Characterisation of the HSP70-HSP90 organising protein gene and its link to cancer

A thesis submitted in the fulfilment of the requirements for the degree of

Master of Science in Biochemistry

Of

Rhodes University

By

Stacey Weeks

April 2015

Abstract

HOP (Heat shock protein 70/ Heat shock protein 90 organising protein) is a co-chaperone essential for client protein transfer from HSP70 to HSP90 within the HSP90 chaperone machine and has been found to be up-regulated in various cancers. However, minimal *in vitro* information can be found on the regulation of HOP expression. The aim of this study was to analyse the HOP gene structure across known orthologues, identify and characterise the HOP promoter, and identify the regulatory mechanisms influencing the expression of HOP in cancer. We hypothesized that the expression of HOP in cancer cells is likely regulated by oncogenic signalling pathways linked to *cis*-elements within the HOP promoter. An initial study of the evolution of the HOP gene speciation was performed across identified orthologues using Mega5.2. The evolutionary pathway of the HOP gene was traced from the unicellular organisms to fish, to amphibian and then to land mammal. The synteny across the orthologues was identified and the co-expression profile of HOP analysed. We identified the putative promoter region for HOP *in silico* and *in vitro*. Luciferase reporter assays were utilized to demonstrate promoter activity of the upstream region *in vitro.* Bioinformatic analysis of the active promoter region identified a large CpG island and a range of putative *cis*-elements. Many of the *cis*-elements interact with transcription factors which are activated by oncogenic pathways. We therefore tested the regulation of HOP levels by rat sarcoma viral oncogene homologue (RAS). Cancer cell lines were transfected with mutated RAS to observe the effect of constitutively active RAS expression on the production of HOP using qRT-PCR and Western Blot analyses. Additionally, inhibitors of the RAS signalling pathway were utilised to confirm the regulatory effect of mutated RAS on HOP expression. In cancer cell lines containing mutated RAS (Hs578T), HOP was up-regulated via a mechanism involving the MAPK signalling pathway and the ETS-1 and C/EBP$^{\beta}$ *cis*-elements within the HOP promoter. These findings suggest for the first time that Hop expression in cancer may be regulated by RAS activation of the HOP promoter. Additionally, this study allowed us to determine the murine system to be the most suited genetic model organism with which to study the function of human HOP.

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Master of Science of Rhodes University. It has not been submitted before for any degree or examination at any other university

…….
Ms Stacey Weeks, April 2015

Table of Contents

Table of Figures

Table of Electronic Figures

List of Tables

List of Abbreviations

17-AAG	17-N-Allylamino-17-demethoxygeldanamycin
17-DMAG	17-(Dimethylaminoethylamino)-17-demethoxygeldanamycin
ACTB	Beta cytoskeletal actin
AGE	Agarose gel electrophoresis
AR	Androgen receptor
ATP	Adenosine triphosphate
ATPase	Adenosine 5´-triphosphatase
BSA	Bovine serum albumin
C/EBP	CCAAT enhancer binding proteins
CDC37	Cell division cycle 37
c-Kit	v-Kit feline sarcoma viral oncogene homologue
CO_2	Carbon dioxide
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNAJC4	HSP40 homologue (subfamily C, member 4)
DP	Aspartic acid and proline rich domains
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGF	Endothelial growth factor
EGFP	Enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
ERK	Extracellular signal-regulated kinase
ERK2i	ERK2 inhibitor
ERα	Estrogen α-receptor
EtOH	Ethanol
FCS	Foetal calf serum
FERMT3	Fermitin family member 3
FGFR-1	Fibroblast growth factor receptor 1
FKBP2	FK506-binding immunophilin 2
FLRT1	Fibronectin leucine rich transmembrane protein
Flt-3	FMS-like tyrosine kinase-3

FTase	Farnesyltransferase
FTIs	Farnesyltransferase inhibitors
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAPs	RAS GTPase activating proteins
GDP	Guanine nucleotide diphosphate
GEFs	Guanine nucleotide exchange factors
GGTase	Geranylgeranyltransferase type I
GPI	Glycodylphosphatidylinositol
GRIA	Ionotropic AMPA glutamate receptor
GRP94	Endoplasmic reticulum HSP90
GTP	Guanine nucleotide triphosphate
HGF	Hepatocyte growth factor
HOP	Heat shock protein 70/ Heat shock protein 90 organising protein
HPRT1	Hypoxanthine phosphoribosyltransferase
HRP	Horseradish peroxidase
HSC70	Heat shock cognate protein 70
HSE	Heat shock element
HSF	Heat shock factor 1
HSP	Heat shock protein(s)
HSP40	Heat shock protein 40 kilodaltons
HSP70	Heat shock protein 70 kilodaltons
HSP90	Heat shock protein 90 kilodaltons
HSR	Heat shock response
JDP	J domain protein
JNK	Jun amino-terminal kinase
KCl	Potassium chloride
KH_2PO_4	Monopotassium phosphate
MACROD1	MACRO domain containing 1
MAPK	Mitogen activated protein kinase
MAPK	Mitogen-activated protein kinase
MEk	Meiosis-specific serine/threonine Kinase
MEKK1	Mitogen-activated protein kinase kinase kinase 1
$MgSO_4$	Magnesium sulphate
MKK4	Mitogen-activated protein kinase kinase 4

MKK7	Mitogen-activated protein kinase kinase 7
MLK3	Mitogen-activated protein kinase kinase kinase 11
MMP2	Matrix metalloproteinase-2
mTOR	Mammalian target of rapamycin
Na_2HPO_4	Disodium hydrogen phosphate
NaCl	Sodium chloride
NEF	Nucleotide exchange factor
NRT	No reverse transcriptase controls
NTC	No template controls
NUDT22	Nucleoside diphosphate-linked moiety X motif 22
O/N	Overnight
PBS	Phosphate buffered saline
PDGFR-β	Platelet derived growth factor receptor beta
PI3K	Phosphatidylinositol 3-kinase pathway
PKC	Protein kinase C
PrP^c	Prion protein
PSA	Penicillin-streptomycin-amphotericin
qRT-PCR	Quantitative real time polymerase chain reaction
RAF	Serine/threonine-specific kinase
RAL	RAS like
RAS	Rat sarcoma viral oncogene homologue
RFU	Relative fluorescence units
RLU	Relative light units
SDS	Sodium-dodecyl-sulphate
SDS-PAGE	Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis
Sp1	Stimulatory protein 1
STIP1	Stress inducible protein 1
TF	Transcription factor
TFBS	Transcription factor binding site
TNBC	Triple negative breast cancer
TPR	Tetratricopeptide repeat
TRAP1	Tumor necrosis factor receptor-associated protein 1
TRPT1	tRNA phosphotransferase
TSS	Transcription start site

VEGFB	Vascular endothelial growth factor B
VEGFR	Vascular endothelial growth factor receptor
αHFPA	Alpha hydroxyl farnesyl phosphoric acid

List of Symbols

>	Greater than
<	Smaller than
%	Percentage
bp	Base pairs
Ct	Cycle threshold
g	Grams
IC_{50}	Drug concentration that caused a 50% reduction in cell viability
kDa	Kilodaltons
L	Litres
M	Molar
mg	Milligrams
min	Minutes
ml	Millilitre(s)
mM	Milimolar
nm	Nanometre
nM	Nanomolar
^{o}C	Degree Celsius
P	Probability
™	Trade marked
U	Units
V	Volts
v/v	Volume per volume
w/v	Weight per volume
x g	Relative centrifugal force to gravity
α	Alpha
β	Beta
Δ	Delta
µ	Micro
µg	Micrograms
µl	Microlitres
µM	Micromolar

Acknowledgements

I would like to thank my supervisor, Dr. Adrienne Edkins, for her guidance, encouragement and support during my MSc. Her passion for her work is contagious and the excellence she demands of herself motivates her students to do the same, as it shows us that we too can do anything we put our minds to if we only believe in ourselves and put in the hard work. I would also like to thank my co-supervisor, Prof. Gregory Blatch, for his enthusiasm for and insight into this project. He has been invaluable in helping me keep my eyes on the bigger picture and thus the value of our contribution to the field. Furthermore, the members of my research unit, BioBRU, have played an immense role in this project as well. The discussion and debates shared in BioBRU have aided both the direction and completion of my work and the excitement with which it has been achieved. Both my family and my fiancé also deserve heartfelt gratitude. Their support and encouragement has seen me through the celebrations and frustrations of my MSc, and their belief in me has inspired me to pursue my dreams. I would also like to thank the DAAD-NRF and Rhodes University for funding.

Chapter 1

Literature Review

1.1 Molecular Chaperones

In order for the cell to function properly, cellular proteostasis needs to be maintained. For this, the integrity of the cellular protein machinery needs to be meticulously maintained. Problems such as protein denaturation, misfolding and aggregation constantly challenge the cell under physiological conditions, and even more so upon environments with extended levels of stress (Buchner, 1996; Beissinger & Buchner, 1998; Jolly, 2000; Tiroli-Cepeda & Ramos, 2011; Escusa-Toret *et al.*, 2013). Examples of such stresses are heat, oxidative stress, heavy metals and toxic substances (Heikkila *et al.*, 1982; Courgeon *et al.*, 1984; Yura *et al.*, 1984; Michel & Starka, 1986). A family of proteins known as molecular chaperones have the responsibility within the cell to facilitate the correct folding of proteins (known as clients) under both normal and stressful conditions (Buchner, 1996; Beissinger & Buchner, 1998; Jolly, 2000; Tiroli-Cepeda & Ramos, 2011; Escusa-Toret *et al.*, 2013). To prevent the aggregation potential of many of these folding proteins, the chaperones are required at stoichiometric ratios to decrease the comparative concentrations of non-native proteins (Kiefhaber *et al.*, 1991). Without these chaperone proteins, the damaging effects of stresses in the cell can be lethal. Extended cell stress can result in defects within the cytoskeleton, prevent the correct localisation of organelles, disrupt intracellular transport processes, deregulation of RNA splicing and mRNA functioning, and enhanced cell membrane permeability. When these changes do not lead to cell death, it leads to a stagnation of both cell growth and proliferation (Lindquist, 1980; Welch & Suhan, 1985; Nover *et al.*, 1989; Vogel *et al.*, 1995; Kampinga & Craig, 2010; Richter *et al.*, 2010). Furthermore, due to the dynamic functions of chaperones many of them require the assistance of accessory proteins known as co-chaperones; which partake in the functioning of chaperones (Caplan, 2003; Calderwood, 2013).

1.2 Heat Shock Proteins

A major subgroup of the molecular chaperone family are selected members of the heat shock proteins (HSP). In response to a variety of extra- and intra-cellular stresses or trauma, cells induce the synthesis of a series of HSPs. HSPs have adapted to protect the cells from the damage caused by abrupt environmental changes (Nicolet & Craig, 1989; Richter *et al.*, 2010). HSPs have the role of folding immature or denatured proteins, while others act as proteases, thereby preventing protein aggregation and retaining proteins in native conformations. This is accomplished by the binding of HSPs to the

hydrophobic surfaces of the partially folded client proteins (Wegele *et al.*, 2004; Song & Masison, 2005). Within the HSP family are two subfamilies of chaperones: Heat shock protein 90 (HSP90) and Heat shock protein 70 (HSP70). HSP90 and HSP70 are ubiquitous, essential and highly conserved proteins (Song & Masison, 2005). HSP90 and HSP70 are found to exist together in a heterocomplex with HSP40 and other accessory proteins, which together act as a protein folding and stabilisation machine (Czar *et al.*, 1994). HSP70 is required for *de novo* and stress related folding of clients, where the clients may be unfolded or misfolded proteins. HSP90 on the other hand is required for conformational regulation of clients and the clients are transferred to HSP90 already partially folded. Some of these clients are labile and require stabilisation by HSP90 until they are required within the cell (Buchner, 1996; Beissinger & Buchner, 1998; Jolly, 2000; Tiroli-Cepeda & Ramos, 2011; Escusa-Toret *et al.*, 2013).

1.2.1 – Heat Shock Protein 70

The human HSP70 group of proteins is comprised of 17 genes and 13 isoforms, some of which are constitutively expressed under normal physiological conditions (known as the heat shock cognate proteins, examples being HSPA5/HSP70-5/GRP78, HSPA8/HSP70-8/HSC70, HSPA9/HSP70-9/GRP75). Others are inducible stress related proteins (examples being HSPA1A/HSP70-1, DNAJB1/HSP40, HSPB1/ HSP27) (Tavaria *et al.*, 1996; Brocchieri *et al.*, 2008; Kampinga *et al.*, 2009). Certain HSP70 isoforms (such as HSPA8/HSC70) have housekeeping functions that aid in protecting proteins during specific cellular processes, such as protein transportation across membranes, proofreading protein structure and preventing the degradation of partially folded proteins during translation (Mayer & Bukau, 2005; Song & Masison, 2005). Whereas upon cellular stress, the expression of certain alternative HSP70 isoforms are induced (such as HSPA1A/HSP72) so as to prevent the aggregation of unfolding proteins and when necessary to refold aggregated proteins (Mayer & Bukau, 2005). Various regulatory proteins are known to be controlled by transient associations with HSP70 (such as steroid hormone receptors, kinases [e.g. CyclinB1] and transcription factors [e.g. c-Myc]) (Mayer & Bukau, 2005). Common traits to these interactions are an initial interaction with HSP70 during the co- or post-translational folding of the newly synthesised client proteins, a sensitivity of these interactions to the phosphorylation of the client proteins or the presence of specific ligands, and the

requirement of HSP90 in the complexes (Bukau *et al.*, 2000; Hartl & Hayer-Hartl, 2002; Wegele *et al.*, 2004; Mayer & Bukau, 2005).

