe, i.e. out of 100 mitotic figures checked, 50 displayed hybridization signals on one pair of homologous chromosomes. Using DAPI banding to identify the specific chromosome, the assignment between the signal from the probe and the long arm of chromosome 8 was obtained. Based on a summary of ten photographs, *MOST-1* was localized to a more precise position at chromosome 8q24.2 (Fig. 13), and no other loci were detected with the probe. The *MOST-1* sequence was also mapped by UniGene computational analysis to chromosome 8q24.3.

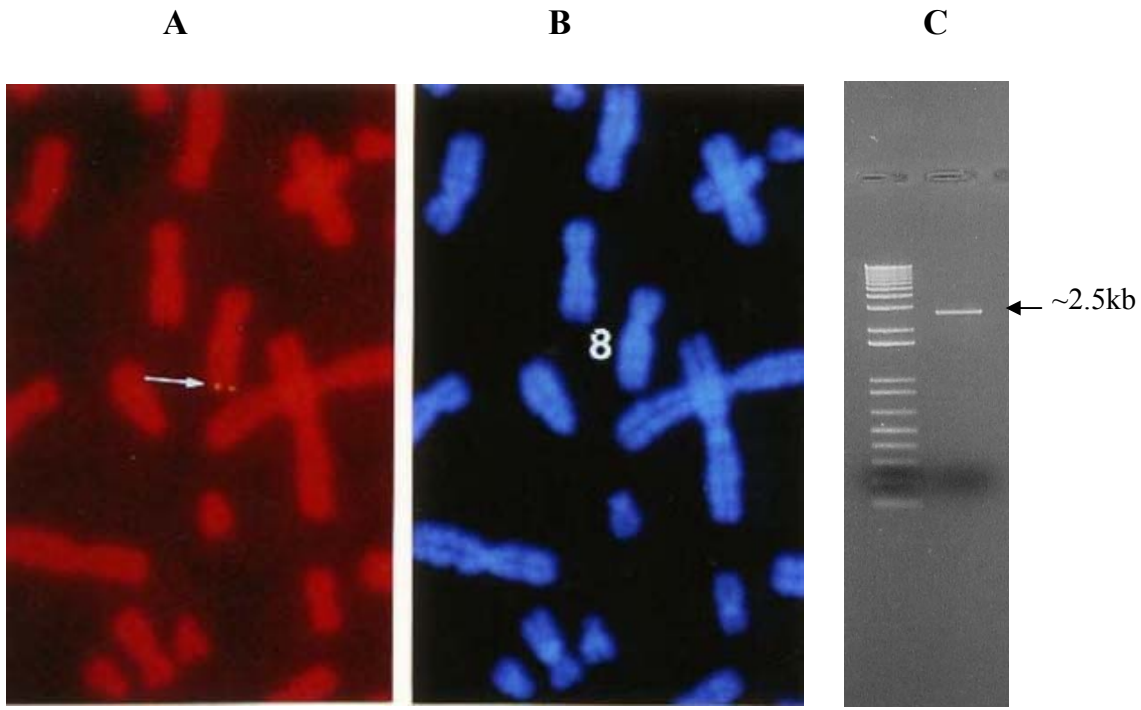
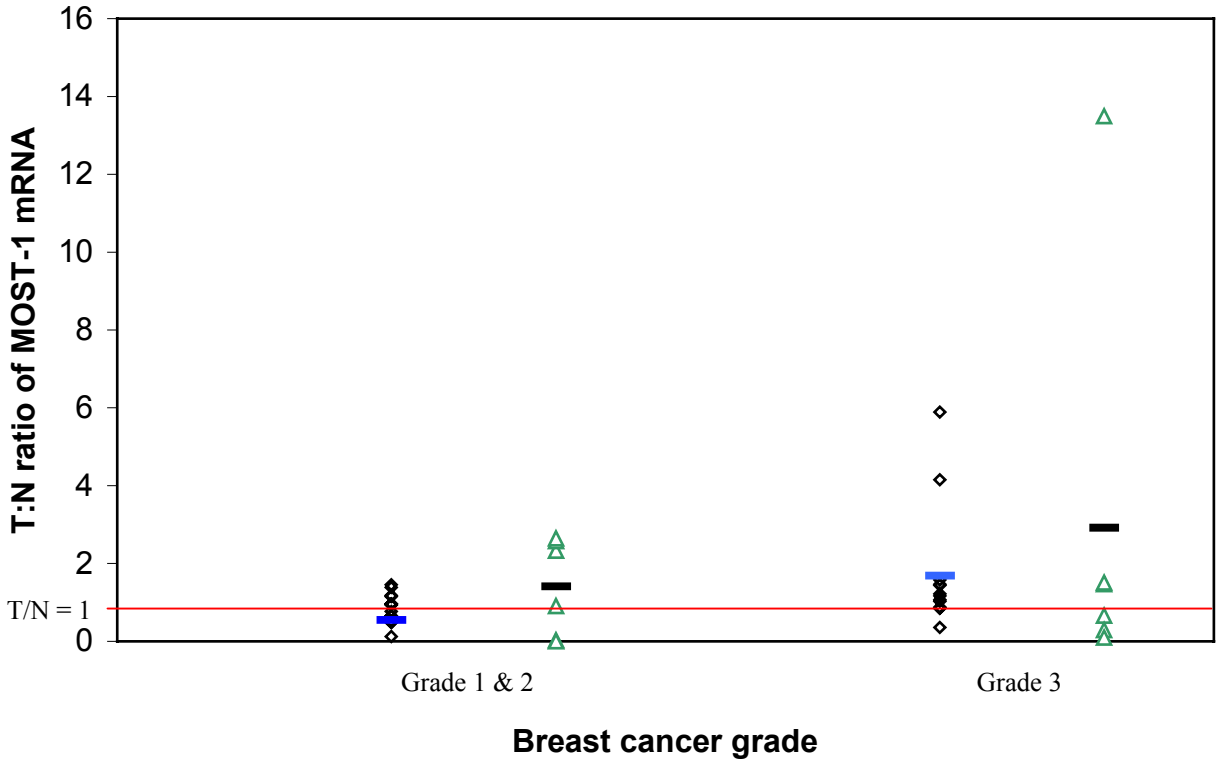


Figure 13: Chromosomal Localization of *MOST-1*. FISH analysis of MOST-1 with specific probe shows MOST-1 localization to 8q24.2, a region reported to be amplified in a number of cancers especially breast and prostate cancers. A: Staining of chromosome with *MOST-1* specific probe on both alleles. B: DAPI staining of chromosomes of same metaphase spread. C: Chromosome probe for *MOST-1* used for FISH experiment.

4.6. Breast Biopsies Screening

Semi-quantitative PCR analysis shows that *MOST-1* is overexpressed in a Subset of Breast Cancers

MOST-1 expression in paired breast tissue biopsies was compared by first normalizing densitometric readings of *MOST-1* with *G3PDH* RT-PCR products, and then calculating the ratios between corresponding tumor and adjacent normal tissues (T:N). Out of 27 paired biopsy samples tested, relative T:N expression of *MOST-1* was increased in 11 pairs (41%), unchanged in 4 (15%), and decreased in another 12 (44%). Real-time RT-PCR performed on 12 randomly selected pairs of tumor and normal tissues revealed *MOST-1* expression data in six grade 1 and 2 cancers and six grade 3 cancers consistent with semi-quantitative RT-PCR results Graph 1. The T:N ratio of *MOST-1* expression demonstrated a correlation with breast cancer grade, i.e. mean ratios for grade 1 and 2 versus grade 3 were 0.84 versus 1.69 by semi-quantitative RT-PCR, and 1.41 versus 2.92 by real-time RT-PCR. Though there is a relative 2 times increase in ratio for both methods, the differences between the means for grade 3 versus grade 1 and 2 cancers were not statistically significant due to low sample numbers.



Graph 1: T/N Ratio of *MOST-1* gene expression in tumor biopsies compared to normal showed increased MOST-1 expression in tumor biopsies. Semi-quantitative analysis of normalized densitometric reading of PCR products in log phase are expressed as open diamonds while real time cT values are indicated in green triangles. Average of semi-quantitative analysis of Grades 1 & 2 and 3 biopsies are 0.837 and 1.687 (indicated by blue lines) respectively; demonstrating a 2 fold increase in MOST-1 expression. Real time analysis of 12 randomly picked samples in the pool showed a similar 2 fold increase in MOST-1 expression levels with mean T/N ratio of 1.415 and 2.91 respectively for Grades 1 & 2 compared to Grade 3 biopsies (indicated by black lines).

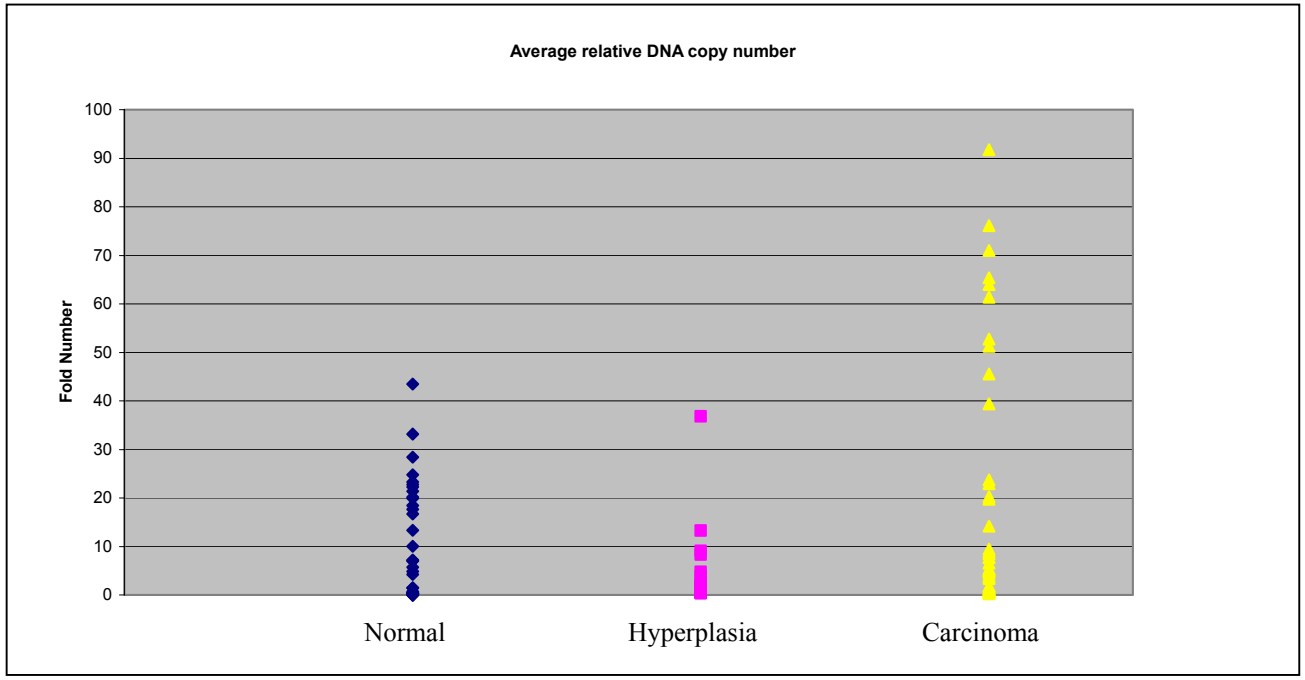
4.7. Prostate Biopsies Screening

Real time PCR analysis *MOST-1* DNA Copy Levels are increased in Prostatic Neoplasia

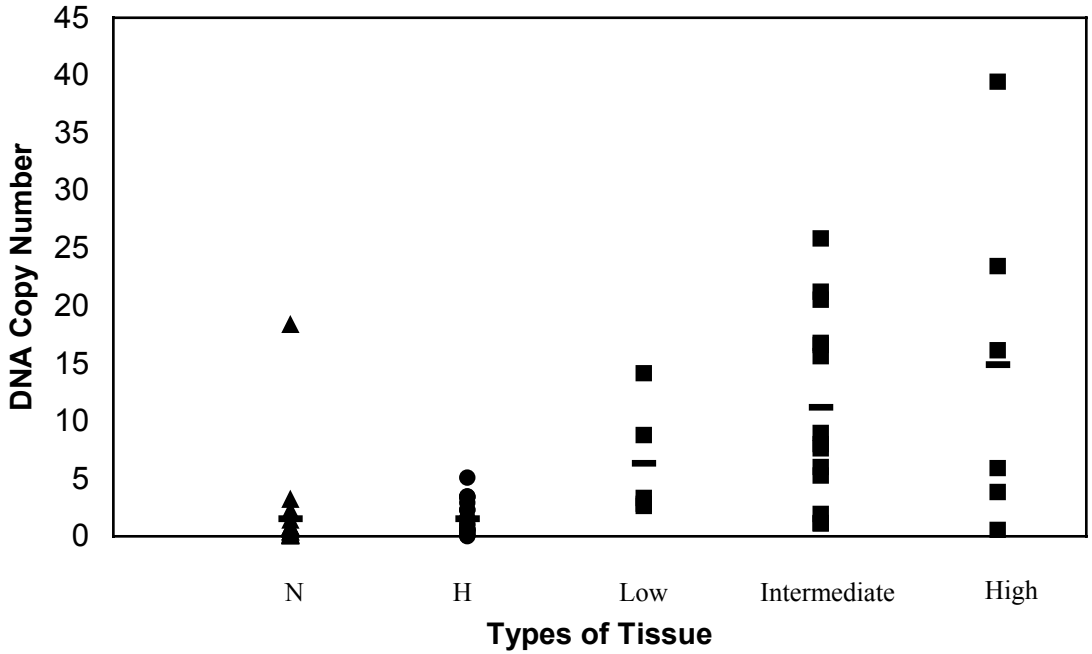
Owing to the relatively low incidence of prostate cancers in Singapore, a statistically sufficient number of paired biopsies of prostatic cancer and adjacent normal tissues were not readily available. Hence, we resorted to analyze DNAs from a total of 61 histological assessed, archival prostate biopsy sections, i.e. 15 normal prostates, 21 benign prostatic hyperplasia and 25 prostatic carcinomas. All DNA samples were subjected to real-time PCR experiments to derive each individual ΔC_T value, i.e. the difference between the threshold cycles for *MOST-1* and *G3PDH* gene detection. The formula $2^{-\Delta C_T}$ was employed for quantitation of the DNA levels of *MOST-1*. By comparing the mean ΔC_T value of normal samples (as the benchmark reference) with those of prostatic neoplasia, relative *MOST-1* DNA levels were increased in cancer tissue compared to hyperplasia and normal tissues as shown in Graph 2A. By subdivision of cancer tissue to low, intermediate and high grade carcinoma based on Gleason scores which correspond to 3-5, 6-7 and 8-9 respectively. *MOST-1* DNA copy number was shown to increase by 1.4-fold in hyperplasia, by 9.9, 7.5 and 4.2-fold in high, intermediate and low grade carcinomas. Multiple comparisons between the mean for normal tissues against the individual means for neoplastic tissues indicated that the differences for hyperplasia or low grade

cancers were statistically significant (p=0.014 and p=0.002, respectively), and for intermediate or high grade cancers were highly significant (p=0.000 for both grades) as shown in Graph 2B.

Graph 2: Relative real time quantification of *MOST-1* in prostate biopsies.



Graph 2A: Threshold cycles of *MOST-1* normalized with G3DPH showing normalized relative DNA copy number in 3 groups of tissues. Results indicate that there is a general trend for increasing DNA copy number in carcinoma compared to normal but not so between hyperplasia and normal.



Graph 2B: Relative DNA copy number in prostate cancer biopsies expressed against normal tissue. DNA fold number from above was compared to mean of normal sample as benchmark to obtain relative DNA copy number. In addition, carcinoma samples were subdivided into low (Gleason score 3-5), intermediate (Gleason score 6-7), and high (Gleason score 8-9) grade carcinoma. Relative DNA copy number was increased by 1.4, 4.2, 7.5 and 9.9-fold in hyperplasia, low, intermediate, and high grade carcinoma respectively.

4.8. Polyclonal antibody generation and verification

4.8.1. Dot Blot analysis

To characterize MOST-1 protein, polyclonal antibody was raised against synthetic peptide TGPEQSDICHTGSEAR. Peptide sequence was subjected to BLASTp analysis against known protein sequence in GenBank prior to synthesis. Criteria for selection of peptide sequence include a hydrophilic and antigenic region in the putative ORF of MOST-1 protein after analysis with Plot Structure (Figure 14). 3 rabbits were immunized with synthetic peptide resulting in 3 polyclonal anti-sera. 2 of the anti-sera immuno-dot blot screening showed hyperimmune antibodies reacting strongly to conjugated and unconjugated peptide, whereas their respective preimmune sera were nonreactive (Figure 15). Polyclonal antibody #235 was then selected for affinity purification using peptide-bound column for use as a probe in Western and confocal analysis.

4.8.2. Polyclonal antibody recognizes aggregated form of MOST-1

During the characterization of affinity purified polyclonal antibody with Western analysis on MOLT-4 lysate showed that it only recognizes aggregated form of MOST-1 (Figure 16A). This prompted the expression of recombinant MOST-1 in TNT experiments which also illustrate the polyclonal antibody recognizing the high molecular weight product despite a ladder of different MOST-1 isoforms ranging from

15kDa to > 100kDa. This finding suggests the propensity of MOST-1 to self-aggregate in solution. Figure 16B shows differential treatment to TNT products illustrating that the aggregated MOST-1 is SDS soluble but upon thermal stress, the aggregated MOST-1 becomes SDS-insoluble. These data is in line with the computational analysis where MOST-1 is predicted to be unstable as a monomer and insoluble. Secondary structure analysis with PredictProtein (www.expasy.ch) shows a stretch of extended βsheets from aa34 to 70 suggesting the propensity of MOST-1 to aggregate in solution.

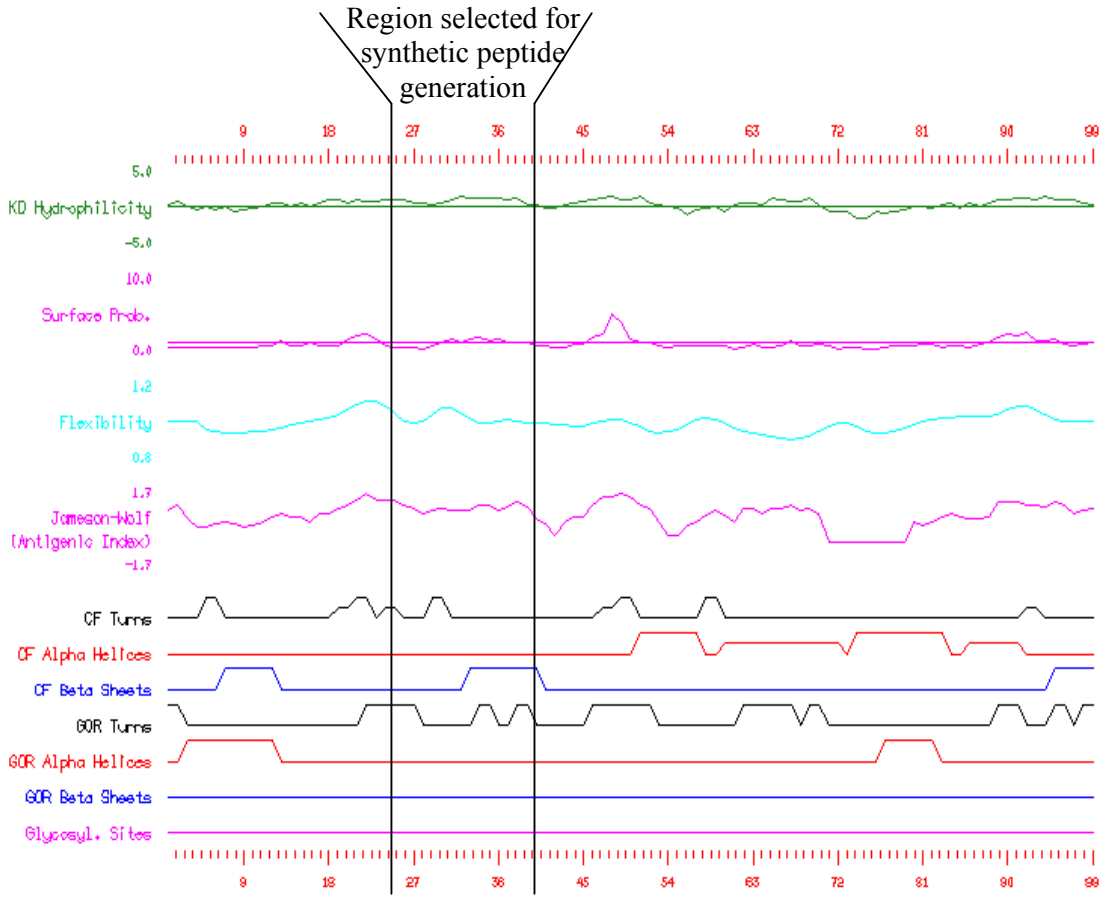


Figure 14: *MOST-1* ORF analysis using PlotStructure. Synthetic peptide region was selected based on its hydrophilic nature and antigenicity as indicated in analysis above.