HSP70 is comprised of two major functional domains, an N-terminal nucleotide binding domain containing the ATP binding domain, and a C-terminal client binding domain that can bind to the hydrophobic residues of partially unfolded proteins (Liberek *et al.*, 1991; Brocchieri *et al.*, 2008). These domains are joined by a third domain known as the linker domain that is important for the mechanism of HSP70 mediated folding. The HSP40 family are co-chaperones that aid the binding of HSP70 to the non-native client proteins (Laufen *et al.*, 1999). The HSP40s capture unfolded substrates via their substrate binding domains and thereby prevent the aggregation of the unfolded proteins while delivering these proteins to HSP70. HSP40s interact with HSP70s via exposed residues on the J-domain of HSP40s (Jiang *et al.*, 2007). Upon delivering the client protein to HSP70, HSP40 stimulates the ATPase activity of HSP70 via the J-domain, which converts HSP70 into the ATP bound state that has a higher affinity for client proteins (Mayer & Bukau, 2005; Jiang *et al.*, 2007; Kampinga & Craig, 2010). Nucleotide exchange factors (NEF) cause the dissociation of ADP (a result of the hydrolysed ATP) from HSP70. The dissociation of ADP allows for the rapid binding of ATP to HSP70 and the subsequent dissociation of the client protein. The client protein is released when refolded or, in the case of a subset of proteins, transferred to HSP90 as partially folded proteins (Wegele *et al.*, 2004; Mayer & Bukau, 2005; Kampinga & Craig, 2010).

1.2.2 – Heat Shock Protein 90

HSP90 isoforms are found in high concentrations within the cytoplasm (HSP90α/β), nucleus (HSP90α/β), the mitochondria (Tumor necrosis factor receptor-associated protein 1, TRAP1) and the endoplasmic reticulum (glucose related protein 4, GRP94), (Richter *et al.*, 2010). Two cytosolic isoforms of the protein exist, HSP90α which is the inducible form and HSP90β which is the constitutive form (Csermely *et al.*, 1998; Li *et al.*, 2012). HSP90β is an essential gene in the mouse, while HSP90α is not (Voss *et al.*, 2000; Grad *et al.*, 2010). HSP90 is also found on the cell surface and is secreted into the extracellular space for certain cell types (Eustace *et al.*, 2004; Sidera & Patsavoudi, 2008; Li *et al.*, 2012). The client proteins of HSP90, of which there are over 300, are involved in almost all cellular processes, from signal transduction, cell cycle

progression and transcriptional regulation. However, they are predominantly comprised of those involved in signalling pathways such as steroid receptors, kinases and transcription factors (TF) (Echeverría *et al.*, 2011; Li *et al.*, 2012). Little is known about how HSP90 recognizes these clients, however, a study by Taipale *et al.* (2012) revealed that HSP90 kinase client recognition is determined by thermodynamic parameters and not sequence motifs. They identified that an additional accessory protein, known as CDC37, is required to recognise kinase families and there upon HSP90 bound kinase clients within that family which were intrinsically unstable.

HSP90 functions as a homodimer; each monomer is comprised of an N-terminal domain which contains an ATP-binding pocket and motifs encoding co-chaperone interaction regions. As with HSP70, the N-terminal domain of HSP90 is connected to the middle domain (M-domain) by a charged linker. The M-domain contains the binding sites for various clients and co-chaperones. The C-terminal domain of HSP90 contains a dimerization motif (Makhnevych & Houry, 2012). The N terminal dimerization of HSP90 and client binding is regulated by ATP binding and hydrolysis, while the C terminal dimerization of HSP90 is constitutive. The release of ADP drives conformational changes in the protein thereby navigating it through high and low affinities for the impartially folded client proteins (Prodromou *et al.*, 1997). However, the ATPase cycle is dependent on the involvement of co-chaperones that affect the conformational dynamics of the HSP90 complex and thus the rate at which the ATPase activity occurs (Trepel *et al.*, 2010).

Various co-chaperones interact with cytosolic HSP90 isoforms, modulating different aspects of the functioning of the HSP90 chaperone machine. Co-chaperones regulate HSP90 through three means: by determining the rate of ATP hydrolysis, determining the conformational flexibility of HSP90 or by binding specific client proteins to deliver them to HSP90. Additionally, co-chaperones aid HSP90 functioning by stimulating the ubiquitination or dephosphorylation of client proteins (Panaretou *et al.*, 2002; Meyer *et al.*, 2004; Roe *et al.*, 2004; McLaughlin *et al.*, 2006; Forafonov *et al.*, 2008; Zhang *et al.*, 2008a; Makhnevych & Houry, 2012). HSP90 co-chaperones fall into two groups, tetratricopeptide repeat (TPR) (e.g. HOP, PP5, Unc45, FKBP52/51, CYP40) and non-TPR containing co-chaperones (e.g. p23, AHA1, and CDC37). p23, cell division cycle 37 (CDC37) and the activator of the HSP90 ATPase protein (AHA1) regulate the

ATPase cycle of HSP90. p23 partially inhibits the ATPase activity of HSP90 by securing HSP90 in an ATP-dependant, stabilized conformation. This conformation of HSP90 has a higher affinity for client proteins, thus p23 enables HSP90 to bind to client proteins (McLaughlin *et al.*, 2006). The C-terminal of CDC37 interacts with the N-terminal of HSP90 and thereby inhibits the ATPase activity of the chaperone by fixing the chaperone in an open conformation. Simultaneously, the N-terminal of CDC37 interacts with protein kinases. By inhibiting the ATPase activity of HSP90 it allows specifically for the binding of kinase client proteins to HSP90 (Roe *et al.*, 2004; Calderwood, 2013). HOP also inhibits ATPase activity of HSP90 and is important for communication between HSP70 and HSP90, however its function will be addressed in greater detail at a later point. AHA1, on the other hand stimulates the ATPase activity of HSP90. The binding of the N-terminal of AHA1 to the middle domain of HSP90 results in a conformational shift of the middle domain of HSP90. This shift that enables the N-terminal domain of HSP90 to interact with ATP, thereby accelerating the progression of the ATPase cycle (Meyer *et al.*, 2004; Li *et al.*, 2012). The co-chaperone CHIP is a ubiquitin E3 ligase that binds to HSP90 (and HSP70) in order to degrade damaged or misfolded client proteins (Xu *et al.*, 2002; Calderwood, 2013). Thus CHIP acts as a protein quality control co-chaperone.

1.3 The Heat Shock Response

One major client of HSP90 is the heat shock factor 1 (HSF1), which is a regulator of the heat shock response (HSR). In the absence of cell stress, HSF1 is maintained in an inactive monomeric form through interaction with HSP90 and HSP70 in the cytoplasm (Fig.1) (Zou *et al.*, 1998; Richter *et al.*, 2010). During cellular stress, damaged proteins compete with HSF1 for HSP90/HSP70 binding. This causes the dissociation of HSF1 from HSP90, which consequently allows the transcription factor (TF) to homotrimerise, undergo activation by hyperphosphorylation and translocate to the nucleus (Fig.1) (Sarge *et al.*, 1993; Richter *et al.*, 2010). Two phosphorylation sites within HSF1 are of particular importance for the transactivation function of HSF1, Ser230 and Ser326 (Holmberg & Hietakangas, 2001; Guettouche *et al.*, 2005; Akerfelt *et al.*, 2010). The trimeric HSF1 then binds to the promoters of various HSPs to induce the up-regulation of their expression in response to the cellular stress (Zou *et al.*, 1998; Guo *et al.*, 2001; Mizrak *et al.*, 2006). The *cis*-elements in the target gene promoters for HSF are known as heat shock elements (HSE) (Fig.1). Binding of HSF to an HSE leads to a change in

chromatin organization within HSP promoters, and an interaction with various components of the chromatin remodelling machinery and the basal transcriptional machinery (Morimoto, 1998). HSF1 activity is modulated by a negative feedback mechanism when the cell returns to a normal state of functioning and proteostasis. The excess of unemployed HSP70 and HSP90 bind HSF-1 and thus downregulate the transcriptional activator (Fig.1) (Gibbs *et al.*, 1993; Mizrak *et al.*, 2006; Richter *et al.*, 2010).

The heat shock factor family is comprised of 4 genes (HSFs 1, 2, and 4; as well as a unique avian HSF3), all of which are ubiquitous except for HSF3 (Morimoto, 1998). HSF1, although constitutively expressed in most tissues, is activated in response to elevated temperatures, heavy metals, amino acid analogs and oxidative stress whereas HSF2 is activated in response to hemi-induced differentiation of erythroleukemiac cells or hemin treatment (Morimoto, 1998; Trinklein *et al.*, 2004; Ostling *et al.*, 2007; Akerfelt *et al.*, 2010). HSF2 is constitutively expressed and was initially thought to only be activated in cells involved in differentiation and development (Goodson *et al.*, 1995). However, HSF2 has been found to be important for both constitutive and stress-inducible expression of HSP (Wilkerson *et al.*, 2007; Sandqvist *et al.*, 2009). It is recruited alongside HSF1 during heat stress; although it depends strictly on the presence of HSF1 to relocate to the nucleus (Trinklein *et al.*, 2004; Ostling *et al.*, 2007). HSF2 aids the regulation of the HSR by forming transcriptionally active heterotrimers with HSF1 (Schuetz *et al.*, 1991; Sarge *et al.*, 1993; Akerfelt *et al.*, 2010). Such a functional relationship between the two TFs offers an efficient way to control gene expression and orchestrate the differential upstream signalling and target gene networks (Akerfelt *et al.*, 2010). The third HSF works in conjunction with HSF1 in avian systems to regulate the HSR (Nakai *et al.*, 1995). HSF4 is expressed in human heart, brain, skeletal muscle, and pancreas tissue; and acts to repress the expression of HSP (Nakai *et al.*, 1997; Pirkkala *et al.*, 2001).

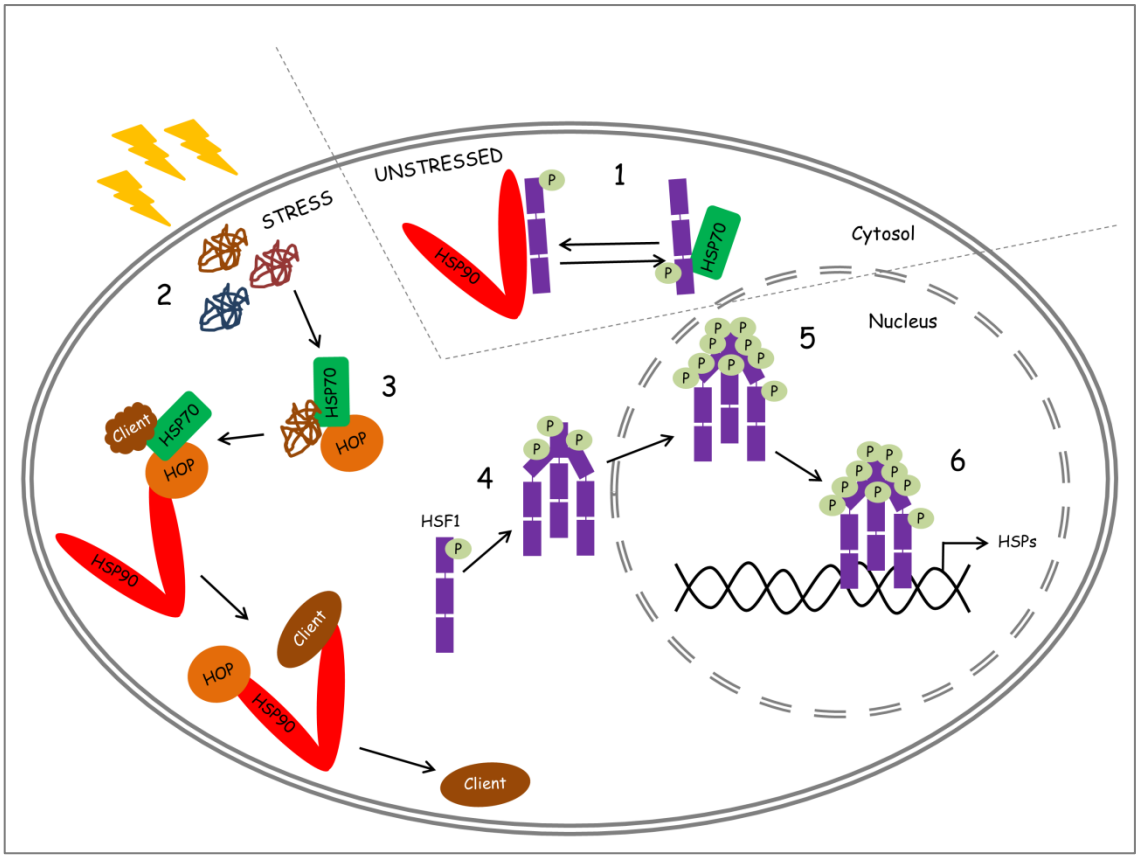


Figure 1: The Heat Shock Response.

A schematic diagram illustrating the regulation of the heat shock response in eukaryotic cells. (1) Under physiological conditions within the cell, HSF1 is bound in an inactive state by the molecular chaperones HSP90 and HSP70. (2) Upon environmental or physiological stress, proteins begin to aggregate. (3) HSP70 and HSP90 are recruited to the damaged proteins and, with the aid of the co-chaperone HOP, refold the client proteins to a mature conformation. (4) During this process, HSF1 has been released from the chaperone complex and is free to trimerise. (5) Once trimerised, the HSF1 complex translocates to the nucleus where it is hyperphosphorylated. The hyperphosphorylation activates the DNA binding capacity of the complex, (6) whereupon HSF1 binds to heat shock elements (HSE) within the promoters of heat shock protein genes. This stimulates an upregulation in the expression of heat shock proteins in order to compensate for the stressed conditions of the cell. When proteostasis is achieved, the HSF1 trimer dissociates and the monomers are bound by HSP90 and HSP70 once more. Adapted from Richter *et al.* (2010) and Morimoto (1998).

1.4 HSP90 as a drug target

Cancer is a disease which has haunted humans for centuries. It is a complex disease that arises as a result of genomic instability and tumour-promoting inflammation (Stratton *et al.*, 2009; Hanahan & Weinberg, 2011). Mutations in the human genome either cause genes to have a dominant gain of function which are then referred to as oncogenes, or they cause genes which were previously considered as tumour suppressor genes to have a recessive loss of function (Hanahan & Weinberg, 2000, 2011). The variety of mutational combinations that are possible has led to the identification of more than one hundred distinct cancers, with further subtypes within each. Cancers are characterised

by deregulated cell physiology that affects both cell proliferation and homeostasis (Hanahan & Weinberg, 2000, 2011). These deregulations are common across all cancers and have become known as the hallmarks of cancer. They are classified into eight alterations within the cell which provide the cell with eight new capabilities: cells are able to evade apoptosis, they are insensitive to growth suppressors, are self-sufficient in growth signals, they undergo sustained angiogenesis, have limitless replicative potential, and they gain the ability to invade tissues and metastasise (Hanahan & Weinberg, 2000, 2011). Additionally, malignant cells are able to avoid immune destruction and have deregulated cellular energetics (Hanahan & Weinberg, 2011). Tumorigenesis occurs by the combination of mutational events within cells that lead to the alteration of multiple signalling pathways. These mutations may occur in the form of point mutations, rearrangements, translocations, deletions and amplifications of genetic material (Stratton *et al.*, 2009; Hudson *et al.*, 2010; Cairns *et al.*, 2011). Some of the effects of these mutational events is deregulation of gene expression and altered chromatin structure (Stratton *et al.*, 2009). Deregulated gene expression alters the balance between tumour-suppressive and tumour progressive gene networks, thereby indirectly enabling cells to proliferate, evade apoptosis and metastasize (Kittler *et al.*, 2013).

Cancer cells have been shown to use the HSP90 chaperone machine to protect mutated and/or overexpressed proteins from degradation or misfolding, thereby stabilizing tumourigenic cells. Consequently HSP90 has been seen as a critical facilitator of cancer cell survival and oncogene addiction (Bagatell & Whitesell, 2004; Sangster *et al.*, 2004; Mizrak *et al.*, 2006; Trepel *et al.*, 2010). In a study conducted by (Kubota *et al.*, 2010), colon cancer tissues were shown to have an up-regulation of components of the HSP90 chaperone machine, specifically HOP, HSP90 and HSC70. Furthermore, several client proteins (such as B-RAF, fibronectin, FKBP52, CDK4) of HSP90 are known to have an involvement in cell proliferation and tumour progression (Scammell *et al.*, 2003; Wegele *et al.*, 2004; da Rocha Dias *et al.*, 2005; Zhao & Houry, 2005; Hunter *et al.*, 2014). Some client proteins have been identified as tumour-promoting signal transducer proteins; thus making chaperones targets for anticancer drugs because of their role in helping to promote tumour cell adaptation to unfavourable environments (Prodromou *et al.*, 1997; Kubota *et al.*, 2010; Horibe *et al.*, 2011). Cells with high levels of HSPs were found to have resistance to apoptosis, due to HSPs directly interacting with the

apoptotic molecules of the cell, as well as blocking the caspase-independent apoptosis pathways (Sreedhar & Csermely, 2004). Alternatively depleting HSP levels in tumour cells induces apoptosis.

Numerous HSP90 inhibitors have been designed as drug targets for cancer treatment. To date there are ninety seven worldwide clinical trials that have been completed, terminated or are in progress for HSP90 drugs as cancer treatments (U.S. National Institutes of Health, 2014). However, many HSP90 inhibitors have failed as effective inhibitors because they induce the HSR (Fukuyo *et al.*, 2010). Many N-terminal HSP90 inhibitors (such as geldanamycin, 17-AAG/DMAG, and radicicol) cause the dissociation of the HSP90 chaperone complex which has a two-fold consequence (Fukuyo *et al.*, 2010). Firstly, the co-chaperone CHIP then preferentially binds to HSP90 causing the ubiquitination of HSP90 client proteins and their subsequent degradation (Schulte *et al.*, 1997; Xu *et al.*, 2002; da Rocha Dias *et al.*, 2005; Grbovic *et al.*, 2006). This is desirable in cancer as many HSP90 client proteins are those supporting the oncogenic process. Secondly, upon the dissociation of the HSP90 chaperone complex, HSF1 is released and goes on to trimerise and induces the expression of additional HSPs such as HSP70 (Fukuyo *et al.*, 2010). This is a negative side effect with respect to cancer treatment. The induction of HSP70 expression has been found to attenuate the cell death effects of the HSP90 inhibitors. However, it has been found that by reducing the expression of HSC70 and/or HSP72 in cancer cells, the degradation of HSP90 clients and apoptosis of the tumour cells is induced (Powers et al., 2008; Davenport et al., 2010).

1.5 The HSP70/HSP90 organising protein

The HSP90 chaperone machine is comprised of various co-chaperones, one of which is the HSP70/HSP90 organising protein (HOP), otherwise known as p60 or the stress inducible protein 1 (STIP1). HOP is a TPR co-chaperone or adaptor protein that is essential for the full functioning of HSP90 with certain client proteins (Hernández *et al.*, 2002; Trepel *et al.*, 2010). Some co-chaperones are required for the mediation between clients and chaperones; and such is the case for HOP. HOP is specifically responsible for catalysing the client transfer from HSP70 to HSP90, by coordinating the cycle of substrate binding and release through modulating the conformational dynamics of the chaperones. In doing so, HOP facilitates the folding of nascent polypeptides, the

productive assembly of multimeric protein complexes, and the disassembly of proteins specific to the HSP70 and HSP90 chaperones (Hernández *et al.*, 2002; Caplan, 2003; Song & Masison, 2005; Trepel *et al.*, 2010). Alternate functions for HOP, presumed to be independent of HSP90 and HSP70 have also been identified. HOP forms a signalling complex with prion proteins on the cell surface. HOP was shown to be a ligand for the glycodylphosphatidylinositol (GPI)-anchored cell surface sialoglycoprotein homologue (Prion Protein, PrP^c); a cell surface ligand for cellular prions (Zanata *et al.* 2002; Martins *et al.* 1997; Lima *et al.* 2007; Arantes *et al.* 2009; Hajj *et al.* 2013). Beraldo *et. al.* (2013) went on to show that extracellular HOP plays a role in neuroprotection specifically against ischemic injury, by preventing neuronal death upon PrP^c activation of extracellular HOP. Recent research has revealed further insight into the roles of HOP within the cell. HOP has been shown to have independent ATPase activity, suggesting that HOP may have a role as a chaperone or at least have additional roles to its HSP70-HSP90 co-chaperone role (Yamamoto *et al.*, 2014). Furthermore, it was recently shown that the co-chaperone activity of HOP can be inhibited in *vivo* by the phosphorylation of multiple sites within the protein. Phosphorylation of human HOP decreases the proteins affinity for both HSP70 and HSP90 (Röhl *et al.*, 2014).

HOP is ubiquitously expressed and is induced upon cellular stress (Nicolet & Craig, 1989). It is located predominantly within the cytoplasm however, HOP has also been identified within the nucleus (Longshaw *et al.* 2004), the Golgi apparatus (Honore *et al*. 1992), in the extracellular environment or associated with the cell membrane (Hajj *et al.* 2013). HOP deletion mutants in yeast caused a clear reduction in the functioning of HSP70 and HSP90 as well as their client proteins (GR and v-Src) (Chang *et al.*, 1997; Song & Masison, 2005). Furthermore, cells become hypersensitive to environmental changes upon the depletion of HOP, due to the consequent reduction of HSP90 functioning (Song & Masison, 2005; Li *et al.*, 2012). HOP was also found to have an essential role in mammalian embryonic development and cell survival in mice, whereupon HOP knockout in mice caused embryonic death (Beraldo *et al.*, 2013). Additional co-chaperones could not compensate for the loss of HOP in murine embryos suggesting unique roles for differing co-chaperones during different stages of development.

The assembly of this multimeric HSP90 chaperone machine is dependent on all three independent TPR domains within HOP. The TPR motif is a degenerate 34-amino acid sequence that forms the scaffold for the various protein interactions (Fig.2) (Das *et al.*, 1998; Smith, 2004). HOP has 9 TPR motifs that are arranged into the three domains (TPR1 and TPR2A, TPR2B) each of which comprise 3 TPR motifs (Prapapanich *et al.*, 1998; Nelson *et al.*, 2003). HOP also contains two DP domains (aspartic acid and proline rich domains) which are otherwise known as STI1 domain sequences as they contain the heat shock chaperone binding motif (Fig.2) (Prapapanich *et al.*, 1998; Nelson *et al.*, 2003). HOP binds to the C terminal regions (terminating with the amino acid sequence EEVD) of both HSP90 and HSP70 simultaneously and thereby creates a platform for the transfer of client proteins between them (Scheufler *et al.*, 2000b; Carrigan *et al.*, 2004; Makhnevych & Houry, 2012). Furthermore, HOP TPR domains can discriminate between the EEVD motifs within HSP90 and HSP70 based on the adjacent residues. The TPR domains recognise the MEEVD and GPTIEEVD motifs in HSP90 and HSP70 respectively, ensuring that TPR1 binds to HSP70 while TPR2A binds to HSP90 (Scheufler *et al.*, 2000a; Odunuga *et al.*, 2003). Additionally, there is possibly a second binding site for HSP70 on HOP in the DP2 domain (Fig.2) (Russell *et al.*, 1999; Carrigan *et al.*, 2004; Onuoha *et al.*, 2008; Li *et al.*, 2012).

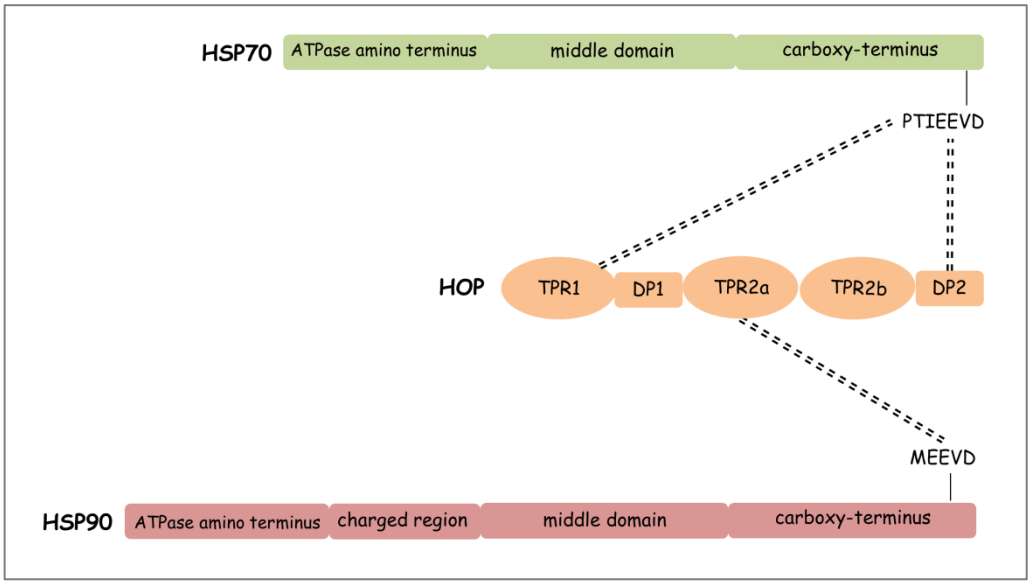


Figure 2: HOP acts as a scaffold for interaction between HSP70 and HSP90.

Schematic diagram of the regions of the co-chaperone HOP that bind to the EEVD motifs found on the molecular chaperones HSP70 and HSP90. Dotted lines illustrate the interactions between the proteins. These interactions between HSP70-HOP-HSP90 allow for the transfer of client proteins from HSP70 to HSP90.

TPR1 and TPR2A mediate the interactions of HSP90 and HSP70 with substrate proteins through an anchoring function (Odunuga *et al.*, 2004; Horibe *et al.*, 2011). Upon the binding of HOP, the HSP90 ATPase activity is inhibited and HSP90 is stabilised in an open conformation due to a rotation in the junction of the middle and C-terminal domains of HSP90 (Southworth & Agard, 2011; Li *et al.*, 2012). In this conformation, client binding residues have converged so that HSP90 is positioned for client loading by HSP70 with an accessible client binding surface. Simultaneously, the N-terminal domain rotates almost 90^{o} and is held in an hydrolysis-incompetent state (Southworth & Agard, 2011). This conformation of HSP90:HOP allows for the interaction with HSP70. HSP70 binds to TPR1 but additionally makes contact with HSP90 (Southworth & Agard, 2011; Kirschke *et al.*, 2014). HSP70 facilitates client transfer to HSP90 by partially unfolding the client protein. HSP90 binds ATP, the hydrolysis of the ATP contributes to the client transfer by a means of coupling the ATP cycles of HSP90 and HSP70 (Kirschke *et al.*, 2014). The coupling of the ATP cycles was visualised by Kirschke *et al.* (2014) using cyroelectron microscopy, whereby they identified points of contact between HSP70 and HSP90 in a glucocorticoid receptor (GR):HSP70:HSP90:HOP complex. Alvira *et al.* (2014) identified that the client protein binds to the side of the HSP90 dimer opposite the HOP attachment site on the dimer. HSP90 then reactivates the client (in this case GR) by refolding it upon the ATP hydrolysis of the chaperone and the co-chaperone HOP (Kirschke *et al.*, 2014). HSP90 moves to a closed conformation upon the binding of ATP, this interaction weakens the interaction of HOP:HSP90 and promotes the exit of HOP from the complex. The refolded client protein is then released after ATP hydrolysis (Li *et al.*, 2012).