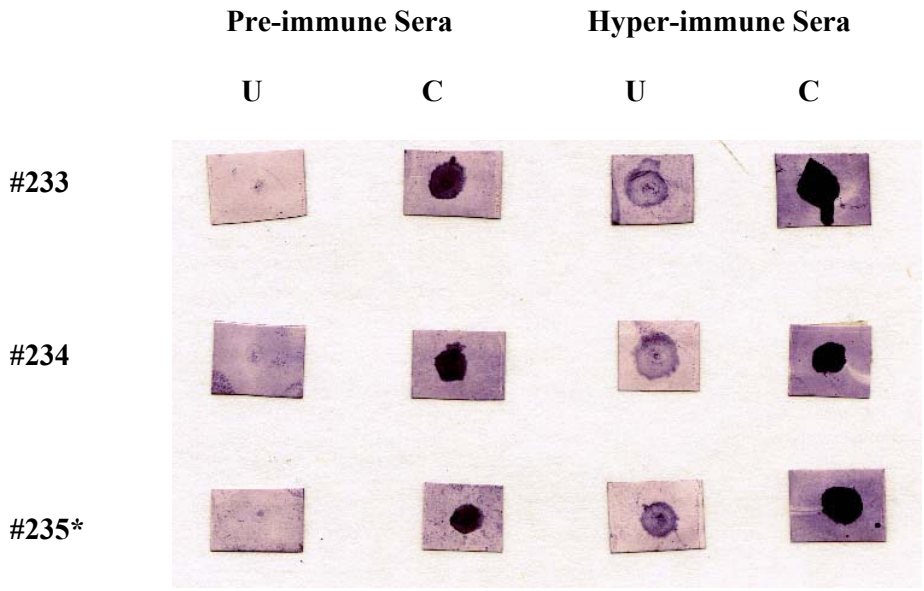


Figure 15: Dot-Blot of 3 rabbit sera after immunization with conjugated peptide. Pre and Post immune sera reactivity to peptide were compared with #235 showing the most-obvious increase to unconjugated peptide suggesting a specific antibody reaction. U: unconjugated peptide dotted on membrane; C: Conjugated peptide dotted on membrane.* #235 showing darkest response to Dot Blot for unconjugated peptide.

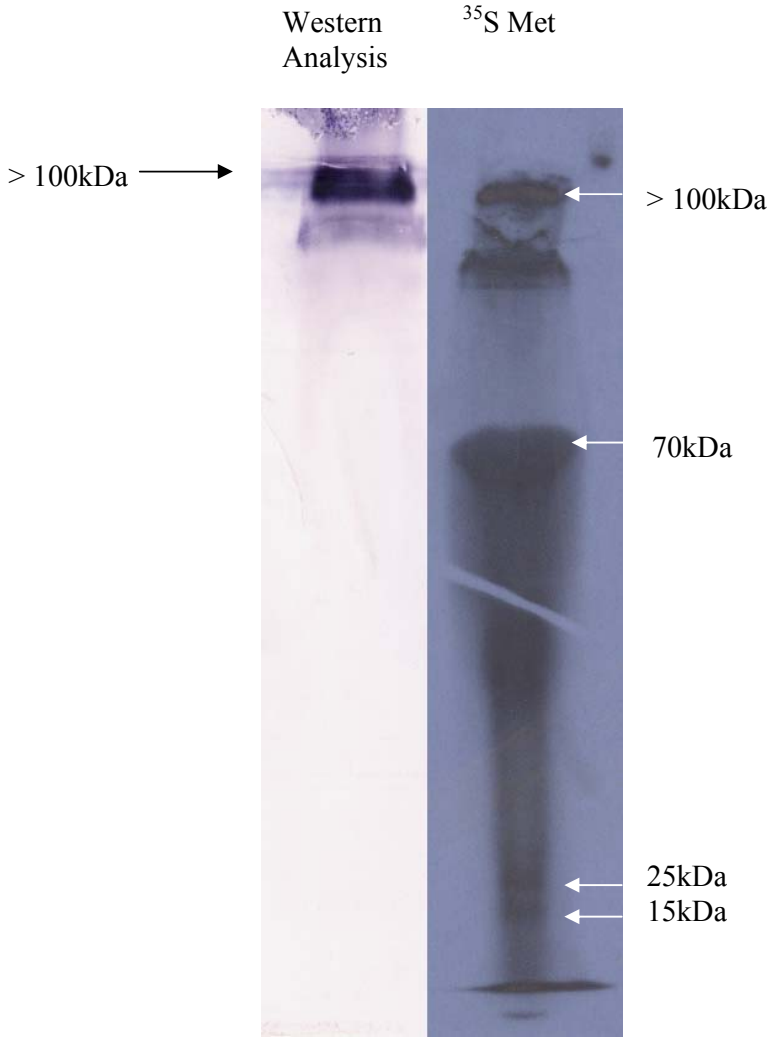


Figure 16A: Polyclonal antibody recognition of high molecular weight MOST-1 protein in TNT experiments. Western analysis using affinity purified antibody showed specific recognition of high molecular weight MOST-1 despite presence of other recombinant species as shown in X-ray film.

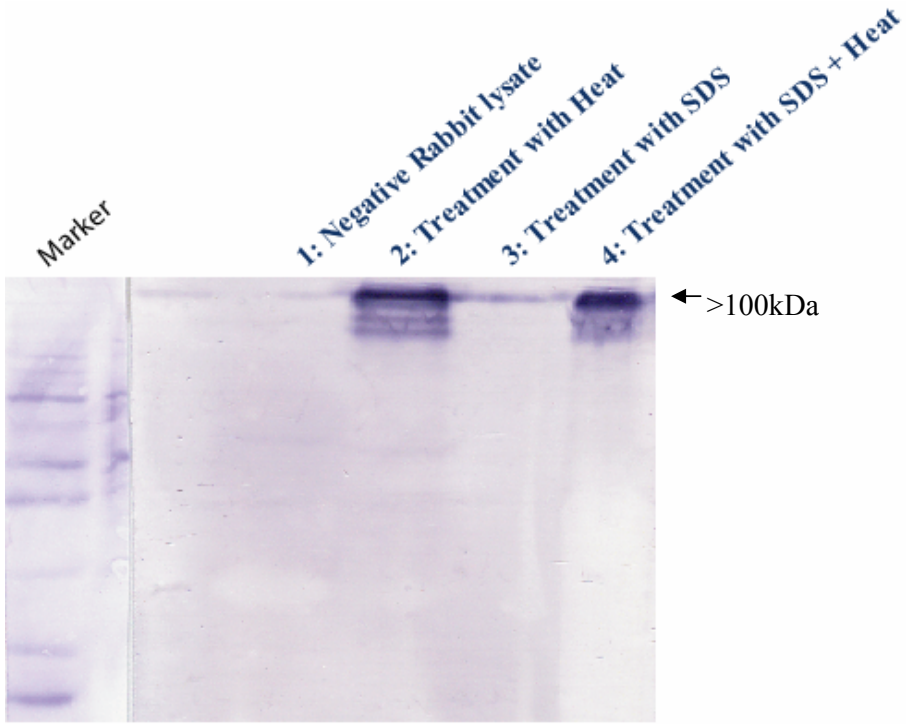


Figure 16B: Differential Treatment of TNT expressed recombinant MOST-1 protein in non-reducing conditions. Western analysis with antibody shows specific recognition of high molecular weight species which is stable to SDS after heat treatment.

4.9. Subcellular localization of MOST-1

Figure 17 shows the cytoplasmic stippling of MOST-1 (stained green) and with respect to the nucleus (stained red with PI) in 4 different cell lines, coinciding with previous observation the MOST-1 is most probably in aggregates in cells. There is a visually higher MOST-1 expression level in corresponding cancer cell lines as compared to their normal-like counterparts suggesting that overexpression of MOST-1 in cancer cells may be a cause of its aggregation.

Western Blot Analysis showed MOST-1 is found predominantly in the microsomal fractions as shown in Figure 18 shows the presence of 2 different size endogenous proteins in MOLT-4 cell lysate with the predominant 100kDa fragment and a lower 70kDa protein present in the microsomal and nuclear fractions.

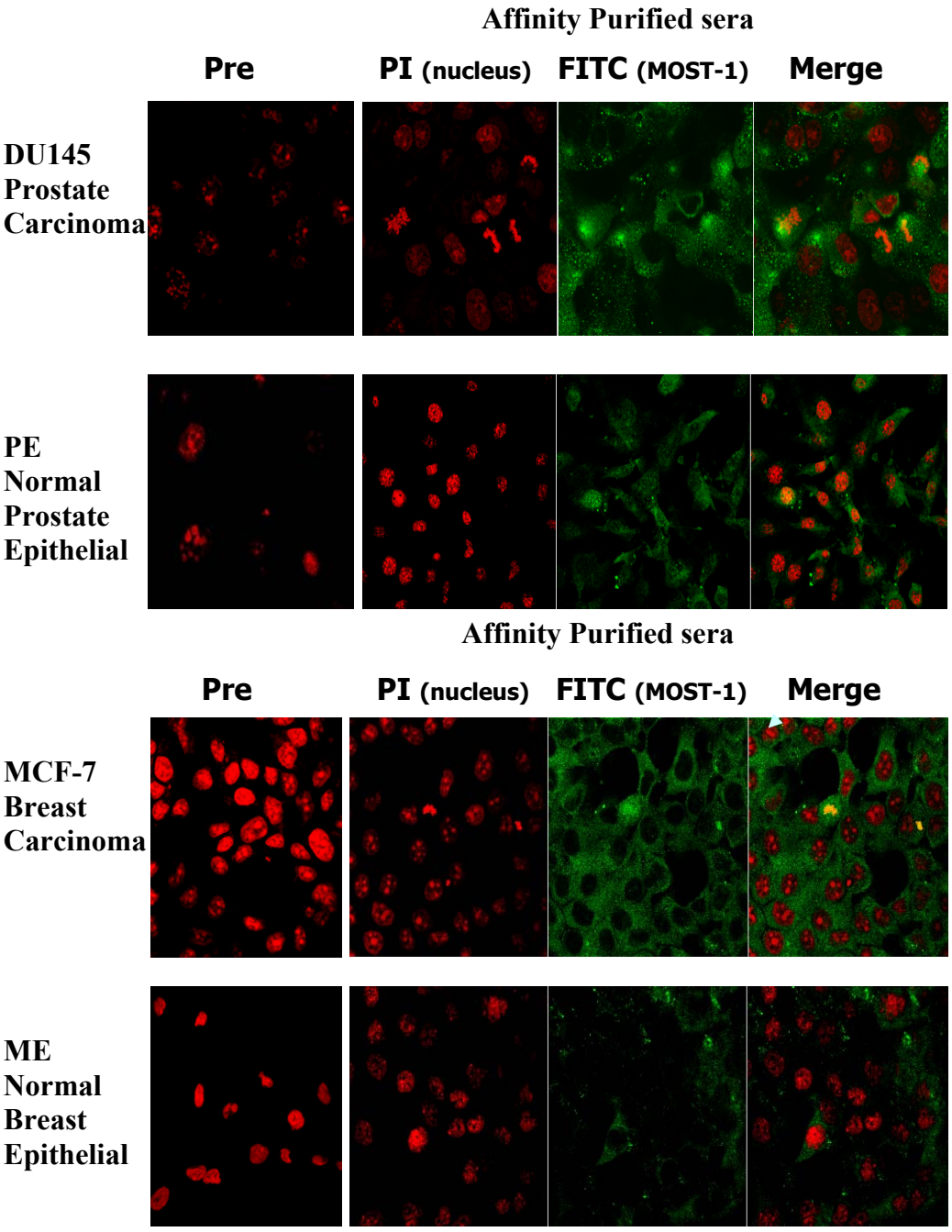


Figure 17: Confocal Microscopy of MOST-1 in various cell lines of breast and prostate origin. MOST-1 appears in cytoplasmic stippling with focal points in dividing cells. There is also a visual increase of MOST-1 in MCF7 breast carcinoma compared to normal breast epithelial cells.

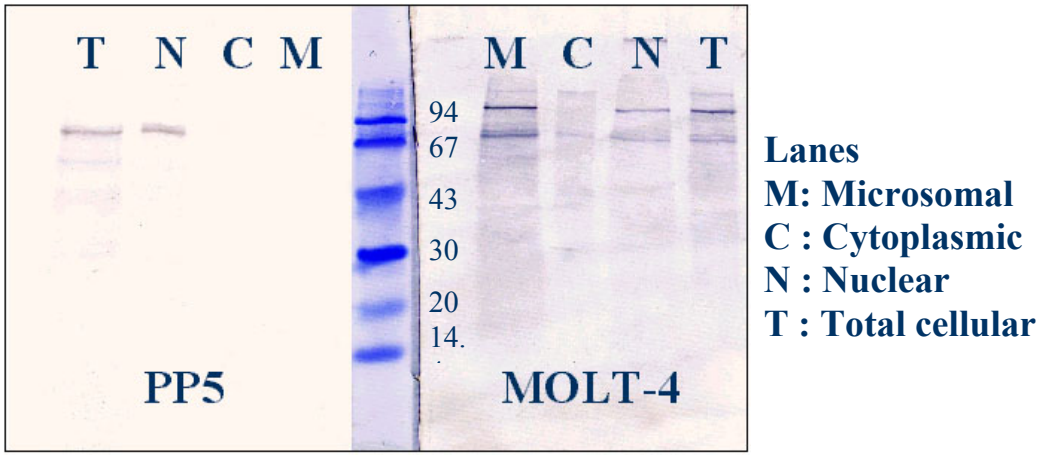


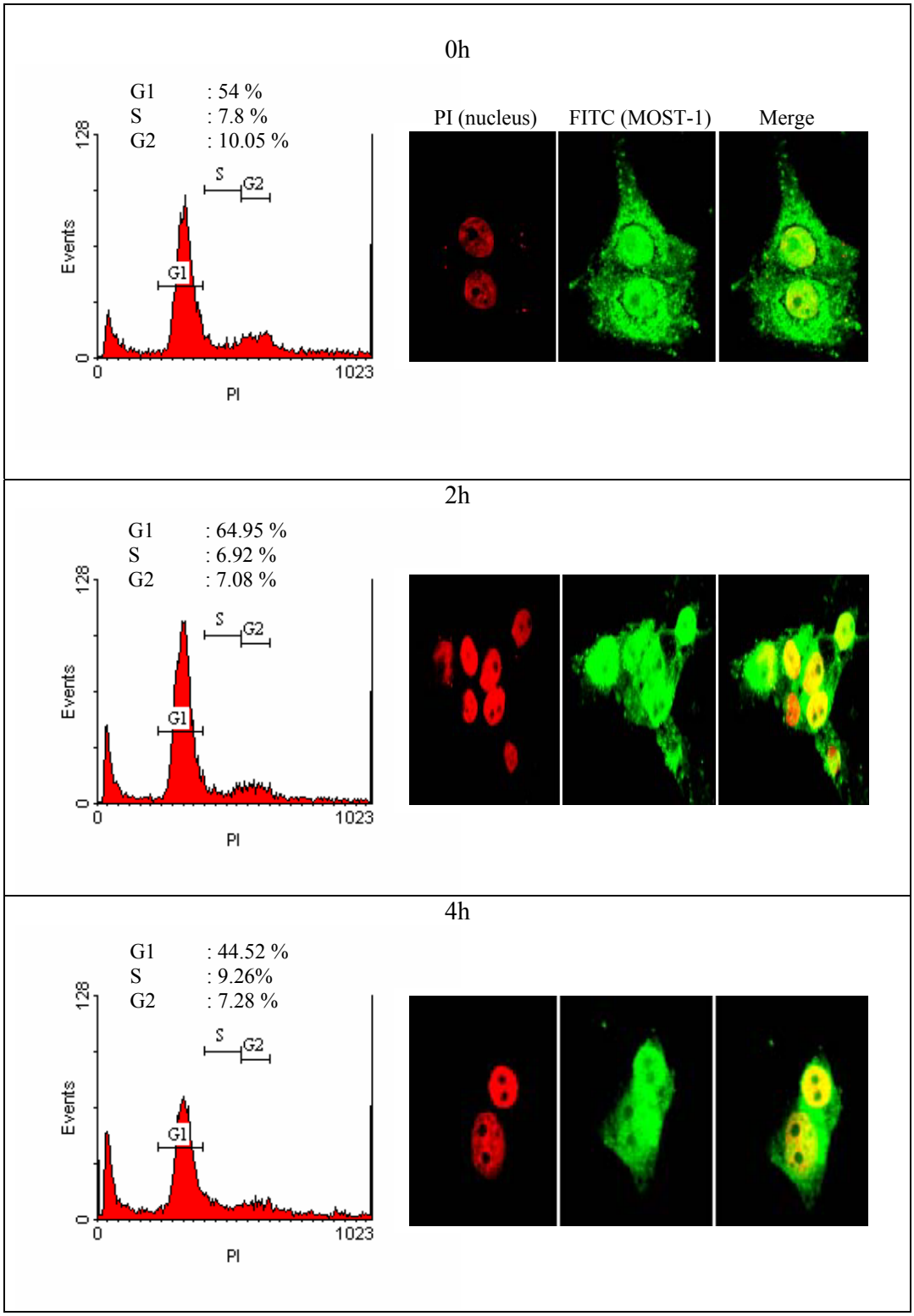
Figure 18: MOST-1 cellular localization studies in PP5 and MOLT-4 cell lines. MOST-1 shows different localization mainly in the nucleus and both microsomal and nuclear fractions respectively. There is also the presence of 2 species, the 70 kDa protein and high molecular weight protein (>100kDa).

4.10. Cell Synchronization Studies

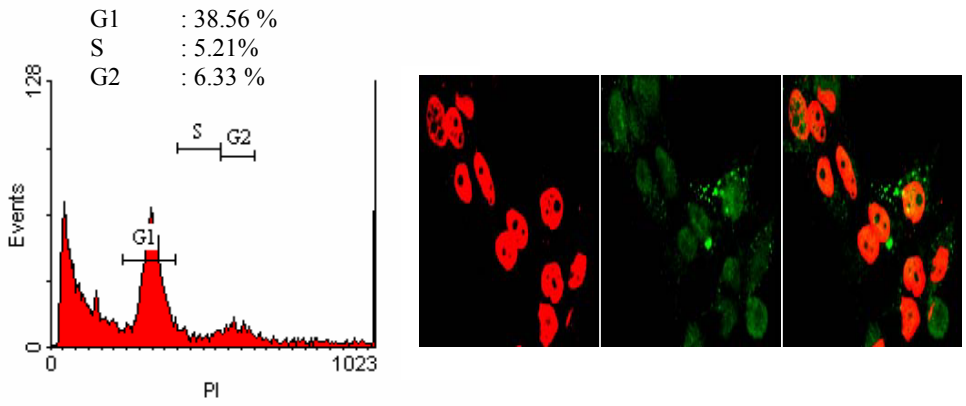
As confocal microscopy showed localization of MOST-1 to a region during cell division, cell synchronization studies were design to investigate the association of MOST-1 to cell cycle. Mimosine has been shown to arrest cells in G1 stage prior to DNA synthesis. Removal of mimosine from media will result in reversible cell cycle progression. To type the population of cells undergoing cell division, PI staining of fixed cells was done. In brief, G2 cells will contain 2 times as much G1 PI due to increase DNA in cells and S phase cells will have an intermediate value between the two. In addition, confocal microscopy was done in parallel to visualize MOST-1 expression. As shown in figure 19, MOST-1 has no clear relation with DNA content of the cells. Confocal microscopy at time intervals however suggest MOST-1 expression is reduced for a 'window' but this timing differs between cancer cells and normal cells as shown in Table 8. Cells appear to be staining more intensely as compared to Figure 17 suggesting an increase in MOST-1 expression level. Closer inspection of the cells shows MOST-1 in vesicle like structures and also in the nucleus. This suggests a role in cell cycle.