1.6 HOP as a novel drug target

While the role of HSP90 in cancer is well established, it is now appreciated that co-chaperones of HSP90, including HOP, may contribute to cancer cell proliferation, migration and drug resistance (Erlich *et al.*, 2007; Forafonov *et al.*, 2008; Holmes *et al.*, 2008; Smith *et al.*, 2009; Horibe *et al.*, 2011). HOP has been identified as overexpressed in certain cancers compared to normal cell equivalents. HOP expression was linked to invasive potential in seven human pancreatic cancers (Walsh *et al.*, 2011); a feature also demonstrated in breast cancer cells. There was an increase in HOP expression in the cancerous Hs578T breast cell line when compared to the non-

malignant Hs578Bst cell line, which was derived from the same patient (Willmer 2011, unpublished data). Kubota and colleagues (2010), showed HOP levels were up-regulated in colonic carcinoma tissue samples compared to non-tumour tissues obtained from the same colonic carcinoma patients (Kubota *et al.*, 2010) and an increase in HOP mRNA and protein expression was detected upon transformation of Rat1 fibroblast cells with the oncogene RAS (van der Spuy 2000, unpublished data). Additionally, HOP has been found in a constitutive complex with HSP90 in cancer cells, whereas complex formation is only induced under specific conditions in normal cells (Kamal *et al.*, 2003). Addition of recombinant HOP to glioma cells induced cell proliferation (Erlich *et al.*, 2007), while the disruption of the HOP-HSP90 interaction reduced proliferation of a range of cancers, including breast, human lung and kidney and sensitised cells to inhibitors against HSP90 (Horibe *et al.*, 2011; Pimienta *et al.*, 2011). Data from knockdown studies have demonstrated that HOP regulates cancer cell migration by numerous mechanisms, including inhibition of matrix metalloproteinases-2 (MMP2) and regulation of components of the cytoskeleton (actin and tubulin) (Walsh *et al.*, 2011; Willmer *et al.*, 2013). The presence of HOP in the extracellular matrix has been linked to the invasive properties of these cells by Walsh *et al.* (2011). While STI1 deletion mutants were not lethal in yeast, it led to a clear reduction in molecular chaperone and client protein function (Chang *et al.*, 1997; Hernández *et al.*, 2002). Furthermore, HOP gene knock out was embryonically lethal in mice (Beraldo *et al.*, 2013). Collectively, these results suggest a role for HOP in malignant cells.

The dependence of cancer cells on molecular chaperones has been deemed the soft spot of cancer (Calderwood, 2013). However, depletion of HSP90 has not proved to be the best form of inhibition due to the consequent induction of an antiapoptotic HSR (Fukuyo *et al.*, 2010). It has been suggested that targeting certain co-chaperones may provide another route to inhibit oncogenic HSP90 chaperoning (Smith & Workman, 2009). The functioning of HSP90 can be hindered by targeting the co-chaperones, such CDC37 or p23, so as to alter its contribution to that of the chaperone's activity (Forafonov *et al.*, 2008; Smith & Workman, 2009; Trepel *et al.*, 2010). This would inhibit the functioning of the chaperones but prevent an induction of the HSR. Various attempts have already been made in targeting the co-chaperones of HSP90. Celastrol was found to disrupt the HSP90/CDC37 complex and Gendunin inhibited the HSP90 co-chaperone p23 (Zhang *et al.*, 2008b; Salminen *et al.*, 2010; Patwardhan *et al.*, 2013).

Unfortunately, Celastrol was found to lack specificity but Gendunin was found to promote the apoptosis of cancer cells. Although Celastrol treatment was not ideal an alternative was found whereby silencing CDC37 or removing the C-terminal of CDC37 (thereby destroying the HSP90 binding capacity of CDC37) diminished and destabilised the kinase client protein binding as well as sensitized the cells to HSP90 inhibitors (Smith *et al.*, 2009, 2013). Additionally, in the absence of p23, mammalian cells were found to be hypersensitive to HSP90 inhibitors (Forafonov *et al.*, 2008). The knockdown of AHA1 caused an increase in the sensitivity of cancer cells to 17-AAG partly due to a 3-fold increase in apoptosis (Holmes *et al.*, 2008). Thus considering the complexities and difficulties that have arisen in designing effective chaperone inhibitors, we support the concept of targeting the co-chaperones and suggest that HOP specifically represents a novel drug target for anticancer therapeutics.

In the case of developing drugs that specifically target HOP, little research has yet been done. However, a study by Horibe and colleagues (2011) revealed that a TPR-modelled peptide inhibited the interaction of HOP with the TPR2A domain of HSP90. The result of this inhibition was the selective cell death of cancerous cells while normal cells remained viable. Furthermore Pimienta and colleagues (2011) developed the compound C9 (1,6-dimethyl-3-propylpyrimido[5,4-e][1,2,4]triazine-5,7-dione) which is an inhibitor of the interaction of HOP with HSP90. C9 caused the cell death of various breast cancer cell lines without inducing the transcription of HSP70. Both of the above mentioned inhibitors, by competitively binding to HSP90, prevent the transfer of client proteins from HSP70 to HSP90 via HOP and thereby indirectly inhibit a function of HSP90 and ensure client proteins of HSP90 remain inactivated. Since the inhibition of this process doesn't induce the prosurvival effect of the transcriptional up-regulation of HSP70 (Pimienta *et al.*, 2011), as do many N-terminal ATP binding site HSP90 inhibitors, the inhibition of the interaction between HOP and HSP90 presents itself as an attractive alternative for the indirect inhibition of HSP90.

1.7 Knowledge Gap and Conclusive Remarks

There is a growing body of evidence linking the co-chaperone, HOP, to the malignant nature of tumours. These data have suggested that HOP may be a key regulator of oncogenesis and may represent a novel drug target for the treatment of cancer.

However, the success rate of the clinical trials of investigational drugs within oncology is the lowest of all trials drugs in clinical development (Kamb *et al.*, 2007). It is thus essential to understand the drug target of choice. Consequently it is vital to understand the factors influencing the gene expression of HOP. To do so, one would be required to locate and characterise the core promoter of the HOP gene. It has been stated that promoters are the best-characterised transcriptional regulatory sequences in genomes due to the predictability of their location immediately upstream of the transcription start site (TSS) (Hiang *et al.*, 2010). In the case of HOP, there is minimal *in vitro* or *in vivo* information to be found on the DNA sequence encoding the promoter. Furthermore, even less is known about the regulatory mechanisms that control the expression of HOP. Since HOP is known to be up-regulated in human cancers it is possible that oncogenic signalling pathways may be involved in the regulation of the HOP promoter.

1.8 Hypothesis and Objectives

This investigation hypothesizes that the HOP core promoter containing the essential transcription factor binding sites and consensus sequences for basal activity will be located 100 bp upstream from the transcription start site (-600 to -100), preceded by proximal region involved in lesser regulatory reactions. The expression of HOP in cancer cells is likely regulated by oncogenic signalling pathways linked to *cis*-elements within the HOP promoter. Hence the aim of this study was to investigate the expression of HOP in cancerous cells. The objectives of this study were:

1. Analyse the HOP gene structure across known orthologues;
2. Identification and characterisation of the HOP promoter;
3. Identification of the regulatory mechanisms influencing the expression of HOP in cancer.

Chapter 2

Materials and Methods

2.1 Materials

Hs578T (cat no.: 86082104) human metastatic breast cancer cell line was purchased from American Tissue Culture Collection (ATCC), as was the MEF-1 SV40 transformed mouse embryonic fibroblast cell line (cat. No.: CRL-2214™). MCF-7 non-metastatic breast cancer cell line, purchased from ATCC (cat no.: HTB-22), was obtained from A/Prof Sharon Prince (University of Cape Town). HeLa (cat no.: CCL-2) human cervical cancer cell line, purchased from ATCC, was obtained from Prof Rosemary Dorrington (Rhodes University). All general reagents were purchased from Sigma-Aldrich (Germany) and Saarchem (Merck; South Africa). Tissue culture media and reagents, including foetal calf serum [FCS], Dulbecco's Modified Eagle Medium [DMEM] with GlutaMAX™-I and 10 X Trypsin- Ethylenediaminetetraacetic acid (EDTA) were from Gibco, (Invitrogen, UK) and Biowhittaker (UK). Insulin was from Novorapid. Tissue Culture plasticware was from Corning Incorporated (USA). Western Blotting power pack, Hybond Support Nitrocellullose and ChemiDoc™-XRS were from Bio-Rad (UK). X-tremeGENE HP DNA Transfection Reagent (cat no.: 06366236001) was purchased from Roche (Switzerland). Primary antibodies were purchased from Cell Signalling (USA); Histone H3 (cat no.: 9715L), Pp44/42 MAPK (phospho-Erk1/2) (cat no.: 4370S), pAKT (cat no.: 4060L), pFAK (cat no.: 3281S), RhoA (cat no.: 9968S), or Santa Cruz Biotechnology (USA); pJNK (cat no.: sc-6254), Pp38 (cat no.: sc-166182), C/EBP$^{\beta}$ (cat no.: sc-150), ETS-1 (cat no.: C-20), while HOP (cat no.: SRA-1500) was purchased from Assay Designs (USA). αTubulin (cat no.: ab7291) and GAPDH-HRP (cat no.: ab185059) primary antibodies as well as HRP conjugated donkey anti-rabbit secondary antibody (cat no.: ab16284) and HRP conjugated goat anti-mouse secondary antibody (cat no.: ab97110), DyLight 550 donkey anti-mouse secondary antibody (cat no.: ab96876) and Dylight 488 donkey anti-rabbit secondary antibodies (cat no.: ab96891) were from Abcam (UK).

2.2 Methods

2.2.1 - Maintenance of cancer cell lines

Hs578T breast cancer cells were maintained in Dulbecco's Modified Eagle Media (DMEM) supplemented with GlutaMAX™-I, 10% (v/v) Foetal Calf Serum (FCS), penicillin-streptomycin-amphotericin (PSA; 100 U/ml) and 0.3 U/ml insulin. HeLa

cervical cancer cells were maintained in DMEM supplemented with GlutaMAX™-I, 10% (v/v) FCS and PSA (100 U/ml). The MCF-7 breast cancer cells were maintained in DMEM with GlutaMAX™-I, 5% (v/v) FCS, PSA (100 U/ml). MEF-1 SV40 transformed mouse embryonic fibroblast cells were maintained in DMEM supplemented with GlutaMAX™-I, 10% (v/v) FCS and PSA (100 U/ml). All mammalian cell lines were maintained at 37 ^{o}C with 9% CO_2.

2.2.2 - Bioinformatic analysis of HOP nucleotide and protein sequences

Sequences for the genomic context analysis and comparison were obtained from Ensembl, UniProtK, HomoloGene and NCBI databases. Accession numbers for sequences were as follows: *Homo sapiens* (NT_033903.7, NM_006819.2), *Mus musculus* (NT_082892.2, NM_016737.2), *Danio rerio* (NM_001007766.2), *Saccharomyces cerevisiae* (NM_001183446.1), *Latimeria chalumnae* (ENSLACT00000019245, H3BB41), *Xenopus tropicalis* (ENSXETG00000007189), *Drosophila melanogaster* (NT_033779.4, NM_058006.4), *Caenorhabditis elegans* (O16259, NM_070921.6), *Pan troglodytes* (XP_508521.2), *Macaca mulatta* (XP_001115412.2), *Rattus norvegicus* (NP_620266.1), *Canis lupus* (XP_854960.1), *Bos taurus* (NP_001030569.1), *Anopheles gambiae* (XP_319365.4), *Schizosaccharomyces pombe* (NP_588123.1), *Magnaporthe oryzae* (XP_364135.1), *Kluyveromyces lactis* (XP_451313.1) and *Ashbya gossypii* (XP_319365.4). Multiple sequence alignments were performed using T-Coffee (Notredame *et al.*, 2000) and pairwise global alignments were performed with BioEdit (Hall, 1999) using default parameters for both. Protein structure was analysed by SMART (Schultz *et al.*, 1998) using the default parameters to identify conserved protein domains. Nucleotide and protein sequences were aligned first by MUSCLE (Edgar, 2004) or Clustal Omega (McWilliam *et al.*, 2013) before the phylogenetic maximum likelihood tree was constructed using MEGA5.2 (Hall, 2013) with bootstrap replicates of 1000 for protein sequences. The tree was compared against trees generated by the neighbour-joining phylogenetic method. Alternate transcripts were collected from ENSEMBL (www.ensembl.org/Homo_sapiens). Co-expression data was gathered from Gemma (Zoubarev *et al.*, 2012) with a stringency of 3 ensuring that genes with at least 3 supporting cases and no negative support were considered. Differential expression data was gathered from Gemma. C0XPRESdb (version 6) (Obayashi *et al.*, 2008) was used

in conjunction with Gemma for co-expression analyses. Cytoscape 3 (Cline *et al.*, 2007) was utilised to construct networks between co-expressed genes or protein-protein interaction networks. The HSP90 Chaperone Machine Interactome Database (Echeverría *et al.*, 2011) was used to obtain the HOP protein-protein interaction data.

2.2.3 - Transcription factor binding site prediction

The presence of transcription factor binding sites in the putative HOP promoter sequence was predicted *in silico* using Cister (Wingender *et al.*, 2000; Frith *et al.*, 2001) using a hidden Markov statistical model, TFSearch (Akiyama, 1995; Heinemeyer *et al.*, 1998) using a cut-off threshold of 85.0, ALIBABA (Heinemeyer *et al.*, 1998) with a cut-off of 80% matrix conservation, and Promo (Messeguer *et al.*, 2002; Farré *et al.*, 2003) using a dissimilarity < 5.0 as a cut-off. CpG islands identified by MethPrimer (Li & Dahiya, 2002) the criteria used being a frequency of observed CG dinucleotides / frequency of expected CG dinucleotides > 0.6; GC % > 50; and length > 100.

2.2.4 - Cloning of HOP promoter luciferase reporter construct

The 1535bp putative promoter sequence was chemically synthesised into a pUC57 vector by GenScript (USA), to create pSAW. The putative HOP promoter sequence was digested out of the plasmid vector using 10 U/µl *Nhe*I (New England BioLabs, cat no.: B7003S) and *Bgl*II (Fermentas, cat no.: ER0082), and ligated into the promoterless pGL4 vector upstream of the firefly luciferase coding region (Promega, cat no.: E6721) to produce pHOP-x, the HOP promoter reporter construct. Digestion products were visualised using 1% (w/v) agarose gel electrophoresis (AGE) in TAE buffer (40 mM Trizma, 1 mM sodium EDTA) at 100 V for 90 minutes (unless stated otherwise). The digested HOP promoter fragment and linearized pGL4 were gel extracted using the ZymoClean™ Gel DNA Recovery Kit (Zymo Research, cat no.: D4002) and ligated together to form pHOP-x using T4 DNA ligase (Fermentas, EL0015) and 10X T4 DNA Ligase buffer (Fermentas, B69). The ligation reaction was incubated overnight (O/N) at 4^{o}C before the recombinant plasmids were transformed into JM109 *E.coli* competent cells (1 x 10^{8} cfu/µg DNA) (Hiang *et al.*, 2010).