Figure 19: Cell Synchronization Experiments

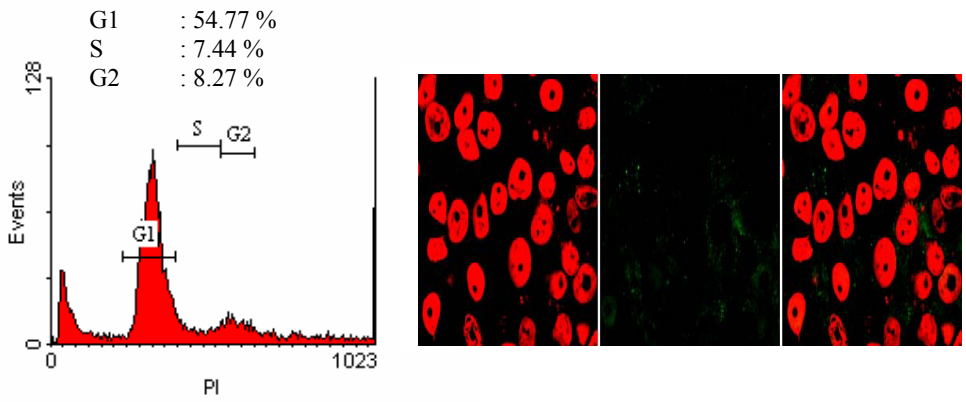
Figure 19A: MCF 7 cells synchronization



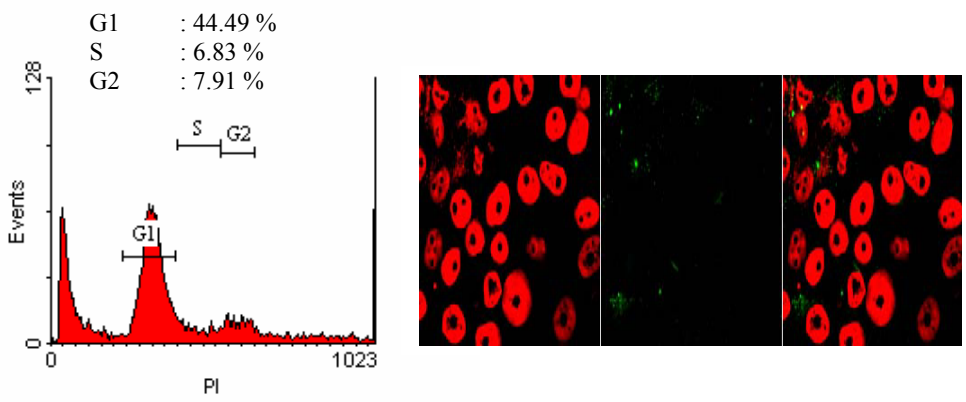
6h


G1 : 38.56 %
S : 5.21%
G2 : 6.33 %
128
0
Events
S
G2
G1
0
1023
PI

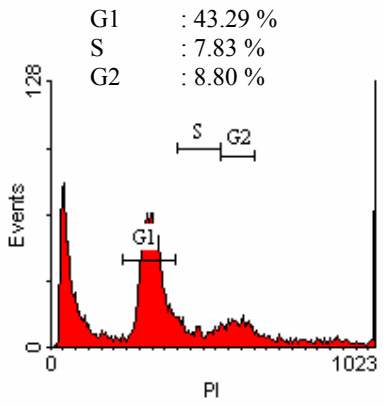
8h

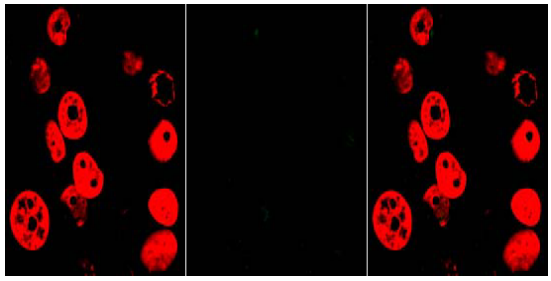

G1 : 54.77 %
S : 7.44 %
G2 : 8.27 %
128
0
Events
S
G2
G1
0
1023
PI

10h

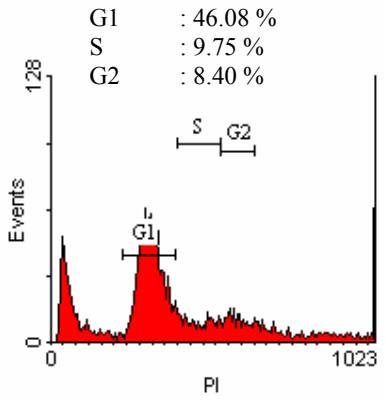

G1 : 44.49 %
S : 6.83 %
G2 : 7.91 %
128
0
Events
S
G2
G1
0
1023
PI

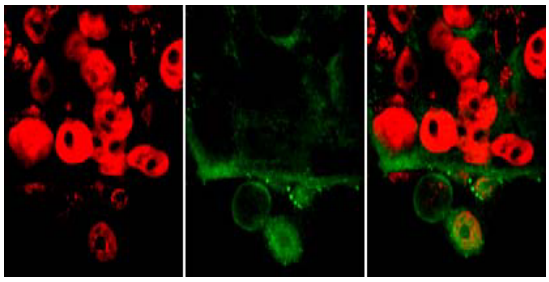
12h


G1 : 43.29 %
S : 7.83 %
G2 : 8.80 %
128
Events
0
S
G2
M1
G1
0
1023
PI

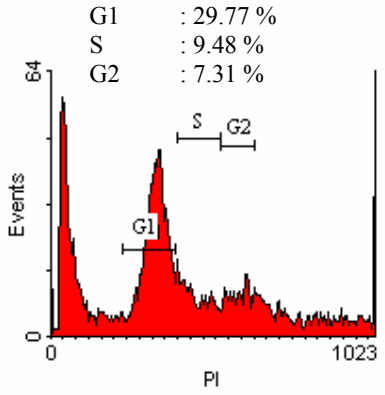


36h


G1 : 46.08 %
S : 9.75 %
G2 : 8.40 %
128
Events
0
S
G2
M1
G1
0
1023
PI



48h


G1 : 29.77 %
S : 9.48 %
G2 : 7.31 %
64
Events
0
S
G2
M1
G1
0
1023
PI

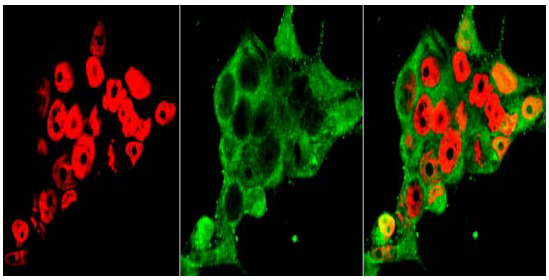
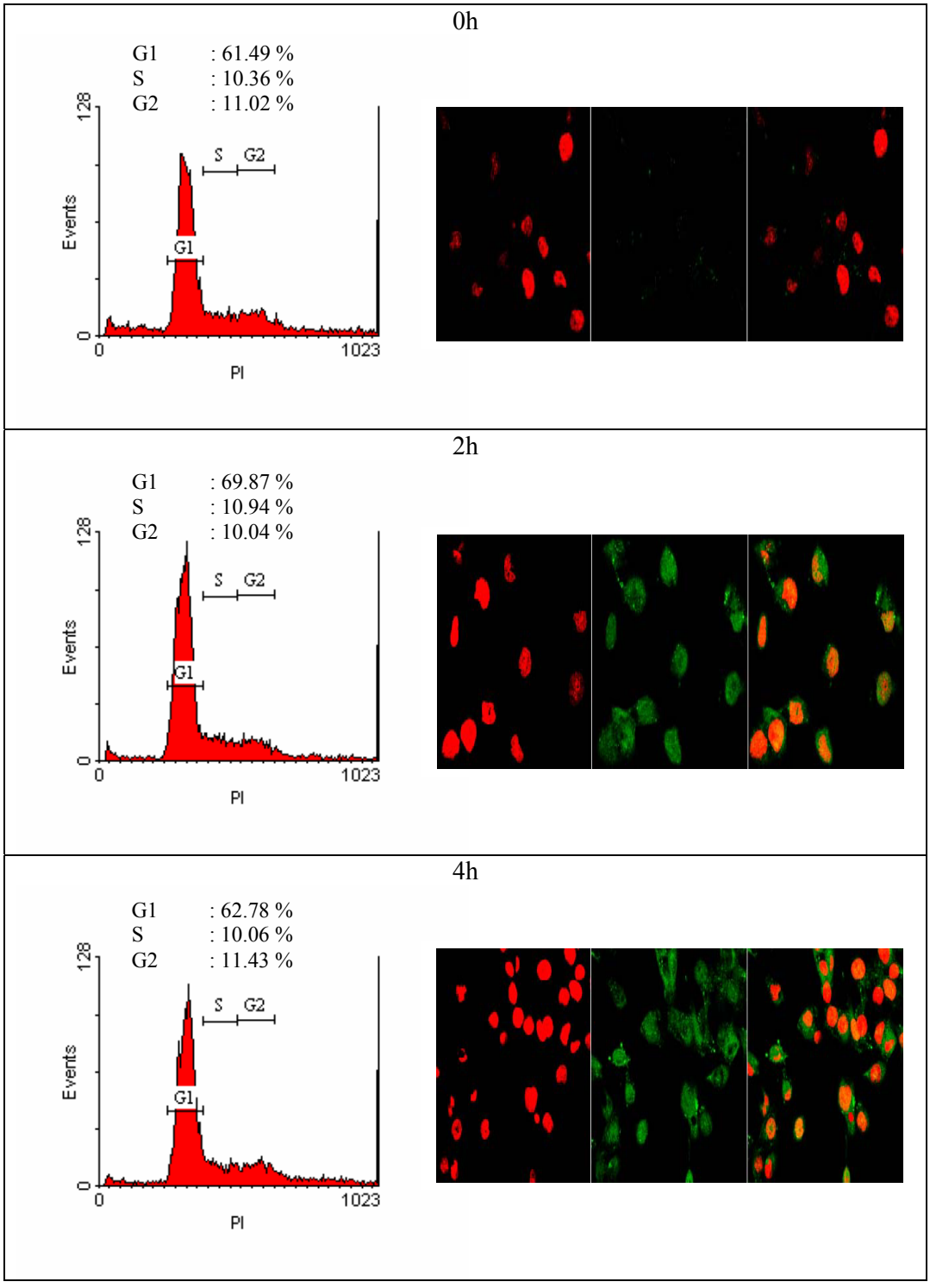
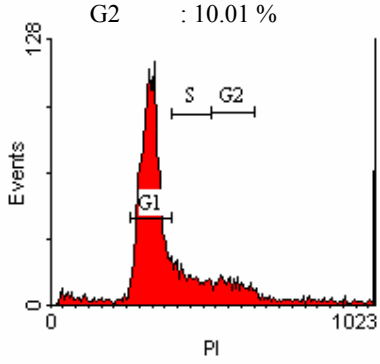


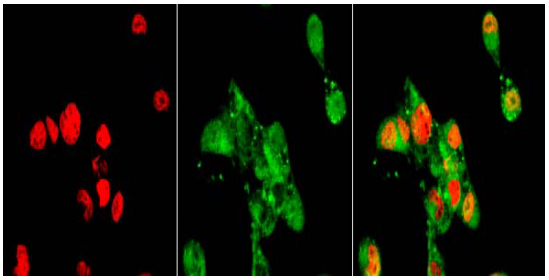
Figure 19B: Mammary normal cells



6h

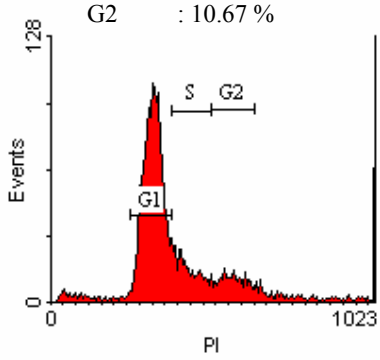
G1 : 67.81 %
S : 12.70 %
G2 : 10.01 %

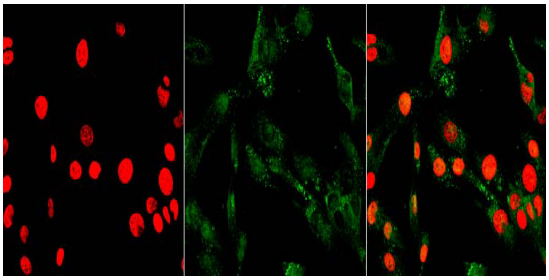




8h

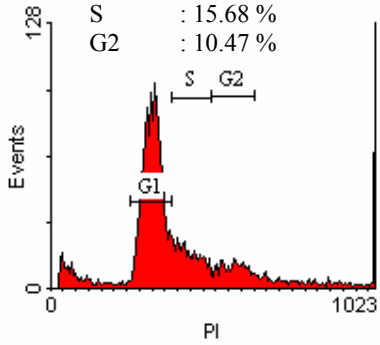
G1 : 65.51 %
S : 15.00 %
G2 : 10.67 %

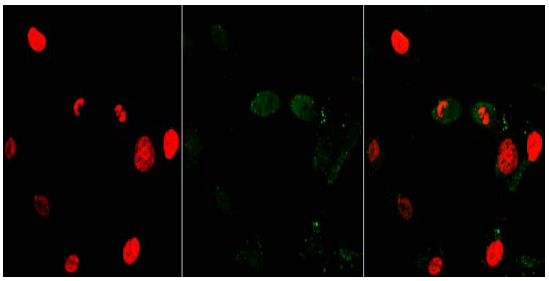




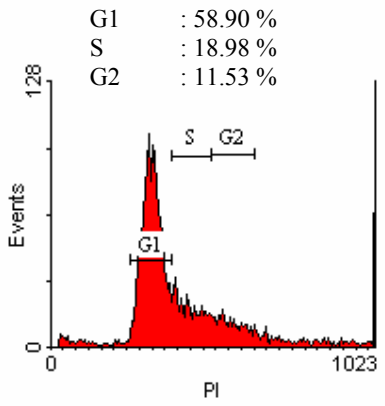
10h

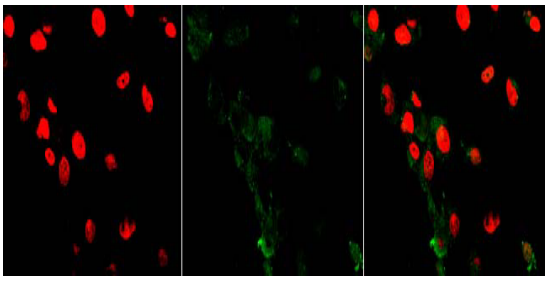
G1 : 57.16 %
S : 15.68 %
G2 : 10.47 %



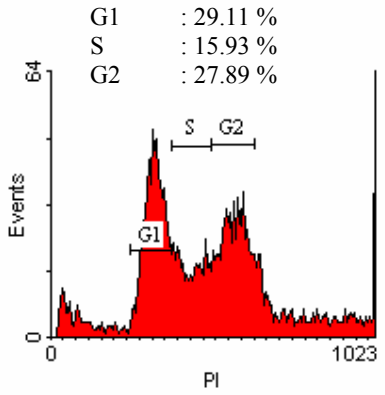


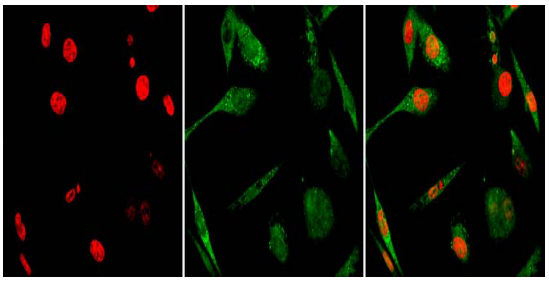
12h


G1 : 58.90 %
S : 18.98 %
G2 : 11.53 %
128
Events
0
S G2
G1
0
1023
PI

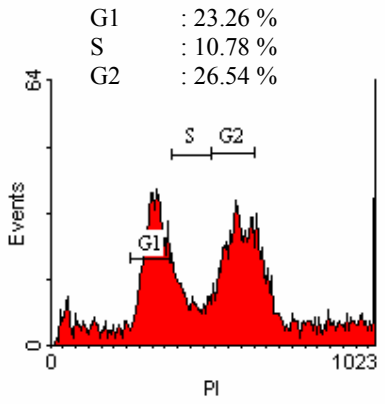


36h


G1 : 29.11 %
S : 15.93 %
G2 : 27.89 %
64
Events
0
S G2
G1
0
1023
PI



48h


G1 : 23.26 %
S : 10.78 %
G2 : 26.54 %
64
Events
0
S G2
G1
0
1023
PI

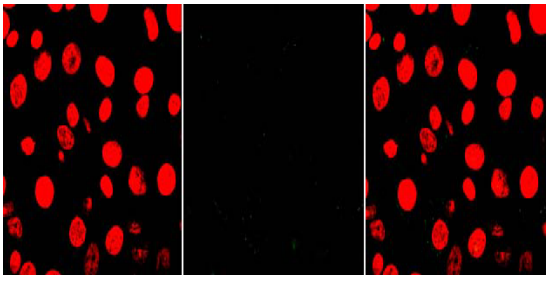


TABLE 9: Summary of cell synchronization comparison of MCF7 and normal mammary cell lines vs. MOST-1 expression levels.

Time (h)	Normal Mammary Cells		Cancer Breast MCF7 Cells	
	Expression Level[a]	Localization[b]	Expression Level[a]	Localization[b]
0	**-**	**-**	**+++**	**Punctate C+ N**
2	**++**	**C Stippling + N**	**+++**	**Punctate C+ N**
4	**++**	**C Stippling + N**	**+++**	**Punctate C+ N**
6	**+++**	**Punctate C+ N**	**++**	**Punctate C+ N**
8	**++**	**C Stippling + N**	**+**	**C Stippling**
10	**+**	**C Stippling**	**+**	**C Stippling**
12	**+**	**C Stippling**	**-**	**-**
36	**++**	**C Stippling + N**	**+**	**Punctate C+ N**
48	**-**	**-**	**++**	**Punctate C+ N**

[a]Expression levels of MOST-1

-: negligible; +: faint; ++: visible expression; +++: high expression.

[b]Localization of MOST-1

punctuate: large round vesicular pattern; stippling: small round vesicular pattern

C: cytoplasmic; N: nuclear

4.11. Yeast two hybrid

A total of 36 clones were obtained and these were subjected to verification via colony lift assays and PCR screening shown in Figure 20. Clones which contain both plasmids and positive for colony lift assays were selected and sequenced. Only clones which are in frame with the tag and expressed a known protein were selected. Of these, overlapping clones were aligned and 7 putative interactors were identified. The interactors were namely creatine kinase(overlapping clones coincide between nt1192-1555, end part of ORF), ferritin (overlapping clones coincide between nt4-219, encompassing the start part of ORF), peripheral benzodiazepine receptor (interaction region between nt156-373, within the ORF), immunoglobulin **C**(mu) and C(delta) heavy chain genes (interaction region of nt15191-15331, part of ORF), SNC73 protein (interaction region of nt1047-1416, within ORF), Gardner feline sarcoma v-FGR oncogene (interaction region of nt1327-1547, within ORF) and telethonin (interaction region of nt336-555, within ORF). The interactors amino acid sequence were aligned (after elimination of vector and tag sequences) for elucidation of potential pattern formation. No obvious pattern could be observe but there seem to be a C – X(>40) – P – X(4-19) – P consensus as shown in figure 21. Table 10 shows the candidate interactors and their putative localization and function.

Co-immunoprecipitation confirmation

Co-immunoprecipitation was carried out with 3 of the 8 interactors, namely human immunoglobulin C (mu) and C (delta) heavy chain, peripheral-type benzodiazepine receptor and ferritin light chain. Pull downs with anti-HA and anti-myc separately were done and detection with tagged antibodies were used to confirm the above 3 interactions. Figure 22A showed the single expression of interactors expression prior to co-immunoprecipitation while figure 22B showed the respective interactors with MOST-1 after pull down with anti-HA and detected with anti-myc and vice versa for figure 22C.

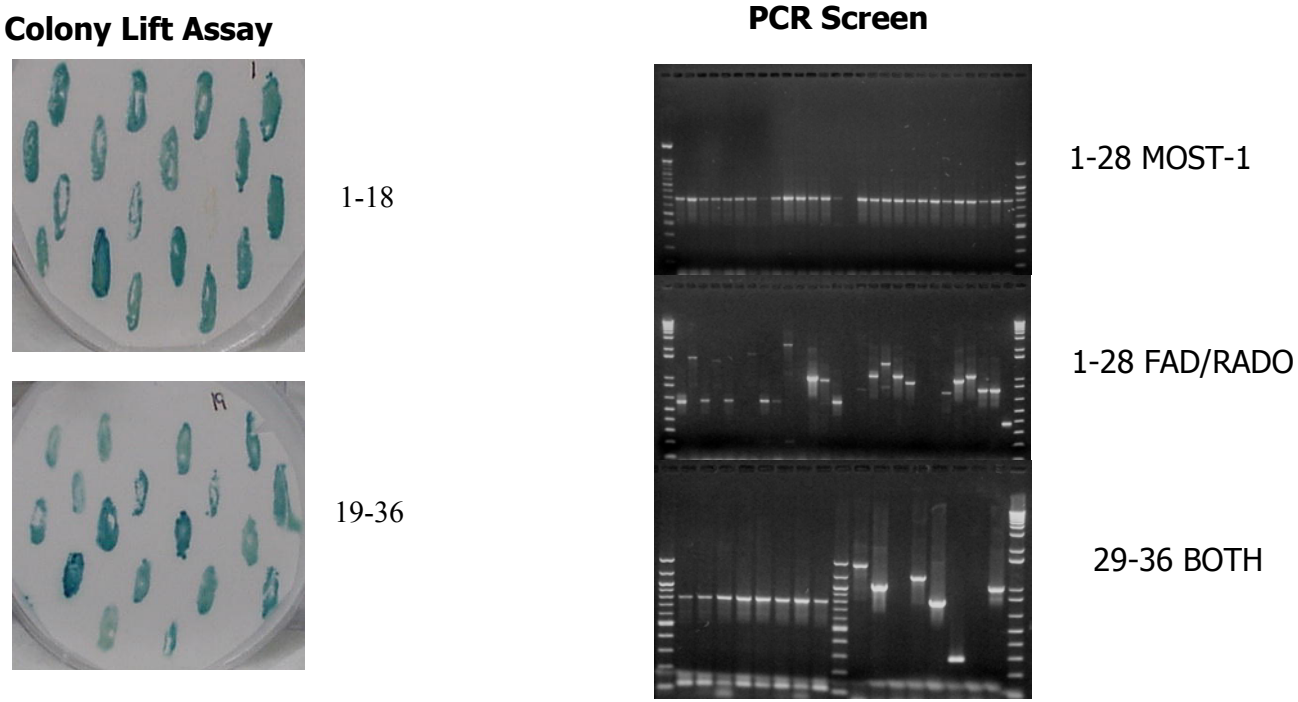


Figure 20: Screening of Y2H interactors. Colony lift assay shows positive clones as dark blue colonies after Buffer Z addition; 25 positive clones were selected out of 36 interactors picked. PCR screen showing presence of both AD and BD plasmid in hybrids was done in which 19 positive clones out of 36 were selected. Overlapping 19 positive clones from the 2 assays were then sequenced and used for BLASTn analysis for identification of in frame proteins.

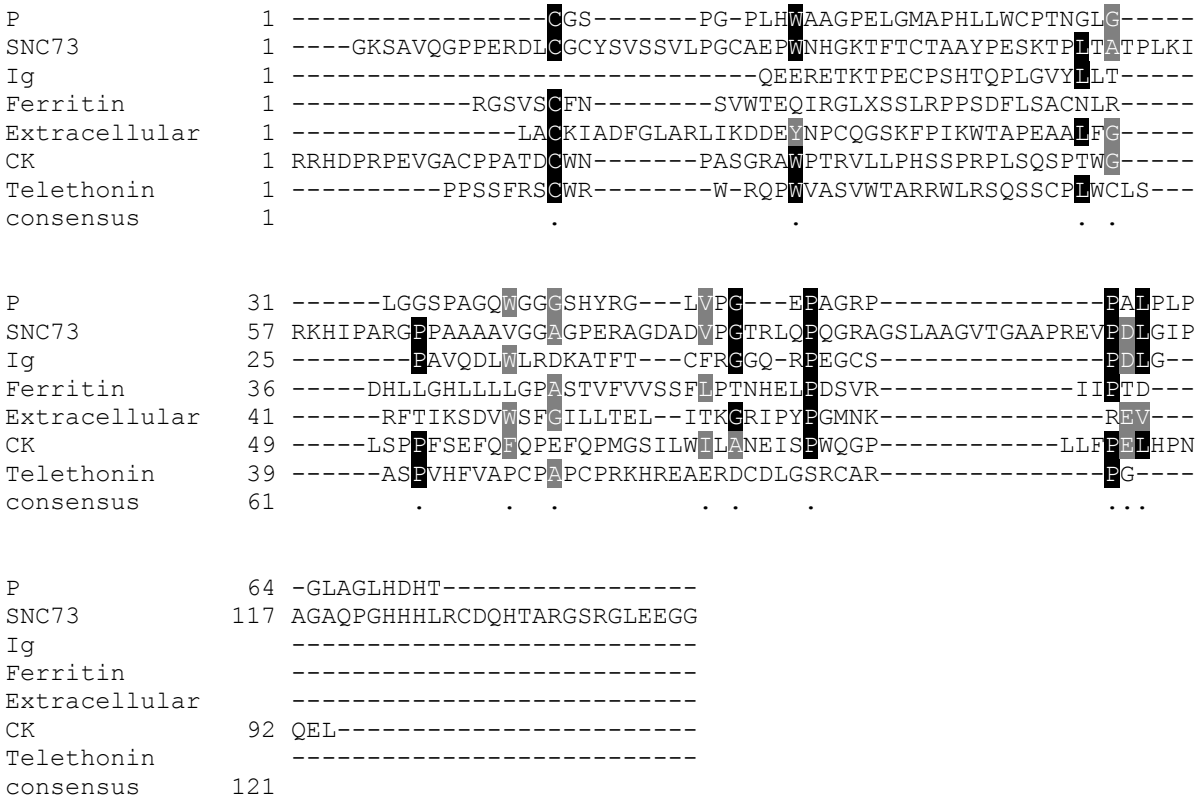


Figure 21: Alignment of Y2H interactors for consensus sequence search. Positive in frame interactors were aligned using Clustalw and consensus sequence was searched to identify possible motif for interaction with MOST-1. Solid Box: Identical amino acid, Shaded box: similar group amino acid.

Table 10: Putative interactors – their localization and function.

Putative Interactors	Cellular Location	Reference
Creatine Kinase (7 clones)	Cytosol and mitochondria	Shen et al, 2002
Ferritin (4 clones)	Discrete pools in cytosol or extracellular	Parthasarthy et al, 2002
Benzodiazepine receptor (Peripheral)	Mitochondria	Chaki et al, 1999
Immunogluobulin (mu and delta) heavy chain regions	Endoplasmic reticulum before secreted	Mielenz et al, 2003
Gardner feline sarcoma viral homologue (v-FGR oncogene)	Membrane bound	Baker et al, 1998
Telethonin (titin-cap)	Z-line of sarcoma	Zou et al, 2003
SNC73 protein	Suggest to be secreted	Hu et al, 2003

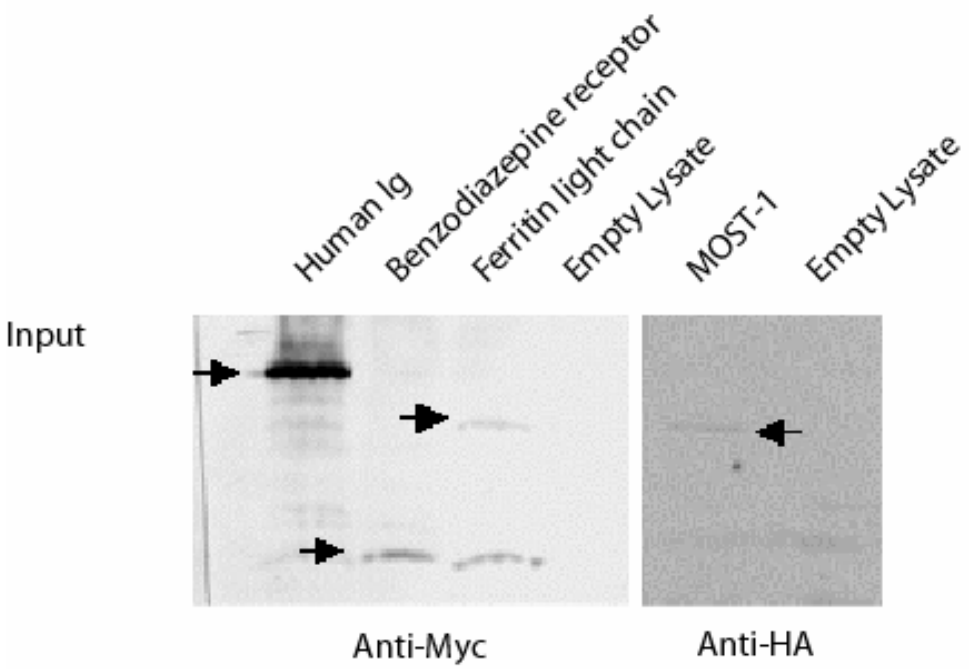


Figure 22A: Single expression of interactors and MOST-1 protein prior to co-immunoprecipitation.

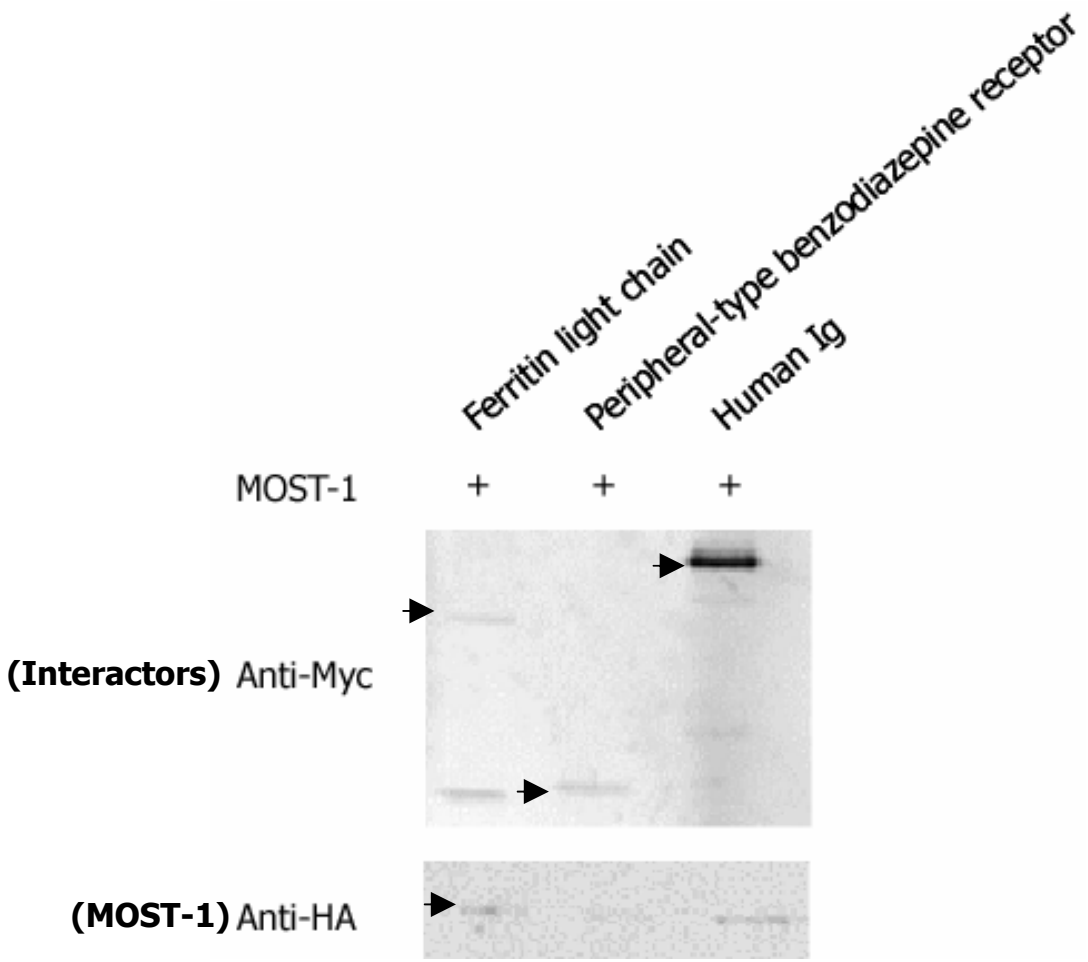


Figure 22B: IP with anti-myc.

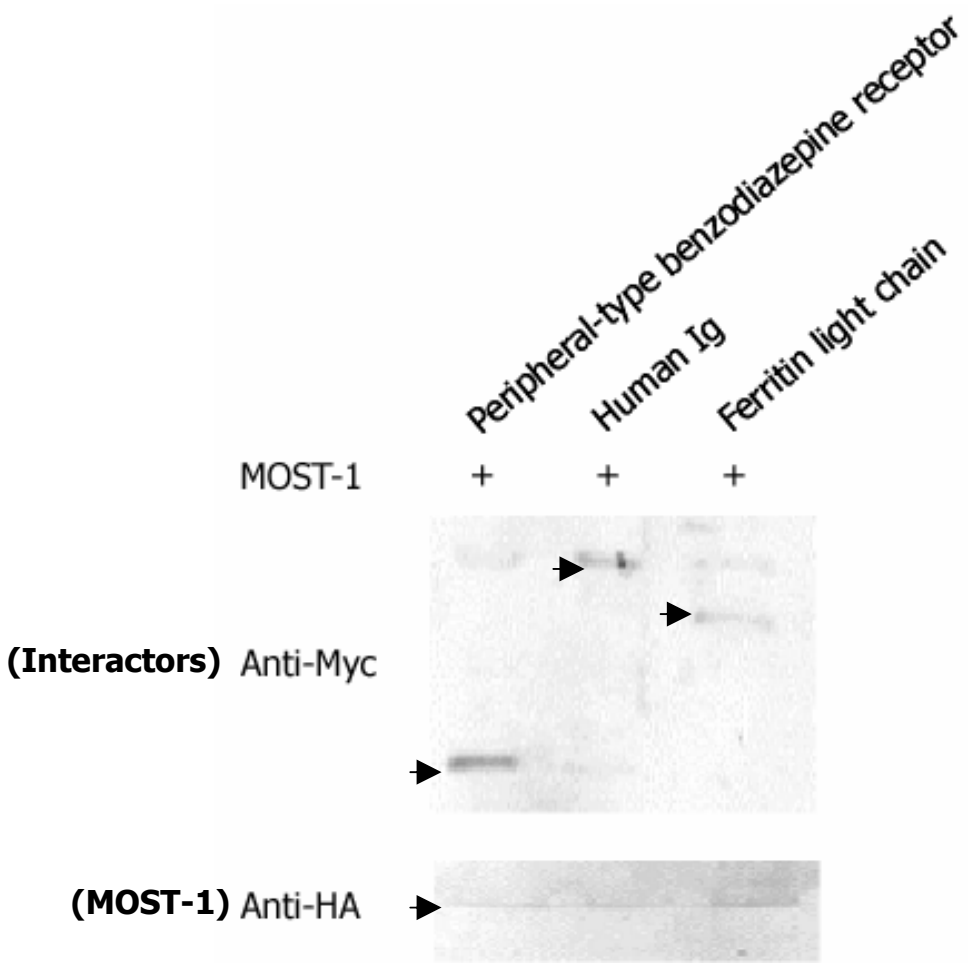


Figure 22C: IP with anti-HA.

4.12. Overexpression and RNA interference studies

RT-PCR was chosen for detection of MOST-1 expression as analysis of cDNA showed 3 RNA degradation signals suggesting a control at transcript level rather than protein level. After mRNA extraction was done, RT-PCR was carried in the log phase of amplification as previously determined for biopsies studies. Figure 23 shows the RT-PCR results of MOST-1 after treatment in the various cells. RNAi was more pronounced in normal cell lines but there is a marked knock down of MOST-1 endogenous transcript levels in the cancer cell lines. Overexpression is not very overt despite the use of vectors with CMV promoters. Since overexpression was done with transient transfection of MOST-1, the low expression level could be due to the control of MOST-1 expression in cells as cotransfection of luciferase showed positive transfection. Since RNAi experiments worked, cell proliferation was measured using BrdU assay while apoptosis was quantitate using TUNEL assay on cells transfected with RNAi. Results are then tabulated in Graph 3.

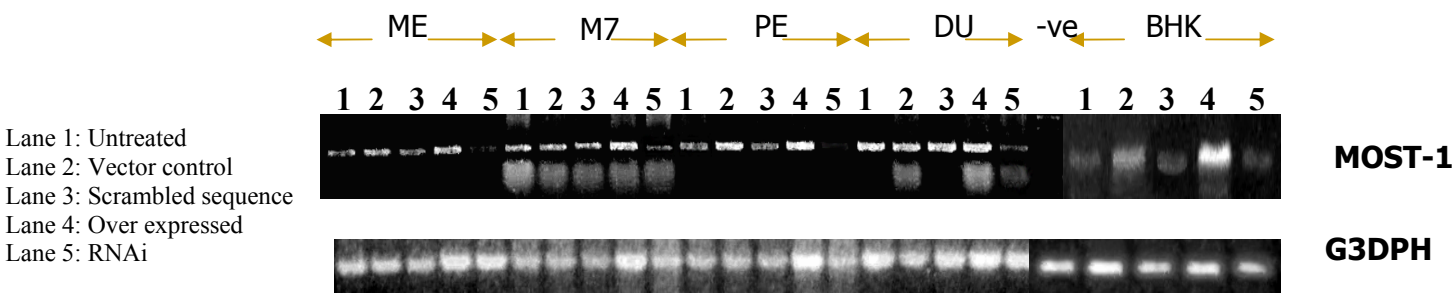
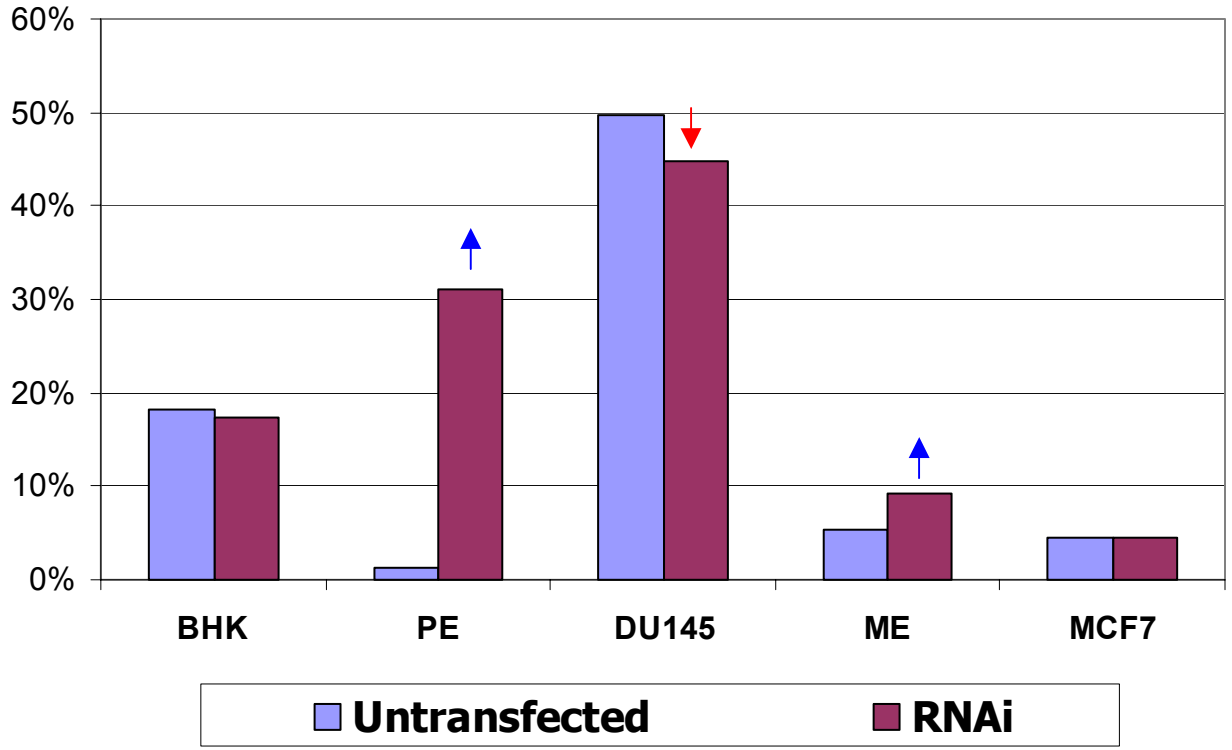
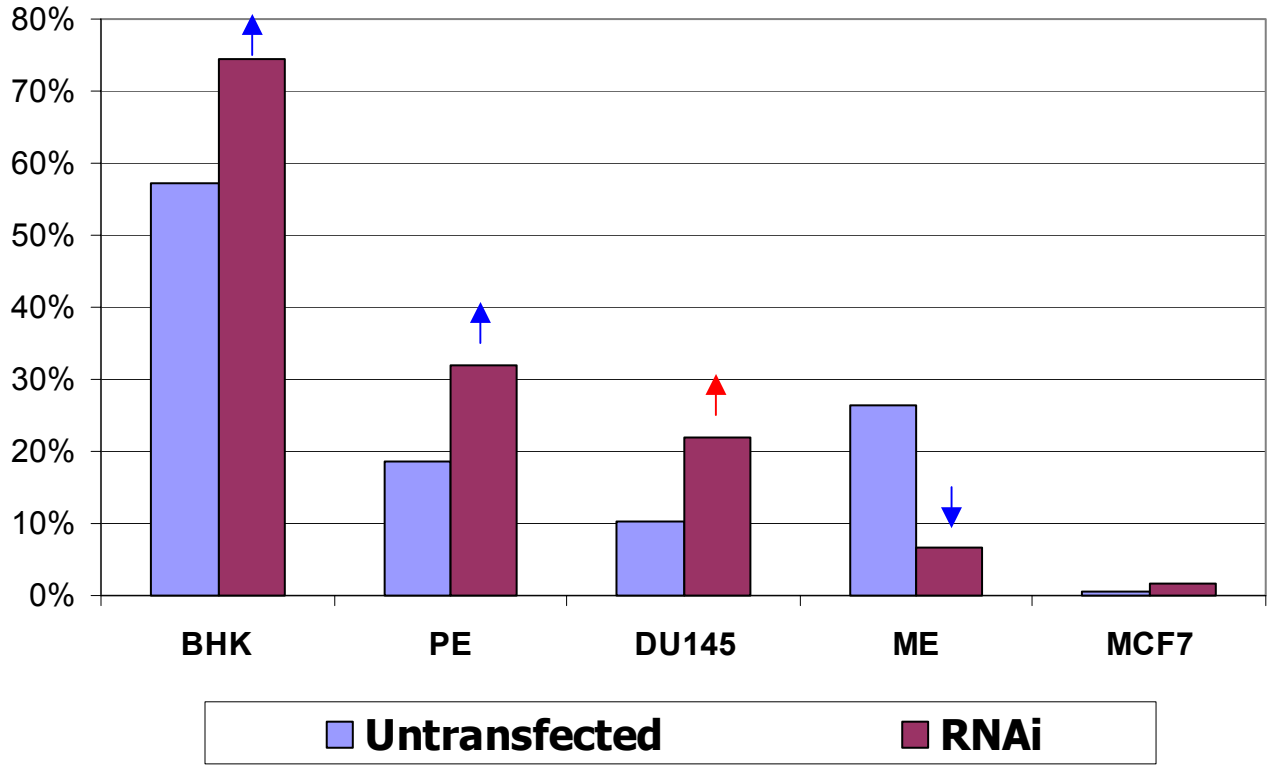