The competent cells were prepared by using the RF1 and RF2 buffer protocol (Hanahan *et al.*, 1991). For transformation, JM109 *E.coli* competent cells were thawed on ice for

10 min, and incubated with 20 ng plasmid DNA on ice for a further 30 min. The cells were heat shocked for 45 seconds at 42°C and incubated on ice for 2 min. An additional 500 µl of 2xYT broth without antibiotics (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl) was added to the cells and incubated for 1 hr at 37°C with shaking. The transformation culture was plated onto 2xYT agar plates (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, 15 g/L agar) containing 100 µg/ml ampicillin and incubated at 37°C O/N. Plasmids were extracted from O/N liquid culture inoculations using the Zyppy Plasmid Miniprep kit according to the manufacturer's instructions (Zymo Research, cat no.: D4013). Restriction enzyme digestions were performed using 10 U *Sca*I (New England BioLabs, cat no.: R0122S), *Sty*I (New England BioLabs, cat no.: R0500S) or *Bgl*II (New England BioLabs, cat no.: R0144S) and incubated O/N at 37°C before being visualised using 1% (w/v) AGE at 100 V for 90 minutes. Endotoxin free pHOP-x constructs were obtained by performing an endotoxin free plasmid extraction using the EndoFree Plasmid Maxi Kit according to manufacturer's instructions (Qiagen, cat no.: 12362). The plasmid was confirmed by sequencing done by Inqaba Biotec (South Africa).

2.2.5 - Transient transfections and luciferase reporter assay

The HOP promoter constructs were transiently transfected into MEF-1, Hs578T, HeLa and MCF-7 cell lines using X-tremeGENE HP DNA Transfection Reagent (Roche, Switzerland) according to the manufacturer's instructions. Briefly, 2 x 10^5 cells were seeded into a 24 well plate (in media without PSA) and incubated at 37°C in a CO_2 incubator O/N. The cell culture medium was removed and 1 µg of the endotoxin free HOP promoter DNA reporter plasmid was diluted with 25 µl of Opti-MEM I Reduced® Serum Medium and 1 µl X-tremeGENE HP DNA Transfection Reagent and incubated for 20 min at room temperature before being added to the cells. The cells were co-transfected with 1 µg of the control reporter plasmid encoding EGFP (pLV-eGFP, Addgene plasmid: 36083). The pLV-eGFP plasmid provided a transfection efficiency control that was used to normalize the results obtained from independent transfections of the pHOP-x reporter construct. Cells transfected with empty pGL4 vectors or water were used as negative controls. The cells were incubated at 37°C in a CO_2 incubator for 48 hours before the promoter activity of the transfectants was determined by measuring the levels of luciferase activity. The cells were lysed with 10% (v/v) Triton-X/H_2O and

each well was split into 3 wells within a black clear bottomed 96 well plate (Corning, cat. no.: 3904). The eGFP within the lysate was excited at 488nm and the emitted fluorescence was recorded at 509nm (relative fluorescence units [RFU]) in a SynergyMx spectrophotometer (BioTek, USA). The sample was subsequently transferred into a white bottomed 96 well plate (Corning, cat. no.: 353296) containing FLAR Buffer (200 mM Tricine, 100 µM EDTA, 2.67 mM $MgSO_4$, 250 µM Adenosine Triphosphate [ATP], 250 µM Luciferin, 17 mM Dithiothreitol [DTT]). The firefly luminescence (relative light units [RLU]) was measured in the SynergyMx spectrophotometer at a sensitivity of 160. The ratio of RLU:RFU was calculated defined as the relative luciferase activity. The results were normalised against the negative controls. The P values for the assay were calculated using a one-way analysis of variance and the Tukey's Multiple Comparison Test. P values of less than 0.05 were considered as statistically significant.

2.2.6 - SDS-PAGE and Western Blot Analysis

Cells were harvested by scraping with a 200 µl pipette tip into sample lysis buffer (0.5 mM Tris, pH 6.8, 10% [v/v] glycerol, 10% [w/v] SDS, 5% [v/v] 2-mercaptoethanol) and protein concentrations determined using the NanoDrop 2000 Spectrophotometer (Thermo Scientific). Samples were prepared for electrophoresis by boiling for 5 minutes in Sodium-Dodecyl-Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) sample buffer (0.5M Tris pH 6.8, 10% [v/v] glycerol, 10% [w/v] SDS, 1% [w/v] Bromophenol blue, 5% [v/v] 2-mercaptoethanol). Equal amounts of proteins were separated by SDS-PAGE according to the protocol described by Laemmli (1970). Proteins were resolved using a 4% (v/v) stacking gel (0.5 M Tris-HCl, pH 6.8) and a 12% (v/v) resolving gel (1.5 M Tris-HCl, pH 8.8) at 100 V for 90 minutes in SDS-PAGE running buffer (0.25 mM Tris pH 8.3, 192 mM glycine, and 1% [w/v] SDS). PIERCE Prestained Protein MW Marker (ThermoScientific, cat no.: 00026612) was used for estimating molecular weights of proteins. Western analysis was performed on resolved proteins from SDS-PAGE according to the method of Towbin *et. al.* (1979). Transfer of proteins from the SDS-PAGE gel to the nitrocellulose membrane was carried out with the semi-dry blot in transfer buffer (0.13 mM Tris-HCl, 100 mM glycine and 20% [v/v] methanol) for 50 minutes at 0.4A. Protein transfer was confirmed with Ponceau staining (0.5% [w/v] Ponceau S in 1% [v/v] glacial acetic

acid). The membrane was blocked for one hour with either 5% (w/v) fat free milk powder (Santa Cruz Biotechnology, cat no.: sc-232S), in Tris-buffered saline (TBS; 50 mM Tris pH 7.5, 150 mM NaCl), or in 5% (w/v) Bovine Serum Albumin (BSA) (Promega, cat no.: R396D) diluted in TBS. The membrane was incubated with primary antibody (1:1000 for HOP, C/EBP$^{\beta}$ and ETS-1; 1:2500 for Histone and RhoA; 1:5000 for pERK1/2, pAKT, pFAK,; 1:100 for pJNK) diluted in either 5% (w/v) milk or 5% (w/v) BSA O/N at 4^{o}C. The membrane was washed with Tris-buffered saline containing Tween-20 (TBST; 0.001% [v/v] Tween-20 in TBS) twice for 20 minutes before incubation with species-matched peroxidase-conjugated secondary antibody (1:5000 for all unless otherwise stated) diluted in 5% (w/v) milk, 5% (w/v) BSA for an hour at room temperature. Membranes were washed in 4 changes of TBST for a total of one hour and detection of proteins carried out using a chemiluminescence developing kit (Clarity™ Western ECL Substrate, BioRad, USA; or Luminal/Enhancer Solution and Stable Peroxide Solution, Thermo Scientific, USA) on AGFA Medical X-ray film (AGFA HealthCare NV, Belgium) or using the ChemiDoc™-XRS and ImageLab Software (BioRad, USA). Densitometry was performed using ImageJ software (Schneider *et al.*, 2012).

2.2.7 - Quantitative Real Time PCR

Cells were fixed in 1 ml TRIzol® Reagent (a mono-phasic solution of phenol and guanidinium thiocyanate). The RNA was isolated from cells using the Trizol (Invitrogen) extraction method and no DNase treatments were performed. RNA was quantified immediately using the Nanodrop2000, and once quantified, the RNA samples were stored at -80°C until used. Samples were used within a week for qRT-PCR. The Nanodrop2000 quantification provided a means of measuring the purity, RNA samples with an A_{260}/A_{280} ratio between 1.8 and 2.0 only were used for qRT-PCR. A total of 100 ng of RNA was used in the KAPA SYBR® FAST One-Step qRT-PCR kit (KAPA Biosystems, cat no.: KK4650) in 20 µl reactions, with each reaction including 1x KAPA SYBR® FAST qPCR master mix (containing KAPA SYBR® Taq DNA polymerase and SYBR Green 1) (KAPA Biosystems, cat no.: KM4108), 200 nM forward primer, 200 nM reverse primer and 1x KAPA RT mix (wildtype M-MuLV Reverse Transcriptase and RNase Inhibitor). In addition to the experimental and control samples, no template controls (NTC) and no reverse transcriptase controls (NRT) were

also included in the qRT-PCR. All kit components and samples were kept on ice during experiment setup. The reaction was run in the MiniOpticon System MJ Minipersonal thermal cycler (BioRad, USA) for 5 min at 42°C for cDNA synthesis. The reverse transcriptase was inactivated by a consecutive 5 min at 95°C. PCR cycling and detection next performed by a further 40 cycles of 10 seconds at 95°C, 30 seconds at 50°C, with a melt curve performed for each from 55°C to 95°C with increments of 0,5°C for 10 seconds.

Data was analysed using the Bio-Rad CFX Manager 3.0 (BioRad, USA). The results were normalized against the quantification of the best of the 3 reference genes; glyceraldehyde-3-phosphate dehydrogenase (GAPDH), beta cytoskeletal actin (ACTB) or hypoxanthine phosphoribosyltransferase (HPRT1). The relative quantification method was used to determine the gene expression levels (Mankame *et al.*, 2010) by calculating the threshold cycle (Ct) values of the housekeeping genes for each sample so as to normalize the HOP Ct value in each sample. The delta Ct value of HOP was used to quantify the relative amount of HOP mRNA expressed under different conditions (Mankame *et al.*, 2010). Primers utilised were commercially available from Integrated DNA Technologies (IDT, USA) and are outlined in Table 1. If the NRT or NTC samples had Ct values that were less than 7 cycles greater than the experimental and control groups, the samples were considered contaminated with genomic DNA and thus excluded from the analysis. Each qRT-PCR experiment had three biological replicates, with three technical replicates for each biological replicate.

Table 1: qRT-PCR HOP and control gene primer details.

Gene Target	GenBank Accession	Amplicon size (bp)	PrimerBank ID	Splice Variant Detection	Forward Primer *					Reverse Primer *				
					Sequence	Tm (°C)	Length (bp)	Location	GC (%)	Sequence	Tm (°C)	Length (bp)	Location	GC (%)
HOP	NM_006819	77	110225356c1	All	CCTTACAGTGCTACTCCGAAGC	62.4	22	68-89	55	ATAGGCAGCAGAACGGTTGC	62.9	20	144-125	55
GAPDH	NM_001256799	197	378404907c1	All	GGAGCGAGATCCCTCCAAAAT	61.6	21	108-128	52	GGCTGTTGTCATACTTCTCATGG	60.9	23	304-282	48
ACTB	NM_001101	250	4501885a1	All	CATGTACGTTGCTATCCAGGC	60.8	21	393-413	52	CTCCTTAATGTCACGCACGAT	60.2	21	642-622	48
HPRT1	NM_000194	131	164518913c1	All	CCTGGCGTCGTGATTAGTGAT	61.9	21	16-36	52	AGACGTTCAGTCCTGTCCATAA	60.5	22	146-125	45

*RT-PCR primers obtained from IDT.

2.2.8 - Effect of mutated RAS on the expression of HOP

Three plasmids were obtained from Addgene and used to determine the effect of mutated RAS on the activity of the HOP promoter. One plasmid encoded wildtype HRAS (mEGFP-HRas, plasmid: 18662), another encoded a mutated HRAS G12V (mEGFP-HRas G12V, plasmid: 18666) that was constitutively active, and another mutated HRAS S17N (mEGFP-HRas S17N, plasmid:18665) that was dominant negative (Yasuda *et al.*, 2006). MCF-7, HeLa and Hs578T breast carcinoma cells were seeded into a 24 well plate at a density of 2 x 10^5 cells/500 µl media. The cells were transiently transfected, as described above, with one of the three RAS plasmids, along with pHOP-x. A total of 500 ng of the plasmids was transfected per well. Luciferase assays were performed as described before to determine the effect of RAS on the promoter activity of HOP in Hs578T and HeLa cells. The quantitative iScript One-Step RT-PCR Kit was utilised to determine the effect of RAS on the HOP mRNA levels. RNA was extracted from MCF-7 cells transfected with the HRAS plasmids using the Trizol method and 100 ng of RNA was used for qRT-PCR. The experimental groups were RNA samples from the cells that had been transfected with a RAS plasmid, while the control group was the RNA samples from untransfected cells. Western Blot analyses of transfected Hs578T and HeLa cell lysates were utilised to determine the effect of RAS on the protein expression of HOP using the procedure described previously.

2.2.9 - Cell Cytotoxicity assays

Cells were seeded at a density of 1 x10^4 cells/well in 50 µl culture media in a 96 well plate and allowed to adhere O/N at 37^{o}C in 9% CO_2. The next day the cells were treated with an additional 50 µl media containing various concentrations of alpha hydroxyl farnesyl phosphoric acid (αHFPA) (Santa Cruz Biotechnology, cat no.: sc-205200), sorafenib (BAY 43-9006) (Santa Cruz Biotechnology, cat no.: sc-220125), U0126 (Cell Signalling Technology, cat no.: #9903) or 3-(2-Aminoethyl)-5-((4-et hoxyphenyl)methylene)-2,4-thiazolidinedione (Sigma, cat no.: A6355) as indicated in figure legends. The cells were incubated at 37^{o}C in 9% CO_2 for 72 hrs, after which 10 µl of WST-1 Cell Proliferation Reagent (Roche, cat no.: 11644807001) (diluted in 40 µl media) was added to each well, and incubated at 37^{o}C in 9% CO_2 for 4 hrs. The absorbance was read at 450nm and dose response cell cytotoxicity curves were

constructed as the log concentration of the drugs versus normalised absorbance. The IC_{50} values were calculated as the drug concentration that caused a 50% reduction in cell viability in comparison to the control cells.