Figure 23: RT-PCR analysis of various cell lines subjected to overexpression and RNAi experiments. Overexpression (lane 4) for all was not overt when compared to vector control and untreated samples (lanes 1 and 2) for densitometric analysis. RNAi treatment (lane 5) was more pronounced compared to scrambled sequence and untreated samples (lanes 1 and 3) suggesting specific knock down due to duplex RNA used.

Graph 3: MOST-1 RNAi effect on cell proliferation and apoptosis



Graph 3A: Mean cell proliferation of RNAi treated cells by BrdU assay. Knock down of RNAi showed an increase in prostate (PE) and mammary (ME) epithelial cells and a decrease for prostate cancer (DU145) cells. There is no apparent effect on BHK and MCF7 cells.



Graph 3B: Mean cell apoptosis of RNAi treated cells measured using TUNEL assay. Knock down of RNAi showed an increase in BHK, prostate epithelial and carcinoma cells and a decrease for mammary epithelial cells. There is no apparent effect on MCF7 cells.

CHAPTER 5. DISCUSSION

Strategy and Isolation of MOST-1

Initial study by Couturier et al showed that DNA of HPV are found integrated in the cell genome in most invasive genital carcinomas as compared to intraepithelial neoplasia where HPV DNA is detected most commonly as episomal molecules. Subsequently, it was found that greater than 99% of all cervical tumors contain HPV DNA and studies have shown that this integration is non random and occurs at loci containing human common fragile sites (CFS). CFSs are described as specific loci that form gaps or breaks in metaphase chromosomes from cells that have been challenged with chemicals that induce replicative stress. There are 89 CFSs in the human genome and CFSs are present in all primates, lower mammals and possibly yeast. The mechanism of CFSs fragility is currently under investigation (Ferber MJ et al, 2003). Feber et al, 2003 reported that 30% of all HPV 18 integration occurred within 8q24 near the *c-myc* proto-oncogene. The *c*-myc locus is coincidently next to the chromosomal location of *MOST-1*, 8q24.2. They have also shown that multiple genes are interrupted by the integration many of which include tumor suppressors, proto-oncogenes and genes maintaining DNA integrity. Integration of HPV genome frequently results in the functional elimination of the viral repressor and remodeling of the E6/E7 promoter, which then lead to the continued expression of E6 and E7 oncoproteins. E6 and E7 oncoproteins of HPV have been shown to direct cell cycle progression and play a major role in HPV-induced carcinogenesis by interfering with the host cell regulatory proteins (Munger et al, 1989). Overexpression of E6 and E7 is insufficient to immortalize primary cells, thus additional alterations such as activating ras mutations are required to transform them (Ferber et al, 2003). In addition, E6 and E7 interact with the tumor suppressors p53 and pRB and thus provide a mechanism

for the cells to accumulate genomic damage (Ferber MJ et al, 2003). Loss of tumor suppressor function, acquisition of unrestrained replicative ability and eventual loss of genomic stability are hallmarks of cancer. Recent reports illustrate that reduction of E6 and E7 by RNAi, in cell line (HeLa) transformed with HPV18, induces senescence. This correlates with previous studies that E6 and E7 immortalize cells. The authors showing that RNAi of E6 and E7 inhibited cellular DNA synthesis and induced morphological and biochemical changes characteristic of cellular senescence further lends support to the above (Hall et al, 2003). This study attempt to isolate genes using HPV primers which have a role in cell cycle from MOLT 4. MOLT-4 is a cancer cell line which was not transformed with HPV and hence do not have HPV genome integration thus reducing the background amplification of HPV genes with the consensus primers. The primers derived from HPV 18 and HPV 11 E6 genes are found to target a sequence in the middle of the ORF of a novel EST, *MOST-1*. Should this be the site of integration, HPV integration would disrupt the gene and prevent its function and thus confirming the studies cited. Similarly, it could also mean that *MOST-1* might function in cell cycle or its regulation. Characterization of this gene was undertaken in the hope of elucidating its function.

***MOST-1* Gene**

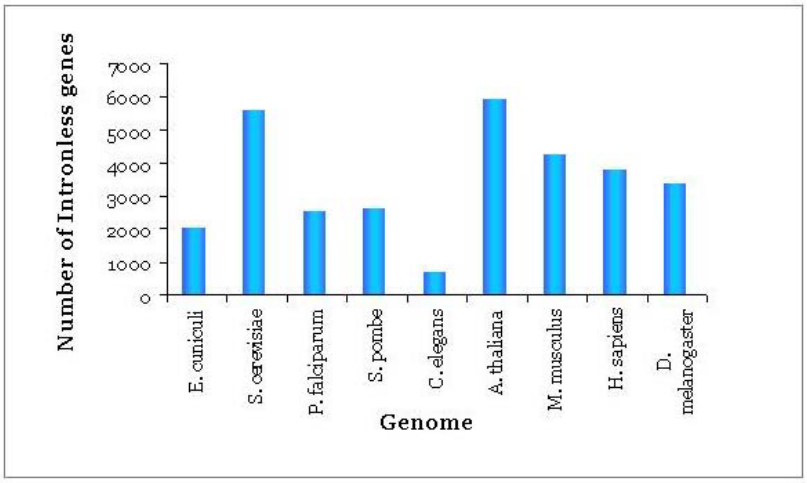
The *MOST-1* gene was found to be intronless. New identification of intronless genes in vertebrates makes exceptions to the rules that all eukaryote mRNA undergoes splicing. Indeed, this property have been predicted to occur in ~5% of the genes (Gentles AJ et al, 1999). Graph 4 shows the prevalence of intronless genes in humans to be approximately more than 10% of known genes in GenBank currently (4000 of the predicted 30, 000 genes). Of the 4000

genes found, 8% of kinases, 44% of helix-loop-helix containing proteins and 35% of cell surface receptors. These genes are commonly found in cell cycle regulation. Histones, cytokines and G-proteins are the main families of intronless genes. Histones are small proteins with an average size of 130 amino acids and have a high degree of protein sequence conservation. G-proteins are thought to be either derived from a single intronless common progenitor or introns are subsequently lost. Human G-protein-coupled receptors, majority related to brain function, function to mediate ligand-induced signaling (Gentles et al, 1999). Some other examples of functional intronless genes are CDR1 (cerebella-disease-related antigen), SOX3, α-interferon β-adrenergic receptor and JUN proto-oncogene. Many of which have been identified to be also related to brain function (Redolfi et al, 1998).

It has been suggested that intronless genes not requiring post-transcriptional splicing, might be more efficiently transcribed with greater abundance and have a higher rate of protein expression. This in turn results in genes which are not liable to differential and aberrant splicing and thus resulting in higher transcriptional fidelity (Gentles AJ et al, 1999). The origin or evolution of intronless gene in mammals is still under debate. There are however several suggestions such as having a single intronless common progenitor relatively recent in evolution with insufficient time to allow gain of introns as in the case of dopamine receptors, retroviruses transforming multi-exon genes when they integrate into host genome nearby as in the case of *src* or via transposable elements with the latter two mechanism having a common

preference towards GC-rich sequences (Gentle et al, 1999). All this evidence support *MOST-1* to be a functional gene despite its intronless nature.

Graph 4: Number of intronless genes compared across genomes



Adapted from http://sege.ntu.edu.sg/wester/intronless/statistics.htm

Structural analysis of *MOST-1* cDNA suggest an unstable transcript, with a strong consensus sequence at the putative methionine start codon (Kozak, 1991) of a potential ORF encoding a relatively small protein. Present within the 3′ UTR of *MOST-1*, three AUUUA motifs, which is a highly conserved sequence repeated three or more times in RNAs encoding many short-lived cytokines and proto-oncogenes (Akashi et al., 1994). Stability of mRNA is determine by the number of AUUUA motifs, which is a highly conserved sequence of AU rich region of ~50bp (called ARE) and is repeated three or more times in the 3' untranslated region (3' UTR) of short-lived cytokines and oncogenes. These motifs confer instability. The rank order of stability was 1x ATTTA = 2x ATTTA (no RNA decay at 4h) > 3x ATTTA = 5x ATTTA ($t_{1/2}$ is 4h) > 7x ATTTA ($t_{1/2}$ is 2h) (Akashi et al, 1994). ARE triggers destabilization by ARE-binding protein attaching to AUUUA motif and cause

deadenylation with poly (A) ribonucleus followed by decay with endonucleases. Deadenylation causes loss of poly(A)-binding protein which stabilizes the 3' region. General model of mRNA stability is conferred by inhibiting function of destabilizing sequences (Lewin, 1997). The structural analysis of *MOST-1* indicates that although this gene is efficiently translated, its expression seems to be controlled at the mRNA level. There is also an Alu repeat at the 3' end of ORF of *MOST-1*. Alu repeats are part of the transposable elements that are found in large number of human protein-coding genes (Nekrutenko A et al, 2001). Alu elements are a family of short interspersed elements which contain eight putative donor sites and three acceptor sites and with ~1.4 million Alu elements interspersed throughout the genome (out of which 1.3% occurs in human coding regions), this provides numerous possibilities of formation of alternative transcripts. At the same time, it has been found that Alu insertion occurs once in every 200 human births suggesting that Alu elements could have contributed significantly to divergence between primates and other mammals because Alu elements are exclusively in primates (Nekrutenko A et al, 2001). Alu elements are scattered by 'retrotransposition' which involved three independent steps: transcription of the Alu repetitive element, reverse transcription of Alu RNA and integration of Alu cDNA (Szmulewicz MN et al, 1998). The significance of Alu is still under debate but in all studies, it is agreed that Alu repeats remain one of the major driving force of evolution. In addition, there is a suggestion that variation of retroelement distribution in the human genome is associated with age and proximity to genes being the oldest Alu elements preferentially found in regions of lower GC. The presence of an Alu repeat at the 3' end of the gene together with *MOST-1* intronless property suggest an evolutionary conservation which in turn suggest the possibility of *MOST-1* being involved in fundamental processes such as cell survival.

The generally low abundance of *MOST-1* mRNA in normal adult tissues was reflected by the necessity for reamplification during screening for tissue expression of *MOST-1*, and by the inability to detect *MOST-1* transcripts by Northern blot analysis. It is then tempting to speculate the reason why *MOST-1* is present is such low amounts in tissues suggesting that if high amounts of *MOST-1* maybe toxic to cells, there may be extensive regulation on its protein and mRNA level. These features are characteristic of many transiently expressed genes encoding small polypeptides involved in signaling or metabolic cascades that are being discovered at an accelerating pace.

Chromosomal localization impact on *MOST-1* function

The chromosomal localization of *MOST-1* to chromosome 8q24.2 is noteworthy as this region (and its vicinity) has been shown by many comparative genomic hybridization studies to be amplified mainly in ~50% breast and ~50-85% of prostate cancers, and also in ovarian, testicular, renal, bladder and colorectal tumors (Visakorpi et al., 1995; Forozan et al., 1997; Kuukasjarvi et al., 1997; Nupponen et al., 1998, 1999; De Angelis et al., 1999; Knuutila et al., 1999; Loveday et al., 1999). Another recent study using FISH in pathologic organ-confined prostate cancer showed a clinical significance of 8q24 overrepresentation together with 8p22 loss results in poor prognosis in patients with high-grade, locally advance prostate cancer (Tsuchiya et al., 2002). Although the localization of the *c-myc* proto-oncogene at 8q24.1 has been cited to account for many poorly differentiated prostate cancers, the entire

long arm of chromosome 8 is usually present at an increased copy number, suggesting that other genes may be involved (Visakorpi et al., 1995; Saramaki et al., 2001). For example, a gene that is differentially expressed during prostate cancer progression has been identified on chromosome 8q11 (Chang et al., 1999). Amplification of *c-myc* may not always correlate with 8q amplification in breast and prostate cancers (Nupponen et al. 1998, 1999). Other genes located at the 8q24.2 region are the *PSGA* gene mentioned above, a prostate-specific gene that apparently plays a role in prostate cancer progression (Reiter et al., 1998), and the brain adenylyl cyclase gene *HBAC1* (Stengel et al., 1992). This locus has also been implicated in various diseases such as partial trisomy 8q where patients suffer from psychomotor retardation (Stengel-Rutkowski et al., 1992), and hereditary spastic paraplegia in which the dominant *HSP* gene has been localized to 8q23-q24 (Hedera et al., 1999). Diseases associated with 8q for which the exact disease locus has yet to be discovered include autosomal recessive achromatopsia with defective photoreceptors (Milunsky et al., 1999), retinitis pigmentosa (Inglehearn et al., 1999), and tibial hemimelia in Langer-Giedion syndrome (Stevens and Moore, 1999). A recent study by Dekken et al showed that the gain of distal 8q discriminates between progressors and nonprogressors. These taken together support a new school of thought that amplification of genes in cancer commonly occurs in large amplicons such that several genes are amplified together. These genes are often involved in either similar pathways or systems.

The amplification of the chromosomal locus coinciding with that of *MOST-1* in breast and prostate malignancies prompted the analysis of quantitative PCR experiments to determine *MOST-1* RNA and DNA levels in frozen breast cancers and archival prostate tumors versus normal tissues. Two fifths of breast cancers exhibited *MOST-1* overexpression, concurring with previous comparative genome hybridization studies that revealed amplification of 8q24.2 in these tumors. Compared with lower grade breast cancers, those of grade 3 showed higher T:N ratios of *MOST-1* transcript levels of up to 13.5. Relative to normal prostate specimens, benign prostatic hyperplasia samples displayed an average 1.4-fold increase in *MOST-1* DNA levels, while up to approximately 10-fold amplification was observed in prostatic cancers especially those of higher grades and Gleason scores. In conclusion, aberrations of *MOST-1* expression and copy number appear to be associated with substantial subsets of breast and prostate carcinomas, especially those of higher grade.

MOST-1 Protein

The next step would then proceed to polyclonal antibody production toward *MOST-1* for protein characterization. Prior to antibody production, the putative ORF of *MOST-1* was analyzed and several interesting predicted sites gave insights to protein function and localization. One of which is the N-myristoylation site occurring at Glycine 2. N-myristoylation is a covalent protein modification that can promote the association of proteins with

membranes (De Jonge et al, 2000). Another site is the casein kinase II (otherwise known as protein kinase II) phosphorylation in which protein kinase II has been proposed to undergo rapid modulations in its association with nuclear matrix and nucleosomes in response to mitogenic signals and phosphorylation of proteins depending on the state of genomic activity. This association may influence the apoptotic activity in the cells (Ahmed et al. 2000). The third site which is a protein kinase C phosphorylation site is also significant as protein kinase C has been shown to exert both inhibitory and stimulatory influences on apoptosis depending on the isoforms involved (Gutcher et al, 2003). Another interesting structure prediction was the extended sheet conformation predicted in almost half the protein amino acid content suggesting a high possibility for aggregation in solution. In addition, protein analysis suggests that MOST-1 is unstable in nature. These findings suggest that MOST-1 may be either associate with membranes or in vesicular structures or have a high propensity to aggregate or oligomerize. Protein oligomerization has been reported in studies with rhodopsin which is coincidentally a G-protein coupled receptor responsible for the capture of light as an initial step in phototransduction. SDS-PAGE analysis of this rhodopsin results in a 'ladder' of protein bands indication of oligomerization. This may explain the high molecular weight seen of *MOST-1* in western analysis of cell lysates. Serine 91 is predicted to be a site of O-glycosylation. Glycosylation was thought to be the factor contributing to the larger than expected molecular size. It has been reported that aberrant glycosylation has been implicated in the oncogenic

transformation and subsequently in the induction of invasion and metastasis , also protein glycosylation have been shown to contribute >50% by weight to the molecular weight of proteins (Fattoraossi et al, 2002; Litman et al, 2002; Carter et al, 2002; Hakomori 2002). *MOST-1* ORF contains ~17% serine/threonine residues which are potential sites for O-linked glycosylation. Deglycosylation assay was done and shown that *MOST-1* is unlikely to be glycosylated.