2.2.10 - Effect of MAPK pathway inhibitors on the expression of HOP

Hs578T cells were seeded at 5 x 10^5 cells/well into a 6 well plate and treated with 3.568 µM Sorafenib, 37.7 µM U0126, 30 µM αHFPA or 60 µM ERK2i and incubated for 48 hrs at 37^{o}C in 9% CO_2. RNA was extracted using the Trizol method and 100 ng of RNA was used for perform qRT-PCR to determine the effect on the mRNA expression of HOP. The experimental groups were RNA samples from the cells that had been treated with inhibitors, while the control groups were the RNA samples from cells that had been treated with the vehicle controls (DMSO, EtOH). Alternatively, cells were harvested for the proteins by scraping into sample lysis buffer and 100 µg of protein was used for Western Blot analyses as previously described to determine the effect of RAS on the protein expression of HOP.

2.2.11 - Time Course Study of the effect of Sorafenib on HOP protein levels

Hs578T cells were seeded into a 6 well plate at a density of 2.5 x 10^5 cells/well and left to adhere O/N at 37^{o}C with 9% CO_2. The cells were treated with 3.568 µM Sorafenib and lysed with cell lysis buffer at varying time points (12hr, 24hr, 36hr, 48hr). Cell lysates were sonicated at 40 Hz for 5 seconds in 10 second intervals for a total of 2 minutes. The lysates were used in SDS-PAGE and Western Blot analysis, consequently making use of ImageJ (Schneider *et al.*, 2012) to perform densitometry to determine the relative levels of HOP at each time point.

2.2.12 - Immunofluorescence staining and Fluorescent Microscopy

Hs578T cells were seeded at a density of 2 x 10^5 cells/ml into a 24 well plate containing a glass coverslip and incubated overnight at 37^{o}C with 9% CO_2. Cells were fixed by a 5 second treatment with ice-cold methanol (-20^{o}C) and allowed to air dry. Cells were blocked with 1% (w/v) BSA in PBS [8 g/l NaCl, 200 mg/ml KCl, 1.15 g/l Na_2HPO_4. 200 mg/ml KH_2PO_4] (BSA/PBS) for 45 minutes at room temperature and incubated with primary antibodies in 1% (w/v) BSA/PBS at 4^{o}C O/N. After incubation, cells were washed twice in 0.1% (w/v) BSA/PBS for 5 minutes followed by incubation with

appropriate species-specific fluorophore-conjugated secondary antibodies at room temperature for 2 hours in the dark in 0.1% (w/v) BSA/PBS. Cells were washed twice with distilled water and the nucleus stained with Hoechst-33342 (DAPI) (1 µg/ml in distilled water) before mounting with DAKO fluorescent mounting medium. Immunofluorescence staining was visualised using the Zeiss AxioVert AI FL-LED inverted fluorescence microscope or the Zeiss LSM 780 Confocal Microscope (Carl Zeiss Microscopy, USA) and analysed using ZEN lite 2012 software (Carl Zeiss Microscopy, USA). Imaging was performed using the 40x, 60x and 100x objectives. Details of individual treatments are described in the figure legends. The fluorescence intensity of a protein within the nucleus and the cytoplasm was determined by plotting the grey value profile of the fluorescence channel of interest for each cell using ImageJ (Fig.3). The borders of the nuclei were determined by overlaying the channel that detected the fluorescence of the Hoechst with that of the channel that detected the fluorescence of the protein of interest (Fig.3A). The grey values were totalled within the nucleus and cytoplasm (Fig.3B). The totalled grey value was divided by the length of the either the nucleus or of the cell (excluding the nucleus, thus the cytoplasm was considered the signal which was on either side of the nuclear signal) (Fig.3B). A minimum of 10 cells were utilised for the analysis.

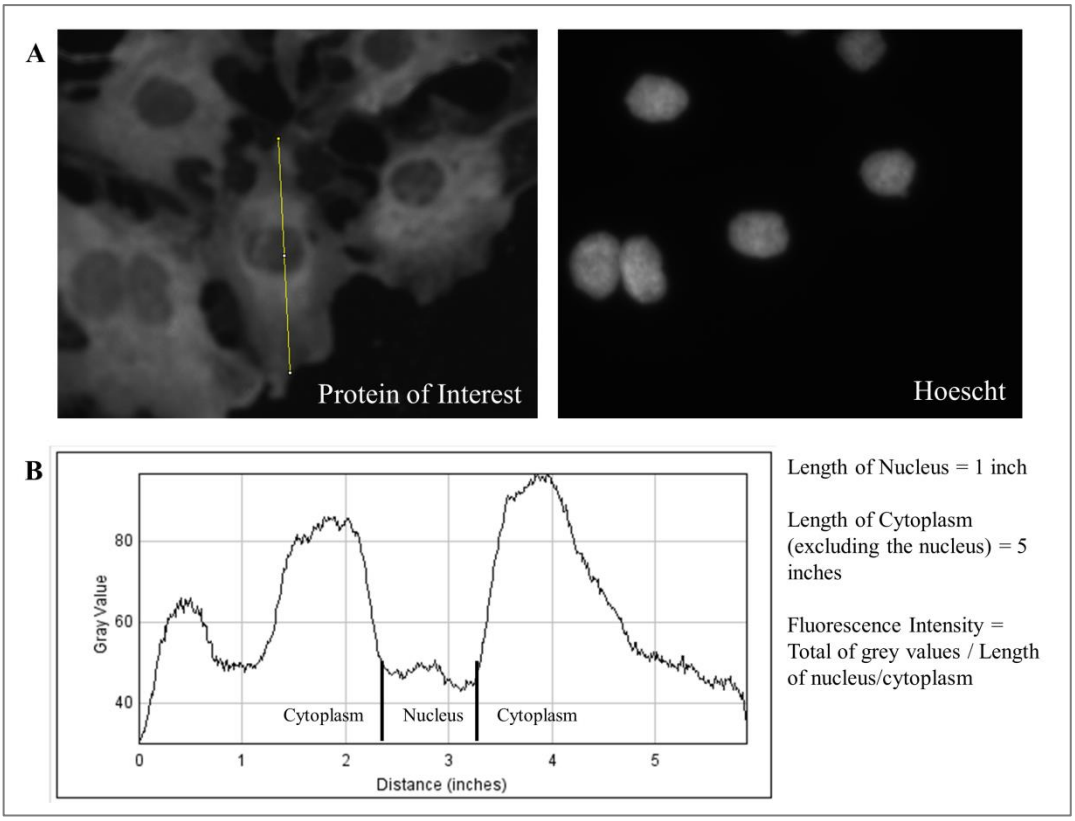


Figure 3: Measurement of fluorescent intensity within a cell.

The location of the nucleus within the cell was determined for one z-stack (A). The profile of the grey values (indicating fluorescence intensity of the protein of interest) within a cell was plotted (B). The grey values were added and totalled for the cytoplasm and nucleus respectively. The totals were divided by the length of either the nucleus or the cell (excluding the nucleus) to give the fluorescence intensity within the nucleus or the cytoplasm.

Chapter 3

Comparison of the HOP gene across species and a bioinformatic analysis of the putative HOP promoter

3.1 Introduction

Model organisms are of vital importance in furthering our knowledge in the study of human diseases and the mechanisms by which genes function. To provide an in depth study of the molecular functioning of a protein of interest, one needs to make use of a model organism that most closely represents the cellular and molecular make-up of the human cell (Armengaud *et al.*, 2014). This refers to the representational scope of the model organism, which describes the extent to which the data obtained from an experimental organism can be projected onto other organisms, or in other words the organisms for which the experimental organism is acting as a proxy (Ankeny & Leonelli, 2011). The representational target in turn refers to the concept, process or entity that is to be explored through the use of the experimental organism (Ankeny & Leonelli, 2011). Thus the preferred range of the representational scope is determined by the representational target. Through a statistical analysis, Nabhan & Sarkar (2014) were able to present a means of predicting the most suitable model organisms for the study of various human cancers and infectious diseases. They performed a structural and functional analysis of disease-related genetic pathways and matched them to molecular interaction networks of 14 model organisms. From there they could predict which model organism was most suitable as a specific disease model. For all 12 cancer subtypes analysed in the study, *Mus musculus* and *Rattus norvegicus* were by far the best disease models. However, in the case of melanoma and renal cell carcinoma it was observed that *Danio rerio* was a more suitable disease model than murine and rat (Nabhan & Sarkar, 2014). Making use of the most suitable model organism or disease model would thus be the most cost effective and time sensitive means of furthering the knowledge of the functioning of HOP in human diseases.

HOP has been either predicted or identified in many model organisms such as the nematode (Song *et al*. 2009), fruit fly (Grigus *et al*. 1998), zebrafish (Woods *et al.* 2005; Tastan Bishop *et al*. 2013), mouse (Blatch *et al*. 1997), rat (Demand *et al*. 1998), frog (Klein *et al*. 2002), fish (Andreassen *et al.* 2009), parasites (Webb *et al*. 1997; Hombach *et al*. 2013) and plants (Zhang *et al*. 2003; Chen *et al*. 2010), as well as in the coelacanth (Amemiya *et al*. 2013; Tastan Bishop *et al.* 2013) and human genomes (Honore *et al.* 1992). This chapter presents a study comparing the HOP gene across species in an attempt to determine what organism is best suited to being a model

organism for the study of HOP in humans. More specifically, since the murine model is the animal model most commonly utilised in studying human disease, we address the question of whether the murine model is a sufficient model for the study of human HOP by determining how similar the HOP genes within the murine and human systems are. Additionally we analyse the factors influencing the transcriptional regulation of HOP, specifically within the human system.

3.2. Results

3.2.1 - Phylogenetic analyses of the HOP orthologues

This study identified that the HOP gene is observed once per genome across all species tested. Table 2 summarises some of the pertinent information on the HOP gene that will be discussed in more detail throughout this chapter. The length of the HOP gene is similar amongst the orthologues, although the HOP gene in yeast, fruit fly and nematode is one tenth the size of the HOP gene found in the other orthologues. Additionally, all orthologues contain 14 exons in the HOP gene, besides the Coelacanth which has 13. The HOP gene in yeast, fruit fly and nematode has only one exon. The HOP gene is found on the positive strand for all orthologues besides mouse and yeast. The HOP mRNA transcript length is similar across all orthologues although it is shorter in yeast and the frog, and it is less than half the size in nematodes. The translated protein sequence of HOP is similar across all orthologues but again is almost half the length in nematodes. The 5` and 3`UTR regions of the HOP gene contain no pattern across the orthologues.

An initial study of the evolution of the HOP speciation was performed across the identified orthologues obtained from HomoloGene, NCBI. Phylogenetic analyses of the protein sequences of the HOP orthologues were performed using MEGA5.2 (Fig.4).

Table 2: Summary of the comparison of the genomic structure and location of HOP orthologues.

Species	Chromosome Location of STIP1	Start Position	Stop Position	Gene length (bp)	Strand	Coding exons	Transcript length (bp)	Protein length (aa)	5` UTR (bp)	3` UTR (bp)
Homo sapiens	11q13	63953587	63972020	18433	+	14	2196	543	147	417
Mus musculus	19A	7020696	7040026	19330	-	14	2173	543	138	403
Danio rerio	XV	69801144	69817741	16597	+	14	2462	542	37	796
Saccharomyces cerevisiae	5	381053	382822	1769	-	1	1770	589	0	0
Latimeria chalumnae	Unassembled genome	1752251	1772059	19808	+	13	2137	576	250	156
Xenopus tropicalis	Unassembled genome	211042	223291	12249	+	14	1914	543	35	247
Drosophila melanogaster	2L	295122	297449	2327	+	1	2023	590	151	400
Caenorhabditis elegans	V	773592	774716	1124	+	1	963	320	0	0

*Data from ENSEMBL and NCBI databases.

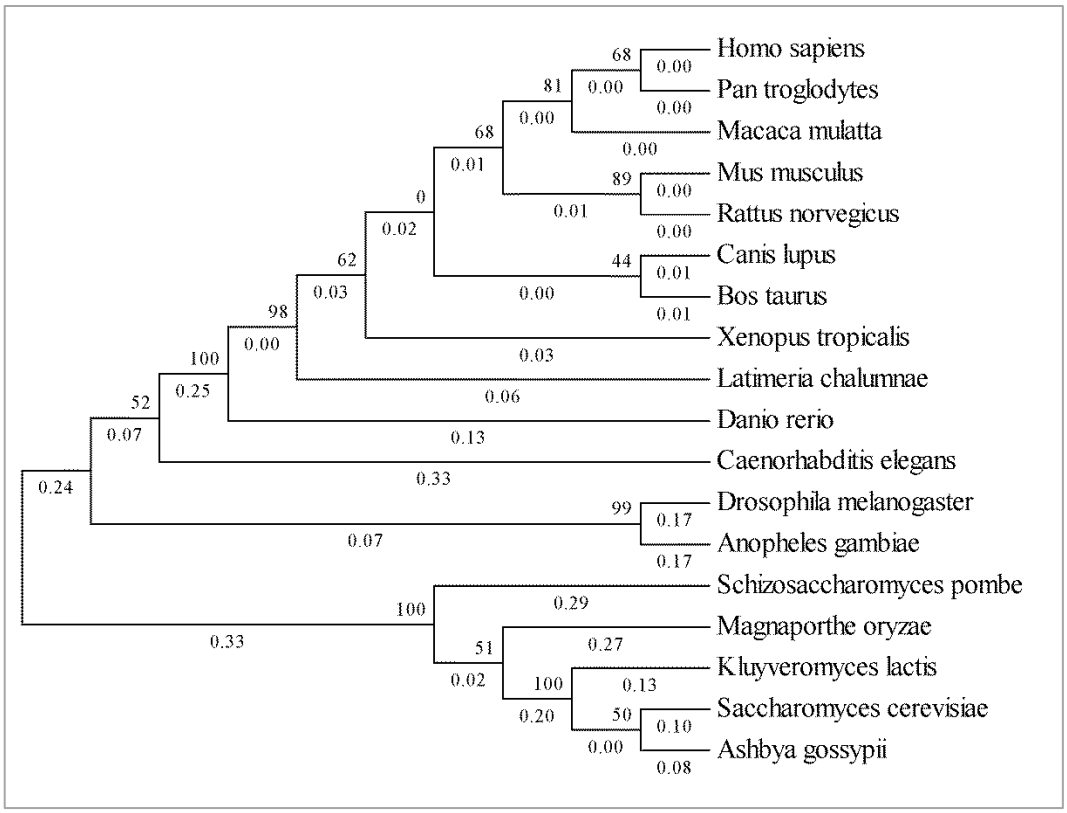


Figure 4: Molecular Phylogenetic analysis of the orthologous HOP protein sequences by the Maximum Likelihood method.

The evolutionary history was inferred by using the Maximum Likelihood method based on the Whelan And Goldman model (Whelan & Goldman, 2001). The optimal tree with the sum of branch length = 2.49973822 is shown. The tree with the highest log likelihood (-3504.7577) is shown. The percentage of trees in which the associated taxa clustered together is shown above the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 3.7753)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 17.1146% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (branch lengths are shown below the branches). The analysis involved 18 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 261 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura *et al.*, 2011).

The maximum likelihood tree was an unrooted tree produced from the protein coding sequences and indicated the suspected evolutionary conservation across the orthologues (Fig.4). Within the protein tree, the evolutionary pathway of the HOP protein could be traced from the unicellular eukaryotes to plants, to fish and amphibians and then to land mammal. All branch lengths between nodes were observed to be less than 0.33, with most mammals having branch lengths of zero (Fig.4). Low branch lengths indicated low numbers of substitutions per site in the amino acid sequences but did not correlate with high bootstrap values (Hall, 2013). Within the land mammals, the HOP protein evolves from cattle, to the murine, to the lower primates and finally to the human. The yeast and fungal species were separated from animals by the first branching of the tree. Certain branches had bootstrap values above the 95% threshold, those being the branching between *Danio rerio* and the transition to amphibians, the coelacanth and the transition to the amphibians, between mouse and rat, and between lower primates and greater primates (including human) (Fig.4). The branching between the various yeasts and fungi were either very high or very low. Those with particularly low bootstrap values (less than 70%) were the branching between *Homo sapiens* and *Pan troglodytes*, between murine and primates, between cattle and the greater mammals, between *Xenopus tropicalis* and mammals, and lastly branching between nematode and amphibians (Fig.4).

3.2.2 - Exon and Protein Organisation of HOP between species

The organisation of the exons and introns within the HOP gene was compared across the orthologues (Fig.5A). The HOP gene transcript in yeast, the fruit fly and the

nematode was observed to be comprised of a single exon, as is commonly observed in these genomes, while the remaining orthologues were comprised of 14 exons, or 13 in the case of the coelacanth orthologue (Table 2, Fig. 5A). The gene length along the chromosome ranged from 1 kb in nematodes to 19 kb in the coelacanth (Table 2). Whereas the mRNA transcript lengths had a smaller size range across the orthologues, with the length ranging from 0.9 kb in nematodes, to 2.9 kb in the zebrafish (Table 2, Fig.5A). All orthologues had 5`UTR and 3`UTR sequences of varying lengths, except yeast and nematode which did not have either according to the ENSEMBL and NCBI databases (Table 2, Fig. 5A). The distribution of the exons had no regularity or pattern across the orthologues.

Next, the domain structure of the HOP proteins was compared across the orthologues. Three tetratricopeptide repeat domains (TPR) were conserved across all species except the nematode which had only two (Fig. 5B). All species were found to have the two conserved aspartic acid and proline rich domains (DP), otherwise known as the heat shock chaperone binding motif (STI1 domain sequences). However, again the nematode had only one conserved domain (Fig. 5B). The human and the mouse each had three TPR domains (TPR1, TPR2A and TPR2B). The individual TPR domains were specifically linked to the exons encoding them (Fig. 5B). TPR2A was found to be the most degenerate domain of the three, with the lowest sequence identity and similarity in the nucleotide and protein composition across the different species. Within human and murine HOP amino acid sequences, which have a high degree of sequence identity in TPR1 (98%) and TPR2B (91%), the sequence identity for the TPR2A domain was 4.7%. The nucleotide sequences encoding TPR1 within the human coding sequence spanned exons 1 and 2; TPR2A spanned exons 5, 6 and 7; while TPR2B spanned exons 8 and 9. TPR1 of the mouse coding sequence spans exons 1 and 2; TPR2A spans exons 5, 6 and 7; while TPR2B spans exons 8, 9, 10 and 11.

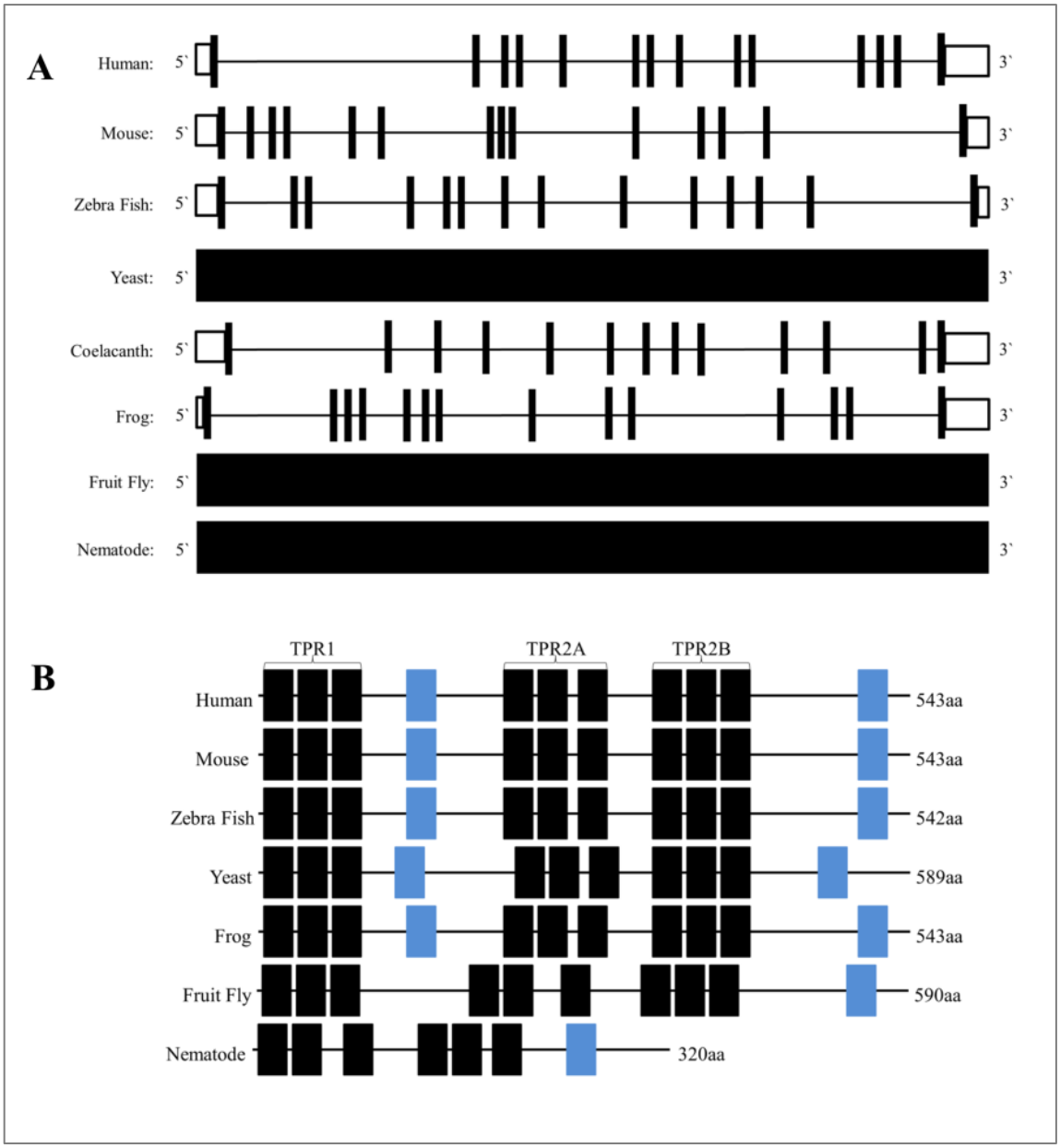


Figure 5: Comparison of the exon organisation and protein structure of HOP orthologues.

A) Schematic diagram depicting the exon structure of the HOP gene across orthologues within *Homo sapiens* (human), *Mus musculus* (mouse), *Danio rerio* (zebrafish), *Saccharamyces cerevisiae* (yeast), *Latimeria chalumnae* (coelacanth), *Xenopus tropicalus* (frog), *Drosophila melanogaster* (fruit fly) and *Caenorhabditis elegans* (nematode) species. The black boxes illustrate the exon positions along the mRNA transcript. The spaces between the black boxes represent the intron sequences. The white boxes illustrate the 5` and 3` untranslated regions (UTR) at either end of the mRNA transcript. B) Schematic diagram comparing the protein structures of HOP. Black boxes represent tetratricopeptide repeat motifs (TPR motifs), three TPR motifs comprising a TPR domain. The blue boxes represent DP (aspartic acid and proline rich domains) or STI1 domain sequences (heat shock chaperone binding motif). The amino acid length of each protein is given at the end of the protein.

3.2.3 - Analysis of the alternate transcripts of the human HOP gene.

Alternate splicing is considered the process of identical mRNA transcripts being spliced in different ways; usually this is a positive tool within organisms as it creates protein diversity from a single gene sequence. However splicing has also been identified with the cause of many diseases (Venables, 2004; Srebrow & Kornblihtt, 2006). This, hence, led to an initial identification of alternate transcripts of HOP within humans (Table 3).

Table 3: Comparison of the alternate protein coding transcripts of human HOP.

Alternate Transcript	Transcript length (bp)	Protein length (aa)	Exons	Sequence Identity to STIP1-001 (%) *		Protein Domains		Evidence **
				Nucleotide	Protein	DP/STI	TPR motif	
STIP1-001	2196	543	14	100	100	2	9	Evidence at the protein and transcript level
STIP1-010	2743	590	14	99.63	99.45	2	9	Evidence at the transcript level
STIP1-003	1900	519	14	100	97.69	2	8	Evidence at the transcript level
STIP1-008	877	137	4	90.34	87.59	1	3	Evidence at the transcript level
STIP1-012	772	236	5	95.5	95.76	1	3	Evidence at the transcript level

*Sequence identity was determined by a pairwise alignment to STIP-001 (allowing the ends to slide). Thus the size difference between transcripts was ignored and the regions contained within the transcripts were focussed on.

** Data obtained from the ENSEMBL database. STIP-001 is the transcript for the known HOP protein. Evidence at the transcript levels indicates evidence of the expression of the mRNA transcript, although not strictly proven at the protein level.

Five putative protein coding, alternate transcripts were identified for human HOP (Table 3). The first isoform (STIP1-001) is the known and functional co-chaperone HOP, while the remaining isoforms had evidence of their existence at the mRNA transcript level. Although the splice variants are truncated to varying degrees, the sections that remain have a high similarity to the known HOP (STIP-001). STIP1-010 and STIP1-003 have similar gene and protein structures to that of STIP1-001. The only difference in STIP1-010 protein was that it had an additional 140 bp in the 5` region of the HOP coding sequence. STIP1-003 was missing 70 bp within the HOP coding sequence, resulting in it missing the third TPR motif within the TPR1 domain. STIP1-008 is a truncated transcript of HOP at the 5` end, such that the nucleotides only encode two of the TPR motifs within the TPR2B domain. STIP1-012 is also a truncated HOP

transcript however is truncated at the 3` end, such that it only encodes the first TPR1 domain.

3.2.4 - Comparison of the genomic context of HOP between orthologues.

Having determined that HOP was conserved amongst different species, the genomic context of the HOP orthologues was compared in relation to the surrounding genes. Synteny was identified within the loci containing the HOP gene for each species. Synteny is the conservation of a series of neighbouring genes on a chromosome across related species, each being derived from a common ancestor (Nadeau, 1989).

Local synteny was observed between the genes neighbouring HOP across the human, mouse, frog and the coelacanth orthologues (Fig.6). The degree of synteny decreased as the distance between the species identified by the phylogenetic tree increased (Fig. 4). The syntenic region within the mouse was found to have the identical order of genes to the human region, however in an inverted orientation and with a differing location of the HOP and fermitin family member 3 (FERMT3) genes (Fig. 6). The human HOP gene was localised between and partially overlapped the nucleoside diphosphate-linked moiety X motif 22 (NUDT22) and an HSP40 homologue (subfamily C, member 4) (DNAJC4), while the murine HOP gene was localised between tRNA phosphotransferase (TRPT1) and fibronectin leucine rich transmembrane protein (FLRT1) (Fig.6). The frog orthologue was also found to retain the same neighbouring genes as human and mouse, however in a shuffled order. Furthermore, in both human and murine chromosomes FKBP2, VEGFB, DNAJC4, NUDT22, FERMT3 and TRPT1 were clustered with smaller distances between them in comparison to the remaining neighbouring genes (Fig.6). The coelacanth region of interest had a mere 3 homologous genes (Fig.6). For both the frog and coelacanth regions, the HOP genes were localised adjacent to FERMT3. The zebrafish orthologue had merely one corresponding gene. No synteny was observed for the yeast, nematode and fruit fly orthologues (Fig.6).