Aggregation and implication of MOST-1 function

Concurrent experiments to characterize the polyclonal antibody did however suggest the ability of *MOST-1* to aggregate especially in cancer cell lines. Misfolded conformation has been reported in other studies with cancer cells. Protein conformation diseases (PCD) in which altered protein conformation has been sighted as a common feature which can result in loss of catalytic activity, structural functions and stability. This in turn will lead to accumulation of abnormal proteins which may results in protein aggregation. PCD can be divided into 2 groups, one in which aggregation of large masses of misfolded protein results in destruction of cells or the other where genetic errors results in misfolded conformation. The former include disorders such as Alzheimer's disease and Parkinson's diseases while the latter include cystic fibrosis, inherited emphysema and many types of cancer (Ishimaru, 2003). As high as 30% of all newly synthesized cellular proteins are defective ribosomal products, these are degraded by proteosomes shortly after their synthesis in

normal or unstressed cells. Thus there is no accumulation of protein aggregates despite their continued production. The cellular 'quality control' machinery suppresses formation of aggregates by ensuring the fidelity of transcription and translation, use of chaperones for correct protein folding and use of proteosomes for degrading improperly folded polypeptides. Because protein aggregates are more stable than the intermediate conformers, to degrade misfolded substrates effectively, the proteosomes must compete for the intermediates before they aggregate (Reviewed by Kopito, 2000). It is however not clear, whether aggregation of misfolded proteins is a primary cause or merely a consequence of conformations diseases but recent reports have indicated that aggregation of misfolded proteins can be a selective process dependent on peptide composition which in turn show specificity in their intracellular distribution and association with other proteins (Milewki et al, 2002). Deaggregation has been shown to be dependent on size with efficiency decreasing as size of aggregates increase (Diamant et al, 2000). On the other hand, one of the recent studies with p53, a tumor suppressor involved in ~50% of human cancers showed that heat denaturation produces granular-shaped aggregates which are similar to those found in early aggregation stages of α-synuclein and amyloid-β. The same study also illustrates that cells with wild-type p53, a tumor suppressor, can also have inactive p53. This is achieved by aggregates "locking" away the available pool of functional protein thus allowing malignant cells to arise (Ishimaru et al, 2003). A study involving guanidinobenzoatase shows that aggregation can result in a higher molecular

weight entity accompanied with a change in catalytic activity (Poustis-Delpont et al, 1994). Protein aggregation appears to be a double-edged sword whereby on one hand, they could be a third line of defense against harmful misfolded proteins by depositing them in a biologically inert form. On the other, occurrence of large protein aggregates could trigger further cell damage and thereby aggravating the deleterious effects of stress situations (Zatloukal et al, 2002). Proteosomes clearance of MOST-1 may have been inhibited due to a faster rate of aggregation to stable compounds. MOST-1 protein distribution in the cytoplasm lends further supports that MOST-1 aggregation appears to be a stable structure within cytoplasmic stippling. Punctate protein pattern has been implicated to be associated with proteins in membrane bound organelles. Examples are cystain C to the endosomes and lysosomes in the cytoplasm of neurons (Deng et al, 2001), NuMA in nucleus (Gobert et al, 2001), CHMP1 subnuclear region (Stauffer et al, 2001), NABC1 to cytoplasm itself (Beardsley et al, 2003) and CHMP4b to perinuclear area (Katoh et al, 2003). Taken together, punctate appearance of proteins in cells occur when they are either overexpressed, misfolded or organelle/vesicle bound. Similarly, the appearance of MOST-1 when overexpressed or deregulated in cancer cells could be an indication that MOST-1 'aggregation' might somehow be aiding in cell survival and in itself maybe be a mechanism to buy some time for the cell to recover and sustain itself. This phenomenon is reiterated in the cell synchronization experiments whereby overexpression of MOST-1 is seen as

large round vesicles formation compared to low expression of MOST-1 in cytoplasmic stippling.

Interactors and their possible function with MOST-1

Y2H screen results in several candidates which seem to echo our initial aim of isolation of a gene involved in cancer progression or cell cycle. The bulk of the interactors are found to be either amplified or overexpressed in cancers and involved in energy metabolism or cell cycle. As the name suggested, creatine kinase isoenzymes catalyze the synthesis of phosphocreatine by adding a phosphate to creatine turning it into the high-energy molecule phosphocreatine. Phosphocreatine is burned as a quick source of energy and it's subsequently used in the regeneration of ATP in cell types where the consumption of ATP is rapid or sudden. Although CK is predominantly found in muscle cells, it has also been proposed that different creatine kinase isoforms serves as an energy shuttle between mitochondria and cytoplasmic creatine kinase in discrete cellular sites of high ATP turnover (Shen et al, 2002, De Groof et al, 2002; Manos et al, 1993). This is supported in studies in which isoenzyme-specific cellular localization and subcellular compartmentation of CK has been shown to have a direct functional coupling of membrane bound CK to the Na+/K(+)-ATPase in cells such as electrocytes, retina photoreceptor cells, brain, kidney, salt glands, placenta, pancreas, thymus, thyroid, intestinal cells etc (Wallimann T et al, 1994). Again, cellular localization appear to be of importance to the type of function of CK

suggesting that interaction with MOST-1 at the correct localization may have some effect in energy metabolism. CK is generally found in the cytosol and mitochondria (Shen et al,2002). With MOST-1 predominantly found in membrane bound sections, it is likely that this interaction is physiologically relevant. Iron is needed for normal cell growth and proliferation but excessive iron caused to formation of free radicals. Ferritin acts as the soluble storage form of iron in tissue that occur in both intracellular and extracellular compartments (Parthasarathy N et al, 2002). Ferritin consists of various combinations of heavy and light chains which results in multiple isoforms (Moroz et al, 2002). High serum levels occur in liver disease, infection, inflammation, or malignancy (Tsuji et al, 2002). A study which investigated the molecular basis for iron depletion in human hepatocellular carcinoma showed tissue ferritin light chain (T-FLC) to be reduced to undetectable levels with 2D gel electrophoreses, however analysis of mRNA of T-FLC exhibits mRNA levels almost the same as those in normal tissues. This suggests that translational or post-translational modification may be the cause of its suppression (Park KS et al, 2002). In addition, Ferritin has been shown to have immunoregulatory properties with effect to control iron overload which seem to impair the generation of cytotoxic T-cells (Walker et al, 2000). This may have implication in cancer progression in which cancer cells escape the host immune system recognition. Perhaps interaction of ferritin with MOST-1 might also 'lock' available ferritin pool which then results in iron overload and hence cancer progression. This might be an explanation to the higher expression of

MOST-1 in high grade breast and prostate cancer. Ferritin intracellular distribution occurs in discrete pools in the cytosol when not secreted (Parthasarthy et al,2002) , this subcellular distribution is seen also for MOST-1, again suggesting a possible physiological relevant interaction. Peripheral benzodiazepine receptor (PBR) appears to function to transport cholesterol across mitochondrial membrane (Chaki et al, 1999). It is a critical component of the mitochondrial permeability transition pore (MPTP) which is a multiprotein complex located between inner and outer mitochondrial membranes. MPTP has been shown to be involved in the initiation and regulation of apoptosis (Galiegue et al, 2003). Interestingly, PBR is a small evolutionary conserved protein which plays important regulation of apoptosis, cell proliferation, stimulation of steroidogenesis, immunomodulation, porphyrin transport, heme biosynthesis, anion transport and regulation of mitochondrial functions (Galiegue et al, 2003). This coincides with the genomic structure of *MOST-1* and the hypothesis that there may be regulatory elements involved in *MOST-1* expression. At the expression level, PBR overexpression has been seen in aggressive phenotype of breast, colorectal and prostate cancer (Papadopoulo V, 2003). Again this coincide with our findings that *MOST-1* is overexpressed in high grade breast and prostate cancers. A recent PBR ligand study shows that ligands of PBR induce apoptosis and cell cycle arrest in esophageal cancer cells (Sutter et al, 2003). Another study suggests that binding of nanomolar concentration of PBR ligands protects cells against apoptosis (Strohmeier R et al, 2002). This contradictory role appears to

be present in MOST-1 overexpression and RNAi studies as well suggesting that the concentration and localization of MOST-1 might be important in determining its function. Another interactor of MOST-1 is SNC73 protein, a novel immunoglobulin protein which has been found to be down-regulated in colorectal cancer. The protein is one of the immunoglobulin heavy chain molecules with its constant region identical to that of IgA1 (Hu et al, 2003). This in addition to the interaction with immunoglobulin C (mu) and C (delta) heavy chain genes and other interactors above which are immunoregulators suggest MOST-1 to be closely associated with immune system and may function to evade immune surveillance of cancer cells. Interestingly, there is an interaction with an oncogene, v-FGR. This oncogene encodes a chimeric oncoprotein composed of feline sarcoma virus derived gag and cellular derived actin and c-Fgr sequences. v-FGR must be myristyolated and membrane bound to induce cellular transformation re-confirming that post-translation modification and subcellular localization is necessary for certain protein function (Baker et al, 1998). Interaction with this protein could have again explaining *MOST-1* amplification in high grade breast and prostate cancers and suggest a role of MOST-1 in cellular transformation. Titin has been reported to control assemble and elasticity in chromosomes. Activated titin kinase phosphorylates telethonin which is a protein of cardiac and skeletal muscle. Telethonin has been identified as a sarcomeric Z-disk protein and has been shown to be involved in the reorganization of the cytoskeleton during myofibrillogenesis (Mayans et al, 1998). However since *MOST-1* is not found

in these tissues under normal circumstances suggest that this interaction might be a non-existent in physiological conditions. A summary of the interactors functions are shown in Table 11.

Table 11: Summary of Y2H interactors function

Interacting partner	Function
Creatine Kinase	Has a role in energy metabolism in cells with high and fluctuating energy requirements e.g. skeletal, cardiac, neural tissues and spermatozoa. B-CK is the isoform found in non-muscle cells and found to be overexpressed in solid tumors and cells lines such as breast and prostate carcinoma. Postulated to be the energy source for malignant cells.
Ferritin	Function to store and release Iron (II) in a controlled fashion but can act as a double-edged sword by encouraging formation of cancer-causing free radicals. High serum ferritin levels are associated with inflammation, liver disease, anemia, thalassemia, malignant disease e.g. leukemia, malignant lymphoma and breast cancer It is also a risk factor for primary hepatocellular carcinoma.
Benzodiazepine receptor (Peripheral)	Found in mitochondria and is a key factor in the flow of cholesterol in liver. Appear to be involved in early apoptosis.
Immunoglobulin (mu and delta) heavy chain regions	Involved in the differentiation of immunoglobulin for the formation of IgM and IgD.
Gardner feline sarcoma viral homologue (v-FGR oncogene)	Belongs to the RNA tumor virus tyrosine-kinase oncogene family which causes C-terminal truncation and thus loss of phosphorylatable tyrosine residues.
SNC73	Encodes one Ig heavy chain with constant region identical to IgA1. Expression is lowered in colorectal cancerous tissue compared to non-cancerous colorectal mucosa.
Telethonin	Sarcomeric Z-disk protein; involved in the reorganization of the cytoskeleton during myofibrillogenesis.

MOST-1 Expression and Cell Cycle

Interestingly, the novel prostate stem cell antigen (PSCA) prostate-specific gene with homology to a G-protein-coupled receptor is also overexpressed in prostate cancers (Xu et al., 2000). PSCA is located in the same locus as MOST-1 8q24.2. This suggest a focus of prostate related genes located in this locus which when deregulated, increases the risk of prostate cancer development. Cell line profiling shows that *MOST-1* was expressed in all the human cancer cell lines tested, but only in specific normal human adult tissues. Another significant finding was the detection of *MOST-1* mRNA in human fetal lung but not in adult lung tissue, suggesting that *MOST-1* expression is dependent on tissue differentiation. This supports the implication of *MOST-1* in the process of carcinogenesis. Its gene expression somehow becomes altered and since with efficient translation due to the ideal Kozak sequence and suggests that the MOST-1 protein may actually enable the better survival of cancer cells.

However, there appears to be no correlation of MOST-1 to cell cycle. This observation is confirmed with cell synchronization experiments on breast normal and cancer cell lines synchronized with mimosine (Figure 15) but MOST-1 appears to accumulate in the cytoplasm during cell senescence, i.e. when majority of the cells are not undergoing cell division. In addition, cell synchronization experiments suggest a biological 'timing' of the expression of MOST-1 both in cancer and normal breast cells given the complete

disappearance of MOST-1 expression at certain times namely, 8-12h for MCF 7 cells and 10-12h and 48h for normal mammary cells.

Subsequently, overexpression and RNAi studies were done and assayed with cell proliferation assay and cell death assay to give an insight on the impact of MOST-1 expression. Overexpression seems to be difficult to achieve which once again lends support that there maybe additional regulatory mechanism which limit the amount of *MOST-1* transcripts in the cells. While RNAi knock down shows a further increase in cell proliferation for prostate normal cells but not for carcinoma cell line, perhaps due to the existing levels of MOST-1 in carcinoma cell line being higher than in prostate normal cell lines. A point to note is the high level of proliferation present in DU145 even without treatment relative to the rest of the cell lines tested. Graph 3B shows an increase in DU145 cell death both when MOST-1 is overexpressed and knock down but more so when MOST-1 is overexpressed. Prostate normal cells show only increase in cell death when MOST-1 is knock down but not when overexpressed. A point to note is the high level of cell death in BHK cells relative to the rest of the cell lines tested. In conclusion, RNAi causes cell population to be two extremes, increase in cell proliferation and increase in cell death. For cancer prostate cell line, DU145, there is an increase in cell proliferation when MOST-1 is overexpressed and yet a corresponding increase in cell death suggesting again two population of cells which react differently to the levels of MOST-1 while RNAi causes increase in cell death. There is no

significant changes for the other cell lines tested due to the high standard deviation in the percentage of cells.

Current Perspectives and Future Directions

Any study with novel gene tends to be a difficult endeavor due to the unlimited possibilities of what the gene function might be. This study has characterize the genetic component of *MOST-1* to correlate with high grade cancer biopsies of breast and prostate. Initial protein and functional study undertaken here succeeded in giving a 'hint' that the function of MOST-1 protein might be to aid in cell survival especially in cancer cells. Perhaps more in depth study would give a better picture to where MOST-1 protein stands in cancer progression. Interactors 'fished' out by Y2H screen suggest that MOST-1 might be involved in energy metabolism which is vital to cell survival, immune evasion which is vital for cancer cell proliferation and progression and cell cycle regulation. In a bid to conserve energy, cells may be required to go into G1 resting phase thus MOST-1 accumulation may act as a link to a sensor of the energy state of the cell and in turn affect cell cycle. This hypothesis may explain the contradictory effects in which MOST-1 has in the functional analysis when it is either overexpressed or knock down. MOST-1 appears to have different functions in different cells at different expression levels at different subcellular localization making it a candidate choice when it comes to overexpression for cancer cell survival. In addition, since cancer is a

multifactor disease, environmental cofactors maybe needed for cell carcinogenesis to proceed following MOST-1 deregulation.

With this study as a platform for the initial characterization of *MOST-1* gene and its potential protein function, we can envisage how different directions will the next phase of research embark to address the issues of MOST-1 with cancer progression, aggregation and evolutionary implication. Since it is found that MOST-1 is likely to be in the microsomal fraction and its interaction with a number of mitochondria and membrane bound proteins, further characterization of its subcellular localization can be done with specific organelle markers either by confocal microscopy or fractionated Western analysis. Investigation into the post-translation modifications predicted should be undertaken. Site directed mutagenesis to the N-myristoylation site can be done and observed to change in MOST-1 localization compared to wild-type. *In vitro* phosphorylation reaction could also be done to address the two other possible post-translation modifications. Taken together, this study have suggest that *MOST-1* can serve as a differentiating tool of breast and prostate tumor progression, is probably evolutionary conserved and may be needed in small amount for cell survival but upon excess, protein aggregation takes place. Aggregation of *MOST-1* may be protective in nature where by free *MOST-1* is 'lock' away in aggregates and non-functional. At the same time, cancer cells have the interesting ability to tolerate high expression of *MOST-1* suggesting that there may be other mechanisms to tolerate the aggregates in exchange for better cell survival.

Future studies involving immunoprecipitation of the aggregated protein and its possible components would allow further insight and confirmation of its interactors and functions. Co-expression of MOST-1 and chaperones could also be done to address its propensity to aggregate or misfold upon over expression. Characterization of the aggregated component with γ-tubulin and ubiquitin for aggresome formation can also be undertaken. At the same time, co-localization of endogenous MOST-1 with its interactors at different cell stage and type would lend insight to its function based on its localization and interaction. Since *MOST-1* gene is locate near HPV 18 integration site, future studies involving the analysis of cervical carcinoma biopsies would address *MOST-1* relevance in HPV-induced carcinogenesis. Figure 24 summarizes the characterization of MOST-1 in this study.

In conclusion, with recent studies showing nonrandom HPV integration into sites which are linked to genetic aberrations in cancers (Feber et al, 2003) and the observation that amplification of 8q can be used to discriminate progressor and nonprogressors in prostate cancer (Dekken et al,2003), the future identification of MOST-1 function and role in carcinogenesis would provide a new biomarker if not potential therapeutic intervention.

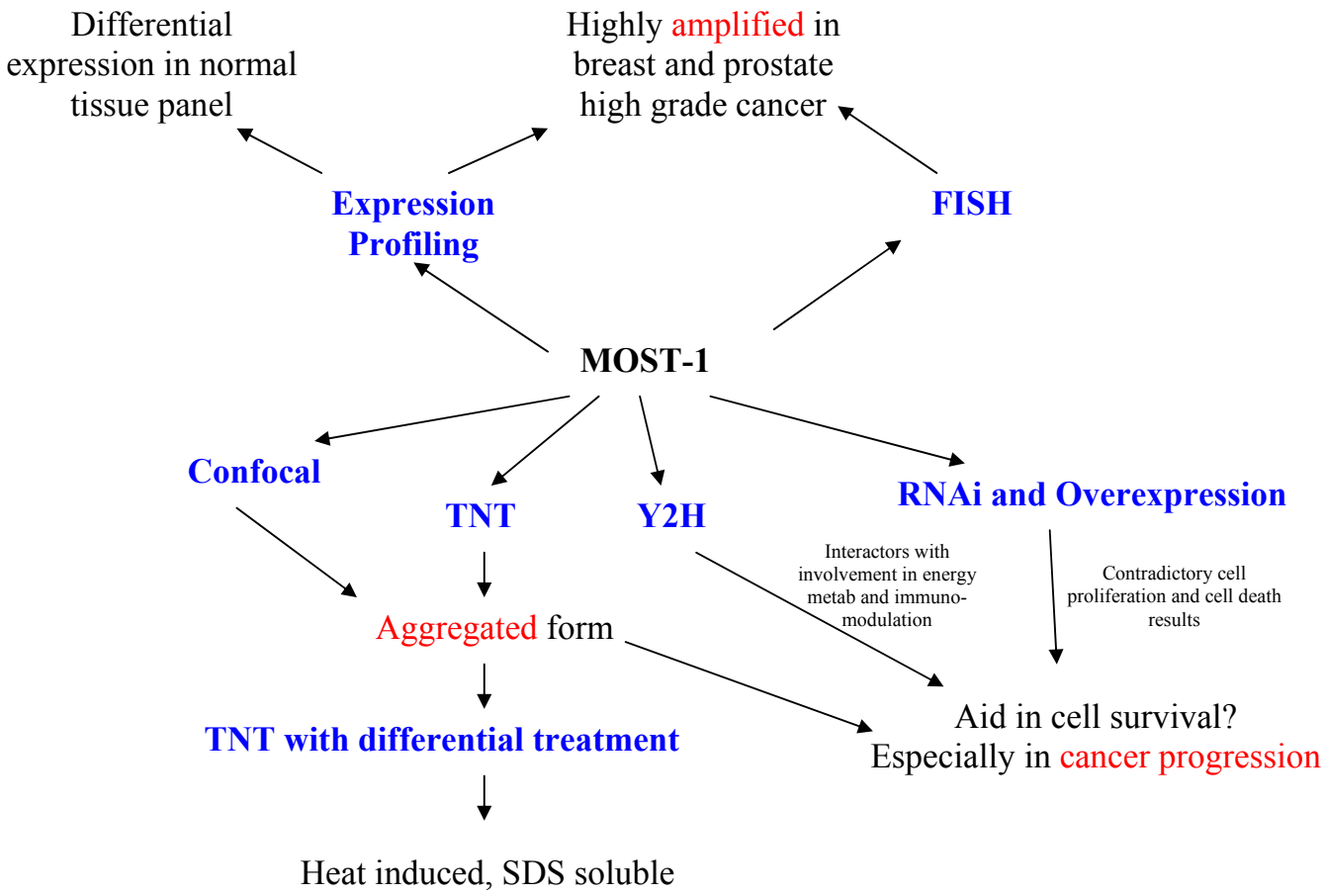


Figure 24: Conclusions of MOST-1 characterization.

REFERENCES

Abdel-Mageed AB, Agrawal KC. 1997. Antisense down-regulation of metallothionein induces growth arrest and apoptosis in human breast carcinoma cells. Cancer Gene Ther 4:199-207.

Adams MD, Kelly JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR, Venter JC. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252:1651-1656.

Ahmed K, Davis AT, Wang H, Faust RA, Yu S, Tawfic S. 2000. Significance of protein kinase CK2 nuclear signaling in neoplasis. J Cell Biochem Suppl. Suppl35: 130-5.

Akashi M, Shaw G, Hachiya M, Elstner E, Suzuki G, Koeffler P. 1994. Number and location of AUUUA motifs: role in regulating transiently expressed RNAs. Blood 83:3182-3187.

Baker SJ, Cosenza SC, Reddy EP. 1998. The role of v-Fgr myristoylation and the Gag domain in membrane binding and cellular transformation. Virology. 249:1-11.

Beardsley DI, Kowbel D, Lataxes TA, Mannino JM, Xin H, Kim WJ, Collins C, Brown KD. 2003. Characterization of the novel amplified in breast cancer-1 (NABC1) gene product. Exp Cell Res. 290 (2): 402-413.

Bosch FX, Lorincz A, Munoz N, Meijer CJLM, Shah KV. 2002. The causal relation between human papillomavirus and cervical cancer. J Clin Pathol 55: 244-265.

Brown D, Jarvis R, Pallotta V, Byrom M, Ford L. RNA interference in mammalian cell culture: design, execution and analysis of the siRNA effect. Ambion Technotes Newsletter. 9(1): 3-5.

Beutner KR, Tyring S. 1997. Human papillomavirus and human disease. Am J Med. 102 (5A): 9-15.

Bishop AJ, Schiestl RII. 2000. Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects. Hum Mol Genet. 2: 545-548.

Chaki S, Funakoshi T, Yoshikawa R, Okuyama S, Okubo T, Nakazato A, Nagamine M, Tomisawa K. 1999. Binding characteristics of [3H]DAA1106, a novel and selective ligand for peripheral benzodiapine receptors. Eur J Pharmacol. 371 (2-3): 197-204.

Chang GT, Tapsi N, Steenbeek M, Blok LJ, van Weerden WM, van Alewijk DC, Eussen BH, van Steenbrugge GJ, Brinkmann AO. 1999. Identification of a gene on human chromosome 8q11 that is differentially expressed during prostate-cancer progression. Int J Cancer 83:506-511.

Chow VTK, Lim KM, Lim D. 1998. The human *DENN* gene: genomic organization, alternative splicing, and localization to chromosome 11p11.21-p11.22. Genome 41:543-552.

Chow VTK, Loh E, Yeo WM, Tan SY, Chan R. 2000. Identification of multiple genital HPV types and sequence variants by consensus and nested type-specific PCR coupled with cycle sequencing. Pathology 32:204-208.

Couturier J, Sastre-garau X, Schneider-maunoury S, Labib A, Orth G. 1991. Integration of papillomavirus DNA near *myc* genes in genital carcinomas and its consequences for proto-oncogenes expression. J Virol. 65 (8): 4534-4538.

Da Silva Veiga LC, Bergamo NA, dos Reis PP, Kowalski LP, Rogatto SR. 2003. DNA gains at 8q23.2: a potential early marker in head and neck carcinomas. Cancer Genet Cytogenet. 146 (2): 110-5.

De Angelis PM, Clausen OPF, Schjolberg A, Stokke T. 1999. Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumor DNA ploidy, genotypes and phenotypes. Br J Cancer 80:526-535.

De Jonge HR, Hogema B, Tilly BC. 2000. Protein N-myristoylation: critical role in apoptosis and salt tolerance. Sci STKE. 63: PEI.

De Groof AF, Fransen JA, Errington RJ, Willems PH, Wieringa B, Koopman WJ. The creatine kinase system is essential for optimal refill of the sarcoplasmic reticulum Ca2+ store in skeletal muscle.

Dekken VH, Alers JC, Damen IAAJ, Vissers KJ, Krijtenburg PJ, Hoedemaeker RF, Wildhagen MF, Hop WCJ, van der Kwast TH, Tanke HJ, Schroder FH. 2003. Genetic evaluation of localized prostate cancer in a cohort of forty patients: Gains in distal 8q discriminates between progressors and nonprogressors. 83 (6): 789-796.

Deng A, Irizarry MC, Nitsch RM, Growdon JH, William Rebeck G.2001. Elevation of Cystatin C in susceptible neurons in Alzheimer's Disease. Am. J Pathol. 159 (3): 1061-1068.

Diamant S, Ben-Zvi AP, Bukan B, Goloubinoff P. 2000. Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J. Biol Chem. 275(28): 21107-21113.

Duensing S and Munger K. 2004. Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. Int J Cancer 109 (2): 157-62.

Ethier S. 2003. Identifying and validating causal genetic alterations in human breast cancer. Breast Cancer Research and Treatment 78: 285-287.

Feber MJ, Thorland EC, Brink AA, Rapp AK, Phillips LA, McGovern R, Gostout BS, Cheng TH, Chung TKH, Fu WY, Smith DI. 2003. Preferential integration of human papillomavirus type 18 near the c-*myc* locus in cervical carcinoma. Oncogene. 22 (46): 7233-7242.

Ferber MJ, Montoya DP, Yu C, Aderca I, McGee A, Thorland EC, Nagorney DM, Gostout BS, Burgart LF, Boix L, Bruix J, McMahon BJ, Cheung TH, Chung TK, Wong YF, Smith DI, Roberts LR. 2003b. Integrations of the hepatitis B virus (HBV) and human

papillomavirus (HPV) into the human telomerase reverse transcriptase (hTERT) gene in liver and cervical cancers. Oncogene 22 (24): 3813-20.

Fiedler M, Muller-Holzner E, Viertler HP, Widschwendter A, Laich A, Pfister G, Spoden GA, Jansen-Durr P, Zwerschke W. 2004. High level of HPV-16 E7 oncoprotein expression correlates with reduced pRb-levels in cervical biopsies. FASEB J. 18 (10): 1120-2.

Forozan F, Karhu R, Kononen J, Kallioniemi A, Kallioniemi OP. 1997. Genome screening by comparative genome hybridization. Trends Genet. 13:405-409.

Galiegue S, Tinel N, Casellas P. 2003. The peripheral benzodiazepine receptor: a promising therapeutic drug target. Curr Med Chem. 10 (16): 1563-72.

Gentles AJ and Karlin S. 1999. Why are human G-protein-coupled receptors predominantly intronless? Trends Genet 15:47-49.

Gobert GN, Hueser CN, Curran EM, Sun QY, Glinsky VV, Welshons WV, Eisenstark A, Schatten H. 2001. Histochem Cell Biol. 115(5):381-95.

Guruprasad K, Reddy BVB, Pandit MW. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. 4(2):155-61.

Gutcher I, Webb PR, Anderson NG. 2003. The isoform-specific regulation of apoptosis by protein kinase C. Cell Mol Life Sci. 60 (6): 1061-70.

Hakomori S. 2002. Glycosylation defining cancer malignancy: new wine in an old bottle. PNAS 99(16): 10231-10233.

Hall AHS and Alexander KA. 2003. RNA interference of human papillomavirus type 18 E6 and E7 induces senescence in HeLa cells. J Virol. 77 (10): 6066-6069.

Hedera P, Rainier S, Alvarado D, Zhao X, Williamson J, Otterud B, Leppert M, Fink JK. 1999. Novel locus for autosomal dominant hereditary spastic paraplegia, on chromosome 8q. Am J Hum Genet 64:563-569.

Heim A, Grumbach IM, Zeuke S and Top B. 1998. Highly sensitive detection of gene expression of an intronless gene: amplification of mRNA, but not genomic DNA by nucleic acid sequence based amplification (NASBA). Nucleic Acids Res 26 (9) 2250-2251.

Heng, HHQ, Squire J, Tsui LC. 1992. High-resolution mapping of mammalian genes by in situ hybridization to free chromatin. Proc Natl Acad Sci USA 89:9509-9513.

Heng HHQ, Tsui, LC. 1993. Modes of DAPI banding and simultaneous in situ hybridization. Chromosoma 102:325-332.

Hoeijmakers JH. 2001. Genome maintenance mechanisms for preventing cancers. Nature 411: 366-374.

Holmquist GP. 1992. Chromosome bands, their chromatin flavors, and their functional features. Am J Hum Genet 51: 17-37.

Hu JB, Zheng S, Deng YC. 2003. Expression of a novel immunoglobulin gene SNC73 in human cancer and non-cancerous tissue. World J Gastroentol. 9 (5): 1054-1057.

Inglehearn CF, McHale JC, Keen TJ, Skirton H, Lunt PW. 1999. A new family linked to the *RP1* dominant retinitis pigmentosa locus on chromosome 8q. J Med Genet 36:646-648.

Ishimaru D, Andrade LR, Teixeira LSP, Quesdo PA, Maiolino LM, Lopez PM, Cordeiro Y, Coasta LT, Heck WM, Weissmuller G, Foguel D, Silva JL. 2003. Fibrillar aggregates of the tumor suppressor p53 core domain. Biochemistry. 42:9022-9027.

Jansen, B. and Wittke, U. 2002. Antisense therapy for cancer – the time of truth. Lancet Oncol. 3: 672-83.

Jin R, Bay BH, Chow VTK, Tan PH, Lin VCL. 2000. Metallothionein 1E mRNA is highly expressed in oestrogen receptor-negative human invasive ductal breast cancer. Br J Cancer 83:319-323.

Jin R, Bay BH, Chow VTK, Tan PH. 2001. Metallothionein 1F mRNA expression correlates with histological grade in breast carcinoma. Breast Cancer Res Treat 66:265-272.

Jin R, Chow VTK, Tan PH, Dheen ST, Duan W, Bay BH. 2002. Metallothionein 2A expression is associated with cell proliferation in breast cancer. Carcinogenesis 23:81-86.

Kang YK, Park JS, Lee CS, Yoem YI, Chung AS, Lee KK. 1999. Efficient integration of short interspersed element-flanked foreign DNA via homologous recombination. J Biol. Chem 274: 36585-36591.

Kang MK and Park NH. 2001. Conversion of normal to malignant phenotype: telomere shortening, telomerase activation, and genomic instability during immortalization of human oral keratinocytes. Crit Rev Oral Biol Med. 12 (1): 38-54.

Kato S, Anderson RA, Camerini-Otero RD. 1986. Foreign DNA introduced by calcium phosphate is integrated into repetitive DNA elements of mouse L cell genome. Mol Cell Biol 6: 1787-1795.

Katoh K, Shibata H, Suxuki H, Nara A, Ishidoh K, Kominami E, Yoshimori T, Maki M. 2003. The ALG-2-interacting protein Alix associates with CHMP4b, a human homologue of yeast Snf7 that is involved in multivesicular body sorting. J Biol Chem. 278 (40): 39104-13.

Kaufmann AM, Backsch C, Schneider A, Durst M. 2002. HPV induced cervical carcinogenesis: molecular basis and vaccine development. Zentralbl Gynakologische. 124 (11): 511-24

Kim HJ, Kim SG. 2002. Alterations in cellular Ca(2+) and free iron pool by sulfur amino acid deprivation: the role of Ferritin light chain down-regulation in prooxidant production. Biochem Pharmacol. 63 (4): 647-57.

Knuutila S, Aalto Y, Autio K, Bjorkqvist AM, El-Rifai W, Hemmer S, Huhta T, Kettunen E, Kiuru-Kuhlefelt S, Larramendy ML, Lushnikova T, Monni O, Pere H, Tapper J,

Tarkkanen M, Varis A, Wasenius VM, Wolf M, Zhu Y. 1999. DNA copy number losses in human neoplasms. Am J Pathol 155:683-694.

Kopito RR. 2000. Aggresomes, inclusion bodies and protein aggregation. Trends in cell biol. 10: 524-531.

Kozak M. 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 266:19867-19870.

Kuukasjarvi T, Tanner M, Pennanen S, Karhu R, Kallioniemi OP, Isola J. 1997. Genetic changes in intraductal breast cancer detected by comparative genomic hybridization. Am J Pathol 150:1465-1471.

Krude T. 1999. Mimosine arrests proliferating human cells before onset of DNA replication in a dose-dependent manner. Expt Cell Research. 247: 148-159.

Larsen F., Gundersen G., Lopez R. and Prydz H. 1992. CpG islands as gene markers in the human genome. Genomics 13: 50-54.

Ledwaba T, Dlamini Z, Naicker S, Bhoola K. 2004. Molecular genetics of human cervical cancer: role of papillomavirus and the apoptotic cascade. Biol Chem 385 (8):671-82.

Leirdal M, Sioud M. 2002. Gene silencing in mammalian cells by preformed small RNA duplexes. Biochem Biophy Res Comm. 295: 744-748.

Lewin Benjamin, 1997. Genes VI. Oxford University Press.

Loveday RL, Greenman J, Drew PJ, Monson JRT, Kerin MJ. 1999. Genetic changes associated with telomerase activity in breast cancer. Int J Cancer 84:516-520.

Lovett M. 1994. Fishing for complements: finding genes by direct selection. Trends Genet. 10: 352-357.

Manos P and Bryan GK. 1993. Cellular and subcellular compartmentation of creatine kinase in brain. Dev Nurosci. 15 (3-5): 271-9.

Mantovani F, Banks L. 2001. The human papillomavirus E6 protein and its contribution to malignant progression. Oncogene 20:7874-7887.

Mayans O, Van der Ven PFM, Wilm M, Mues A, Yong P, Furst DO, Wilmanns M, Gautel M. 1998. Structural basis for activation of the titin kinase domain during myofibrillogenesis. Nature. 395: 863-869.

McGlennen RC. 2000. Human papillomavirus oncogenesis. Clin Lab Med 20 (2): 383-406.

McManus MT, Sharp PA. 2002. Gene silencing in mammals by small interfering RNAs. Nature Genetics. 3: 737-747.

Mercola D and Welsh J. 2004. From mRNA to tumor suppressor. Nature Genetics. 36 (9): 937-938.

Mielenz D, Ruschel A, Vettermann,C, and Jäck HM. 2003. Immunoglobulin μ heavy chains do not mediate tyrosine phosphorylation of Igα from the ER/Cis-Golgi 1. J. Immunol. 171 (6): 3091-101.

Michaelson JS and Leder P. 2003. RNAi reveals anti-apoptotic and transcriptionally repressive activities of DAXX. J Cell Sci. 116 (2): 345-352.

Milewski, MI, Mickle JE, Forrest JK, Stanton BA. 2002. Aggregation of misfolded proteins can be a selective process dependent upon peptide composition. J Biol Chem. 277 (37): 34462-34470.

Milunsky A, Huang XL, Milunsky J, DeStefano A, Baldwin CT. 1999. A locus for autosomal recessive achromatopsia on human chromosome 8q. Clin Genet 56:82-85.

Morin MJ. 2000. From oncogene to drug: development of small molecule tyrosine kinase inhibitors as anti-tumor and anti-angiogenic agents. Oncogene 19: 6574-6583.

Nakamura S. 1993. Possible role of phosphorylation in the function of chicken MyoD1. J Biol Chem. 268 (16):11670-7.

Nekrutenko A and Li WH. 2001. Transposable elements are found in a large number of human protein-coding genes. TIG. 17 (11): 619-621.

Nupponen NN, Kakkola L, Koivisto P, Visakorpi T. 1998. Genetic alterations in hormone-refractory recurrent prostate carcinomas. Am J Pathol 153:141-148.

Nupponen NN, Porkka K, Kakkola L, Tanner M, Persson K, Borg A, Isola J, Visakorpi T. 1999. Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. Am J Pathol 154: 1777-1783.

Onyango P. 2002. Genomics and cancer. Curr Opin Oncol 14:79-85.

Papadopoulo V. 2003. Peripheral benzodiazepine receptor: structure and function in health and disease. Ann Pharm Fr. 61 (1):30-50.

Park KS, Kim H, Kim NG, Cho SY, Choi KH, Seong JK, Paik YK. Proteomic analysis and molecular characterization of tissue Ferritin light chain in hepatocellular carcinoma. Hepatology. 35(6): 1459-66.

Parthasarathy N, Torti SV, Torti FM. 2002. Ferritin binds to light chain of human H-kininogen and inhibits kallikrein-mediated bradykinin release. Biochem J. 365: 279-86.

Pett MR, Alazawi WO, Roberts I, Dowen S, Smith DI, Stanley MA, Coleman N. 2004. Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. Cancer Res 64 (4): 1359-68.

Porkka KP, Tammela TL, Vessella RL, Visakorpi T. 2004. RAD21 and KIAA0196 at 8q24 are amplified and overexpressed in prostate cancer. Genes Chromosomes Cancer. 39 (1): 1-10.

Poustis-Delpont C, Thaon S, Auberger P, Gerardi-laffin C, Sudaka P, Rossi B. 1994. Monomeric 55-kDa guanidinobenzoatase switches to a serine proteinase activity upon tetramerization.

Quek HH, Chow VTK. 1997. Genomic organization and mapping of the human HEP-COP gene (COPA) to 1q. Cytogenet Cell Genet 76:139-143.

Redolfi E, Montagna C, Mumm S, Affer M, Susani L, Reinbold, R, Hol F, Vezzoni P, Cimino M and Zucchi I. 1998. Identification of *CXorf1*, a novel intronless gene in Xq27.3, expressed in human hippocampus. DNA and Cell Biology 17 (12): 1009-1016.

Reiter RE, Gu Z, Watabe T, Thomas G, Szigeti K, Davis E, Wahl M, Nisitani S, Yamashiro J, Le Beau MM, Loda M, Witte ON. 1998. Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer. Proc Natl Acad Sci USA 95:1735-1740.

Roberts G. 2003. Gene Isolation. Edited by Timothy Paustian, University of Wisconins-Madison.

Rubenstein P, Smith P, Deuchler J, Redman K. 1981. NH2-terminal acetylation of Dictyostelium discoideum actin in a cell-free protein-synthesizing system. J Biol Chem. 256 (15):8149-55.

Rubin MA, Varambally S, Beroukhim R, Tomlins SA, Rhodes DR, Paris PL, Hofer MD, Storz-Schweizer M, Kuefer R, Fletcher JA, His BL, Byrne JA, Pienta KJ, Collins C, Sellers WR, Chinnaiyan AM. 2004. Overexpression, amplification, and androgen regulation of TPD52 in prostate cancer. Cancer Res 64: 3814-3822.

Saramaki O, Willi N, Bratt O, Gasser TC, Koivisto P, Nupponen NN, Bubendorf L, Visakorpi T. 2001. Amplification of EIF3S3 gene is associated with advanced stage in prostate cancer. Am J Pathol 159:2089-2094.

Schena M. 1996. Genome analysis with gene expression microassays. BioEssays 18: 427-431.

Scacheri PC, Rozenblatt-Rosen O, Caplen NJ, Wolfsberg TG, Umayam L, Lee JC, Hughes CM, Shanmugam KS, Bhattacharjee A, Meyerson M, Collins FS. 2004. Proc Natl Acad Sci U S A. 101 (7): 1892-7.

Shen W, Willis D, Zhang Y, Schlattner U, Wallimann T, Molloy GR. 2002. Expression of creatine kinase isoenzyme genes during postnatal development of rat brain cerebellum: evidence for transcriptional regulation. Biochem. J. 367:369-380.

Shigeeda N, Uchida M, Barett JC, Tsutsui T. 2003. Candidate chromosomal regions for genes involved in activation of alternative lengthening of telomeres in human immortal cell lines.

Sim DLC, Chow VTK. 1999. The novel human HUEL (C4orf1) gene maps to chromosome 4p12-p13 and encodes a nuclear protein containing the nuclear receptor interaction motif. Genomics 59:224-233.

Sim DLC, Yeo WM, Chow VTK. 2002. The novel human HUEL (C4orf1) protein shares homology with the DNA-binding domain of the XPA DNA repair protein and displays nuclear translocation in a cell cycle-dependent manner. IJBCB 34: 487-504.

Smit AF. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev 9: 657-663.

Squier TC. 2001. Oxidative stress and protein aggregation during biological aging. Exp Geront. 36: 1539-1550.

Stauffer DR, Howard TL, Nyun T, Hollenberg SM. 2001. CHMP1 is a novel nuclear matrix protein affecting chromatin structure and cell cycle progression. J Cell Sci. 114: 2383-2393.

Stengel D, Parma J, Gannage MH, Roeckel N, Mattei MG, Barouki R, Hanoune J. 1992. Different chromosomal localization of two adenylyl cyclase genes expressed in human brain. Hum Genet 90:126-130.

Stengel-Rutkowski S, Lohse K, Herzog C, Apacik C, Couturier J, Albert A, Belohradsky B. 1992. Partial trisomy 8q. Two case reports with maternal translocation and inverted insertion: phenotype analyses and reflections on the risk. Clin Genet 42:178-185.

Stevens CA, Moore CA. 1999. Tibial hemimelia in Langer-Giedion syndrome – possible gene location for tibial hemimelia at 8q. Am J Med Genet 85:409-412.

Stoler MH. 2000. Human papillomaviruses and cervical neoplasia: a model for carcinogenesis. Int J Gynecol Pathol 19:16-28.

Strausberg RL. 2001. The Cancer Genome Anatomy Project: new resources for reading the molecular signatures of cancer. J Pathol 195:31-40.

Strohmeir R, Roller M, Sanger N, Knecht R, Kuhl H. 2002. Modulation of tamoxifen-induced apoptosis by peripheral benzodiazepine receptor ligands in breast cancer cells. Biochem Pharmacol. 64 (1): 99-107.

Sutter AP, Maaser K, Barthel B, Scherubl H. 2003. Ligands of the peripheral benzodiazepine receptor induce apoptosis and cell cycle arrest in oesophageal cancer cells: involvement of the p38MAPK signaling pathway. Br J Cancer. 89 (3): 564-72.

Tan JMM, Tock EPC, Chow VTK. 2003. The novel human MOST-1 (C8orf17) gene exhibits tissue specific expression, maps to chromosome 8q24.2, and is overexpressed/amplified in high grade cancers of the breast and prostate. J Clin Pathol: Mol Pathol. 56: 109-115.

Tham KM, Chow VTK, Singh P, Tock EPC, Ching KC, Lim-Tan SK, Sng ITY, Bernard HU. 1991. Diagnostic sensitivity of polymerase chain reaction and Southern blot hybridization for the detection of human papillomavirus DNA in biopsy specimens from cervical lesions. Am J Clin Pathol. 95:638-646.

Thorland EC, Myers SL, Persing DH, Sarkar G, McGovern RM, Gostout BS, Smith DI. 2000. Human papillomavirus type 16 integrations in cervical tumors frequently occur in common fragile sites. Cancer Res 60: 5916-21.

Tsuchiya N, Slezak J.M, Lieber M.M, Bergstralh E.J, Jenkins R.B. 2002. Clinical significance of alterations of Chromosome 8 detected by fluorescence in situ hybridization analysis in pathologic organ-confined prostate cancer. Genes, Chromosome and Cancer 34:363-371.

Tsuji Y, Ayaki H, Whitman SP, Morrow CS, Torti SV, Torti FM. 2000. Coordinate Transcriptional and Translational Regulation of Ferritin in Response to Oxidative Stress. Mol Cell Biol. 20**:** 5818-5827.

Tuschl T. 2001. RNA interference and small interfering RNAs. Chembiochem. 2: 239-245.

Vackova I, Engelova M, Marinov I, Tomanek M. 2003. Cell cycle synchronization of porcine granulose cells in G1 stage with mimosine. Animal Rep Science. 77: 235-245.

Venter J.C, et al. 2001. The Sequence of the human genome. Science. 291 (5507): 1304-1351.

Vidal M and Legrain P. 1999. Yeast forward and reverse 'n'-hybrid systems. Nucleic Acids Research. 27 (4): 919-929.

Visakorpi T, Kallioniemi AH, Syvanen A, Hyytinen ER, Karhu R, Tammela T, Isola JJ, Kallioniemi OP. 1995. Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. Cancer Res 55:342-347.

Wallenburg JC, Nepveu A, Chartrand P. 1987. Integration of a vector containing rodent repetitive element in the rat genome. Nucleic Acids Res. 15: 7849-7863.

Wallimann T and Hemmer W. 1994. Creatine kinase in non-muscle tissue and cells. Mol Cell Biochem. 133-134: 193-220.

Walker EMJ, Walker SM. 2000. Effects of iron overload on the immune system. Ann Clin Lab Sci. 30 (4): 354-65.

Walter P and Blobel G. 1983. Preparation of microsomal membranes for cotranslational protein translocation**.** Methods Enzymol. 96, 84-93.

Weber-Mangal S, Sinn HP, Popp S, Klaes R, Emig R, Bentz M, Mansmann U, Bastert G, Bartram CR, Jauch A. 2003. Breast cancer in young women (< or = 35 years of age): Genomic aberrations detected by comparative genomic hybridization. Int J Cancer. 107 (4): 583-92.

Wright FA, Lemon WJ, Zhao WD, Sears R, Zhuo D, Wang JP, Yang HY, Baer T, Stredney D, Spitzner J, Stutz A, Krahe R, Yuan B. A draft annotation and overview of the human genome. Genome Biol 2:research0025.1-0025.18.

Xu LL, Stackhouse BG, Florence K, Zhang W, Shanmugam N, Sesterhenn IA, Zou Z, Srikantan V, Augustus M, Roschke V, Carter K, McLeod DG, Moul JW, Soppett D,

Srivastava S. 2000. *PSGR*, a novel prostate-specific gene with homology to a G protein-coupled receptor, is overexpressed in prostate cancer. Cancer Res 60:6568-6572.

Zatloukal K, Stumptner C, Fuchsbichler A, Heid H, Schnoelzer M, Kenner l, Kleinert R, Prinz M, Aguzzi A, Denk H. 2002. p62 is a common component of cytoplasmic inclusions in protein aggregation diseases. Amer J Pathol. 160(1) 255-263.

Zeng Y, Cullen BR. 2002. RNA interference in human cells is restricted to the cytoplasm. RNA 8: 855-860.

Zhou A, Scoggin S, Gaynor RB, Williams NS. 2003. Identification of NR-κB-regulated genes induced by TNFα utilizing expression profiling and RNA interference. Oncogene. 22: 2054-2064.

Zou P, Gautel M, Geerlof A, Wilmanns M, Koch MH, Svergun DI. 2003. Solution scattering suggests cross-linking function of telethonin in the complex with titin. J Biol Chem. 278 (4): 2636-44.

Zur Hausen H. 1996. Papillomavirus infections – a major cause of human cancers. Biochim Biophys Acta 1288: F55-78.

Appendix 1: Mammalian cell tissue culture media

1.1.10X Phosphate Buffered Saline (PBS) (1 liter) pH 7.4

NaCl	80g
KCl	2g
Na_2HPO_4	14.4g
KH_2PO_4	2.4g

NaCl was dissolved separately before adding to the already dissolved 3 chemicals. pH was then adjusted to 7.4. 1X PBS was then prepared by diluting the stock to 10L and then filtered sterilized with 0.2µm bell filter (Whatman #6727-5002).

1.2. Eagle's modified essential medium

With Earle's BSS and 2mM L-glutamine

$NaHCO_3$	1.5g/L
1M Hepes	2%
Sodium pyruvate	1mM
Nonessential amino acids	0.1mM
Fetal bovine serum	10%

1.3.Dulbecco's modified eagle's medium (Sigma #D7777)

$NaHCO_3$	3.7g/L
1M Hepes	2%
Fetal bovine serum	10%

1.4.RPMI-1640 medium (Sigma #R4130)

$NaHCO_3$	1.5g/L
Hepes	25mM
Fetal bovine serum	10%

1.5.Penicillin/Streptomycin (1000X)

Penicillin G sodium BP (Glaxo)	5×10^6 U
Streptomycin sulphate BP (Glaxo)	5g

Appendix 2: Buffers and Reagents for Genome Work

Template isolation

2.1. SSE Extraction Buffer (pH 7.0)

NaOAc	0.3M
SDS	0.5%
EDTA	5mM

2.2. SSE equilibrated phenol

Phenol was melted at 55°C for 1 hour or until liquid. SSE pH 7.0 was added and stirred for 2 hours prior to phase separation O/N. pH of Phenol was monitored with successive change of SSE until pH of Phenol = 7.0.

2.3. 1X Tris-EDTA Buffer (TE; pH 8.0)

Tris-HCl	10mM
EDTA	1mM

2.4. RNA Extraction Buffer A (pH 7.5)

Tris-HCl	10mM
NaCl	0.15M
$MgCl_2$	1.5mM
Nonidet P-40	0.65%

2.5. RNA Extraction Buffer B (pH 7.5)

Urea	7M

SDS	1%
EDTA	10mM
NaCl	0.35M
Tris-HCl	10mM

Northern Blot Analysis

2.5 20XSSC, pH 7.0

NaCl	3M
Na Citrate	0.3M

2.6 Wash Solution 1

SSC	2X
SDS	0.05%

2.7 Wash Solution 2

SSC	0.1X
SDS	0.1%

Appendix 3: Buffers and Reagents for Proteome Work

Western Analysis

3.1 10X Sample Buffer

1M Tris, pH 6.8	625ul
10% SDS	2ml
2-β-mercaptoethanol	0.5ml
Glycerol	1ml
Bromophenol Blue	0.0025g
ddH_2O	Top up to 10ml

3.2 10X Laemmli Running Buffer

Tris	30.3g
Glycine	144.2g
SDS	10g
ddH_2O	Top up to 100ml

Dilute 10x for each run.

3.3 Acrylamide Monomer (30%)

Acrylamide	30g
N,N-Bisacrylamide	0.8g
ddH_2O	Top up to 100ml

Filtered through 0.45um filter and de-gassed.

3.4 Blocking Solution

Bovine Serum Albumin	3%
Gelatin	0.25%
NaCl	150 mM
Tris-HCl (pH 7.4)	15 mM

3.5 TBST, pH 7.4

NaCl	150 mM
Tris-HCl (pH 7.4)	15 mM
Tween-20	0.05%]

3.6 Alkaline phosphatase buffer, pH9.5

NaCl	2.922g
$MgCl_2$	0.508g
1M Tris HCl, pH 7.5	50ml
ddH_2O	450ml

Adjust to pH 9.5 and add ddH_2O to a final volume of 500ml

3.7 CAPS TRANSFER BUFFER (1 liter)

CAPS	1X
Methanol	10% w/v
ddH_2O	Top to 1 liter

3.8 10X CAPS (1 liter)

3-cyclohexylamino-1-porpanesulfonic acid	22.1g
ddH_2O	950ml

pH was adjusted to 11.0 with 5N NaOH before adding ddH_2O to 1 litre.

3.9 NBT (30mg/ml)

NBT was dissolved in 1ml of 70% DMF.

3.10 BCIP (15mg/ml)

BCIP was dissolved in 1ml of 100% DMF.

3.11 NBT/BCIP Color Development solution

Just prior to use, mix 1ml of NBT and 1 ml of BCIP in 100ml of Tris Buffer

3.12 Tris Buffer for Color Development (pH 9.5)

0.1M Tris, 0.5mM $MgCl_2$ dissolved in ddH_2O.

Protein extraction

3.13 Lysis Buffer I

Sodium Deoxycholate	1%
Triton-X 100	1%
Tris-HCL pH7.	0.1 M
NaCl	0.15 M

3.14 Lysis Buffer II

Nonidet P-40 [NP 40]	0.02%
NaCl	150 mM
$MgCl_2$	2 mM
$CaCl_2$	3 mM
DTT	3 mM

Indirect Immunofluorescence

3.15 3% Paraformaldehyde (3%PFA) Fixative

40ml of PBS was heated to 60°C before 3% of PFA (Sigma) was added and mixed for 30min. 10M NaOH was added until solution becomes clear and pH was adjusted to pH 6.1. Additional PBS was then added to make a total volume of 50ml. Solution was aliquot and stored at -20°C until use. Freeze-thaw cycle was not permitted.

3.16 Permeabilizing Agent 0.1% Triton X-100

10% Triton X-100 (in ddH_2O)	1ml
PBS	99ml

3.17 Blocking Solution

0.1% FCS in PBS.

3.18 Quenching Buffer

NH_4CL	0.267g
PBS	100ml

Yeast two hybrid

3.19 Yeast Peptone Dextrose (YPD)

Yeast extract	1%
Peptone	2%
Dextrose	2%

18 g/L of agar was added when needed.

3.20 Dropout media

All reagents for synthetic dropout (SD) media/agar were purchased from Clontech and prepared according to recommended protocols.

3.21. β-Galactosidase filter assay (Colony Lift)

3.21.1. Z Buffer (pH 7.0)

$Na_2HPO_4.7H_2O$	16.1g/L
$NaH_2PO_4.H_2O$	5.5g/L
KCl	0.75g/L
$MgSO_4.7H_2O$	0.246g/L

3.21.2. X-Gal Stock Solution

5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) was dissolved in DMF at a concentration of 20mg/ml and stored in dark at -20°C.

3.21.3. Z Buffer/X-Gal solution

Z buffer	100ml

B-mercaptoethanol	0.27ml
X-gal stock solution	1.67ml

Cell synchronization studies

3.22. Mimosine Stock Solution

L-Mimosine from Koa Hoale Seeds (Sigma, M0253)	25mg
Minimal Essential Media (MEM)	12.6ml

3.23. Propidium Iodide (PI) DNA Staining Solution

0.1% (v/v) Triton X-100 (in PBS)	10ml
DNase-free RNase A (Sigma; treated in 95°C for 5min)	2mg
PI (1mg/ml)	200ul

Appendix 4A: Densitometric reading of tissue screening

Normal Tissue cDNA	Average MOST-1 densitometric reading	Average G3DPH densitometric reading	Ratio of MOST-1/G3DPH
Brain	3.863636	197.4474	0.019568
Heart	48.51667	189.6081	0.255879
Kidney	69.20635	221.4767	0.312477
Liver	66.40952	216.6216	0.306569
Lung	30.55	182.5426	0.167358
Pancreas	74.03922	224.7609	0.329413
Placenta	113.2449	206.9568	0.547191
Skeletal muscle	27.39394	194.4567	0.140874
Colon	129.8857	213.3025	0.608927
Ovary	122.6579	215.2632	0.569804
Peripheral blood leukocyte	14.975	182.7232	0.081955
Prostate	14.78431	228.4171	0.064725
Small intestine	29.64912	177.2917	0.167234
Spleen	0.628788	173.3736	0.003627
Testis	64.7	195.0083	0.331781
Thymus	45.75556	165.315	0.276778

Appendix 4B: Densitometric reading of tissue screening

Cell lines cDNA	Average MOST-1 densitometric reading	Average G3DPH densitometric reading	Ratio of MOST-1/G3DPH
CaSki	191.6825	225.5429	1.176648
HeLa	214.61	202.33	1.060693
Hep3B	185.23	194.84	0.950677
HepG2	208.18	192.74	1.080108
HL60	228.03	241.1	0.94579
Kato	162.57	213.45	0.76163
Mahlavu	147.21	229.67	0.640963
Molt-4	216.67	231.31	0.936708
PP5	158	228.89	0.690288
Raji	232.46	241.85	0.961174
SiHa	216.28	240.68	0.898621
T24	196.07	229.07	0.855939
U937	232.33	145.72	1.594359
MDA-MB-231	223.21	105.98	2.106152
MCF7	230.49	204.42	1.127532
DU145	218.77	192.77	1.134876
PC3	211.67	176.83	1.197025
Hs578t	224.92	196.47	1.144806
Myoepithelial	213.96	185.85	1.151251
MRC-5	230.18	170.15	1.352806

Appendix 5: Breast Biopsies Quantification

Sample Number	Histological Grade	Ratio T/N
1.	1	0.664111
2.	1	0.483586
3.	1	0.489579
4.	2	0.933669
5.	2	0.648018
6.	2	1.179893
7.	2	0.120742
8.	2	1.152283
9.	2	0.964672
10.	2	1.379124
11.	2	0.518576
12.	2	0.959337
13.	2	0.773453
14.	2	1.451111
15.	3	0.87505
16.	3	0.35832
17.	3	4.152415
18.	3	5.890413
19.	3	1.569398
20.	3	1.171173
21.	3	0.850314
22.	3	1.220777
23.	3	1.03262
24.	3	1.444185
25.	3	1.061707
26.	3	1.459491
27.	3	0.850746

Appendix 6: Prostate Biopsies Quantification

Sample Type	Average ΔCT value	$2^{-\Delta CT}$
Normal	-4.507	22.73747
	-2.5055	5.678461
	-4.203	18.41743
	-0.695	1.618884
	-1.512	2.852051
	0.4905	0.711778
	-0.3435	1.268831
	-4.2695	19.28624
	-2.6835	6.424125
	-3.738	13.3429
	4.1995	0.054428
	-4.629	24.74388
	-0.282	1.215879
	1.14	0.45376
	-0.7345	1.663821
	-0.342	1.267513
	-1.3495	2.548238
	-1.4095	2.656451
	1.6785	0.312407
Hyperplasia	-3.8995	14.92335
	-2.3875	5.232499
	1.1775	0.442117
	-0.8785	1.838463
	-0.133	1.096572
	-0.719	1.646041
	-0.137	1.099616
	0.6635	0.631345
	0.152	0.900002
	-1.1925	2.285484
	1.7475	0.297817
	1.7995	0.287274
	2.3425	0.197168
	1.7705	0.293107
	-0.849	1.801252
	1.204	0.43407

	1.541	0.343647
	0.6375	0.642826
	-1.319	2.494931
	-2.307	4.94853
	-0.696	1.620007
Low	-1.389	2.618971
(Gleason Score 3-5)	-3.825	14.17228
	-1.489	2.806943
	-1.7375	3.334568
	-3.139	8.809133
Intermediate	-3.1645	8.966221
(Gleason Score 6-7)	-3.967	15.63817
	-4.0695	16.78965
	-4.4065	21.20746
	-2.998	7.988917
	-2.936	7.652865
	-2.588	6.012646
	-4.053	16.59872
	-4.359	20.52059
	-4.6935	25.87523
	-2.3945	5.257948
	-0.2275	1.170804
	-0.947	1.92786
	-0.1735	1.127791
High	-1.949	3.861068
(Gleason Score 8-9)	-4.552	23.45787
	-5.302	39.45127
	-2.561	5.901166
	-4.013	16.14483
	0.8355	0.560389

Appendix 7: Biopsies information

Clinicopathological data of invasive ductal breast cancer patients from the Singapore General Hospital involved in this study graded according to Bloom et al 1957.

Patient age, yr	Tumor Size, cm	Histological Grade
49	6	2
46	2.5	2
48	2.5	2
55	3	3
50	3	2
41	2	3
58	5	3
55	4	3
42	2.5	3
43	3	3
64	2.8	2
61	5	2
57	2	2
45	6	3
51	2.5	3
62	3.5	2
40	3.7	2
77	3	1
71	4	1
52	1.8	2
33	6.5	3
39	4	3
86	5	3
49	3.4	2
52	4.5	3
60	1.75	2
73	2.8	1

Clinicopathological data of prostate cancer biopsies from the National University Hospital involved in this study graded according to Gleason Score.

Patient age, yr	Race	Histological Grade	Gleason Score
77	Indian	Adenocarcinoma	5
74	Chinese	Adenocarcinoma	7
60	Chinese	Adenocarcinoma	7
75	Chinese	Adenocarcinoma	5
72	Chinese	Adenocarcinoma	7
65	Chinese	Adenocarcinoma	8
77	Indian	Adenocarcinoma	5
67	Malay	Adenocarcinoma	9
61	Chinese	Adenocarcinoma	3
64	Chinese	Adenocarcinoma	9
65	Chinese	Adenocarcinoma	6
70	Malay	Adenocarcinoma	6
Unknown	Chinese	Adenocarcinoma	6
74	Chinese	Adenocarcinoma	5
80	Chinese	Adenocarcinoma	9
74	Malay	Adenocarcinoma	7
63	Eurasia	Adenocarcinoma	6
84	Chinese	Adenocarcinoma	7
66	Chinese	Adenocarcinoma	7
84	Chinese	Adenocarcinoma	7
81	Eurasia	Adenocarcinoma	7
72	Malay	Adenocarcinoma	8
63	Eurasia	Adenocarcinoma	7
70	Chinese	Adenocarcinoma	7

Publications List

- **Tan JMM.**, Tock EPC and Chow VTK (2003). The novel human MOST-1 (C8orf17) gene exhibits tissue specific expression, maps to chromosome 8q24.2, and is overexpressed/amplified in high grade cancers of the breast and prostate. *Journal of Clinical Pathology: Molecular Pathology* 56:109-115.

- **Tan JMM.,** Chow VTK. Proteomics and functional analysis of MOST-1 protein. Manuscript in preparation.

Conference Abstracts

- **Tan JMM,** Chow VTK. (MOST-1, a novel human intronless gene that exhibits tissue-specific expression, maps to chromosome 8q24.2 and is amplified in a subset of breast cancers) 2nd SSBMB/SSMB/BRETSS Combined Annual Scientific Meeting. 8-9 September 2000. Singapore

  Awarded 1st Runner-up for best poster.

- **Tan JMM,** Chow VTK. (The novel, intronless, differentially-expressed human gene maps to chromosome 8q24.2, is amplified in a subset of breast and prostate cancers.) International Conference on Fundamental Sciences: Biological and Chemical Sciences. 21-24 May 2001. Singapore

- **Tan JMM,** Chow VTK. (The novel, intronless, differentially-expressed human gene maps to chromosome 8q24.2, is amplified in a subset of breast and prostate cancers, and interacts with multiple proteins.) 5th NUS-NUH FOM ASM. 29-30 June 2001. Singapore

- **Tan JMM,** Chow VTK. (MOST-1, a novel intronless, differentially-expressed human gene maps to chromosome 8q24.2, is amplified in a subset of breast and prostate cancers, and interacts with multiple proteins.) Biomics Conference : Genes to proteins to structure to drugs. 19-21 November 2001. IBC Life Sciences. Frankfurt Germany