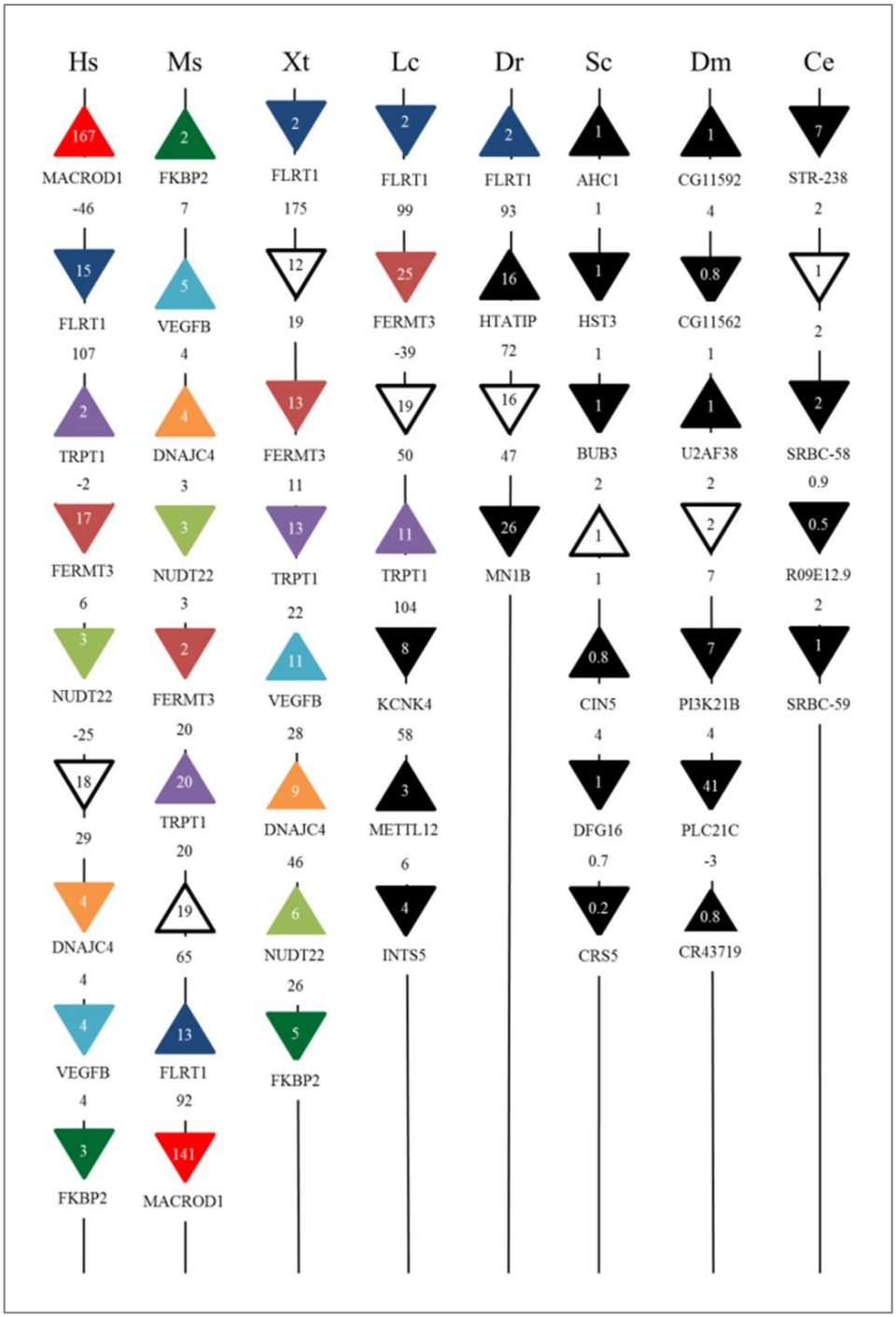


Figure 6: Conserved synteny of the HOP gene.

Schematic diagram of the genomic representations of HOP across *Homo sapiens* (human), *Mus musculus* (mouse), *Danio rerio* (zebrafish), *Saccharamyces cerevisiae* (yeast), *Xenopus tropicalus* (frog), *Drosophila melanogaster* (fruit fly) and *Caenorhabditis elegans* (nematode) species. Triangles represent genes, the point of the triangle points to gene orientation on the chromosome. The white triangle in all cases represents the HOP gene. Coloured triangles represent genes with synteny across at least two species. The names of the genes are listed below the corresponding triangles. Numbers within triangles represent the size of the gene in kb. Numbers between the triangles represent the distance between the genes in kb. Negative numbers represent overlapping genes and the distance across which they overlap.

Table 4: The differential expression of HOP mRNA in cancers.

Up-regulation of HOP in cancers							
Case		**Syntenic genes with same exp. pattern**	**Reference**	**Case**		**Syntenic genes with same exp. pattern**	**Reference**
Compared to normal tissue	Microsatellite Instability (MSI colorectal cancer)	1	(Alhopuro *et al.*, 2012)	**Compared to wildtype**	Coronary Endothelial Cell (induced hypoxia)	4	(Chi *et al.*, 2006)
	Lung adenocarcinoma (late)	2	(Lo *et al.*, 2012)		Smooth Muscle Cell (induced hypoxia)	3	(Chi *et al.*, 2006)
	Cytomegalovirus Infected alpha-beta T cells	1	(Baitsch *et al.*, 2011)	**Compared to surroundin g tissue**	Cholangiocarcinoma Surrounding Liver	3	(Andersen *et al.*, 2012)
	Cytomegalovirus Infected alpha-beta T cells	1	(Baitsch *et al.*, 2011)		Cholangiocarcinoma Tumour	2	(Andersen *et al.*, 2012)
	Gastric Adenocarcinoma	1	(Cho *et al.*, 2011)		Interstitial Lung Disease	0	(Lindahl *et al.*, 2013)
	Gastrointestinal Stromal Tumour	1	(Cho *et al.*, 2011)	**Other**	Undifferentiated stem cells vs. Differentiated stem cells	5	(Kim *et al.*, 2012)
	Recurrent Bladder Tumour	4	(Kim *et al.*, 2010)		TaY-E10 vs MT-2 vs MT-4 (mock vs 17-AAG, 100nM)	0	(Kurashina *et al.*, 2009)
	Bladder Mucosae Surrounding Cancer	4	(Kim *et al.*, 2010)		TaY-E10 vs MT-2 vs MT-4 (mock vs 17-AAG, 300nM)	0	(Kurashina *et al.*, 2009)
	Urinary Bladder Cancer	4	(Kim *et al.*, 2010)		TaY-E10 vs MT-2 vs MT-4 (mock vs 17-AAG, 500nM)	0	(Kurashina *et al.*, 2009)
	Universal Human Reference RNA	0	(Skotheim *et al.*, 2005)		TaY-E10 vs MT-2 vs MT-4 (mock vs 17-DMAG, 150nM)	1	(Kurashina *et al.*, 2009)
	Embryonal Carcinoma	0	(Skotheim *et al.*, 2005)	**Down-regulation of HOP in cancers**			
	Stomach Cancer	0	(Leung *et al.*, 2002)	**Case**		**Syntenic genes with same exp. pattern**	Reference
	Liver Failure: Acute	2	(Nissim *et al.*, 2012)	**Compared to normal tissue**	Cytomegalovirus Infected alpha-beta T cells	3	(Baitsch *et al.*, 2011)
	BRCA-1	0	(Richardson *et al.*, 2006)		Cytomegalovirus Infected alpha-beta T cells	4	(Baitsch *et al.*, 2011)
	Glioblastoma	8	(Sun *et al.*, 2006)		Leukaemia	2	(Paul *et al.*, 2011)
	Astrocytoma	8	(Sun *et al.*, 2006)		Testicular Seminoma	0	(Skotheim *et al.*, 2005)
	Oligodendraglioma	8	(Sun *et al.*, 2006)	**Aggressive vs. Non-aggressive**	HT-1080 (fibrosarcoma) vs. colorectal adenocarcinoma	1	(Haga *et al.*, 2010)
	Oesophageal Carcinoma	1	(Shimokuni *et al.*, 2006)		Ovarian Serous Tumour Malignant	6	(Anglesio *et al.*, 2008)
	Interstitial Pneumonia	0	(Lindahl *et al.*, 2013)		Normal Testicular Germ Cell (vs. neoplastic tissue)	0	(Skotheim *et al.*, 2005)
	Inactive Ulcerative Colitis	3	(Planell *et al.*, 2013)	**Compared to WildType**	Breast Epithelial Cell (induced hypoxia)	5	(Chi *et al.*, 2006)
	Ulcerative Colitis Involved Active	2	(Planell *et al.*, 2013)		MT-4 (Interleukin 2-independent ATL cell line)	1	(Kurashina *et al.*, 2009)
	Ulcerative Colitis Non-involved Active	5	(Planell *et al.*, 2013)		MT-2 (Interleukin 2-independent ATL cell line)	0	(Kurashina *et al.*, 2009)
	Non-small Cell Lung	2	(Sanchez-Palencia *et al.*, 2011)		Colon Cancer (TCF4 knock-out)	2	(Wang *et al.*, 2013)
	Cholanaiocarcinoma surrounding Liver	3	(Andersen *et al.*, 2012)	**Other**	TaY-E10 vs MT-2 vs MT-4 (mock vs 17-AAG, 700nM)	0	(Kurashina *et al.*, 2009)
	Cholanaiocarcinoma Tumour	3	(Andersen *et al.*, 2012)				
	Oesophagus Squamous Cell Carcinoma	1	(Tong *et al.*, 2012)				

In an attempt to further elucidate any relationships or connections between the HOP syntenic genes, an *in silico* analysis of HOP expression in human cancers was performed (Table 4). Deregulated HOP expression was identified using the GEMMA gene expression database and tool (Zoubarev *et al.*, 2012), any correlating deregulation of the expression of the syntenic genes was highlighted. More specifically, the number of syntenic genes that had the same expression pattern as HOP was recorded. The data depicted in Table 4 shows 38 studies with an up-regulation of HOP in cancers, with only 12 studies showing down-regulation of HOP. Within the majority of the studies observed at least one of the HOP syntenic genes had a similar expression profile (Table 4).

3.2.5 - The co-expression profile of HOP in cancer

The high synteny previously observed between the mammals (Fig.6) and the similar expression patterns observed between the syntenic genes in human cancers (Table 5) led to the investigation of the possibility of the co-expression of the syntenic genes. Using the Gemma gene expression database and tool (Zoubarev *et al.*, 2012), all human cancer datasets were searched for co-expression evidence between the syntenic genes. It has been previously observed that genes sharing similar functions were often found in clusters and were co-expressed in mammalian genomes (Michalak, 2008).

HOP was observed to be co-expressed with both MACROD1 and TRPT1 only, however with minimal supporting evidence (Table 5, Fig. 7). Several of the remaining genes from the syntenic locus were found to have *in vitro* evidence in a variety of datasets to suggest co-expression between them (Table 5, Fig. 7). NUDT22 was co-expressed with TRPT1, DNAJC4 and FKBP2; also having the greatest number of supporting datasets. DNAJC4 was co-expressed with VEGFB, MACROD1 and FKBP2. The only genes that were not co-expressed with any of the others were FERMT3 and FLRT1.

Table 5: Co-expression of HOP syntenic genes in human cancers.

Query Gene	Query Gene NCBI Id	Coexpressed Gene	Coexpressed Gene NCBI Id	Specificity	Positive Support *	Negative Support	Datasets tested
NUDT22	84304	TRPT1	83707	0,6005	22	0	178
NUDT22	84304	DNAJC4	3338	0,8861	9	0	146
NUDT22	84304	FKBP2	2286	0,6767	9	0	181
TRPT1	83707	DNAJC4	3338	0,8861	7	0	157
VEGFB	7423	DNAJC4	3338	0,8861	6	0	228
TRPT1	83707	FKBP2	2286	0,6767	6	0	186
TRPT1	83707	VEGFB	7423	0,8109	6	0	189
DNAJC4	3338	FKBP2	2286	0,8861	4	0	227
NUDT22	84304	VEGFB	7423	0,8109	4	0	178
MACROD1	28992	DNAJC4	3338	0,8861	3	0	221
VEGFB	7423	FKBP2	2286	0,8109	2	0	265
STIP1 **	10963	MACROD1	28992	0,9059	2	0	256
TRPT1	83707	STIP1 **	10963	0,9059	2	0	186

*Stringency set to a minimum requirement of 2 datasets with supporting evidence. Specificity refers to the measure of specificity of the co-expression of the gene with others in its taxon. Data obtained from Gemma gene expression database and tool.

**HOP is referred to as STIP1 in the study.

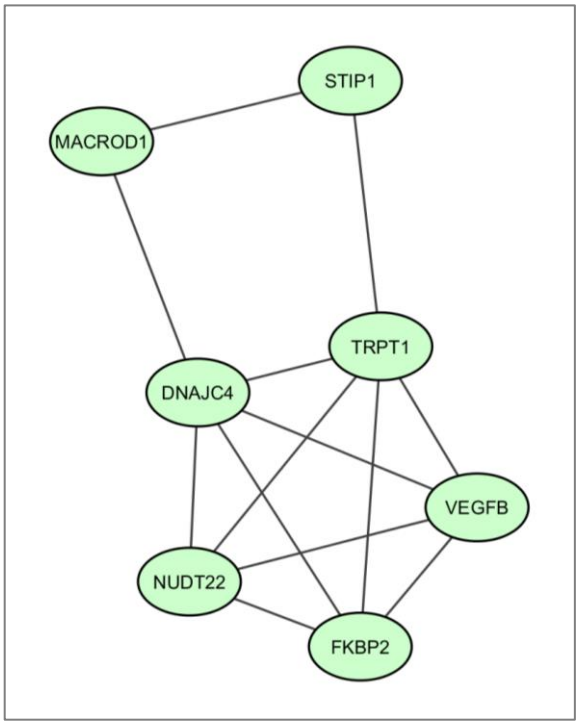


Figure 7: A co-expression network of HOP syntenic genes.

The co-expression network between the HOP syntenic genes was compiled from data gathered from the Gemma (Zoubarev *et al.*, 2012) and C0XPRESdb (version 6) (Obayashi *et al.*, 2008) databases and visualised using Cytoscape 3.0 (Cline *et al.*, 2007). The nodes represent proteins that are connected by edges. These edges represent *in vitro* evidence for co-expression between the proteins. HOP is shown as STIP1.

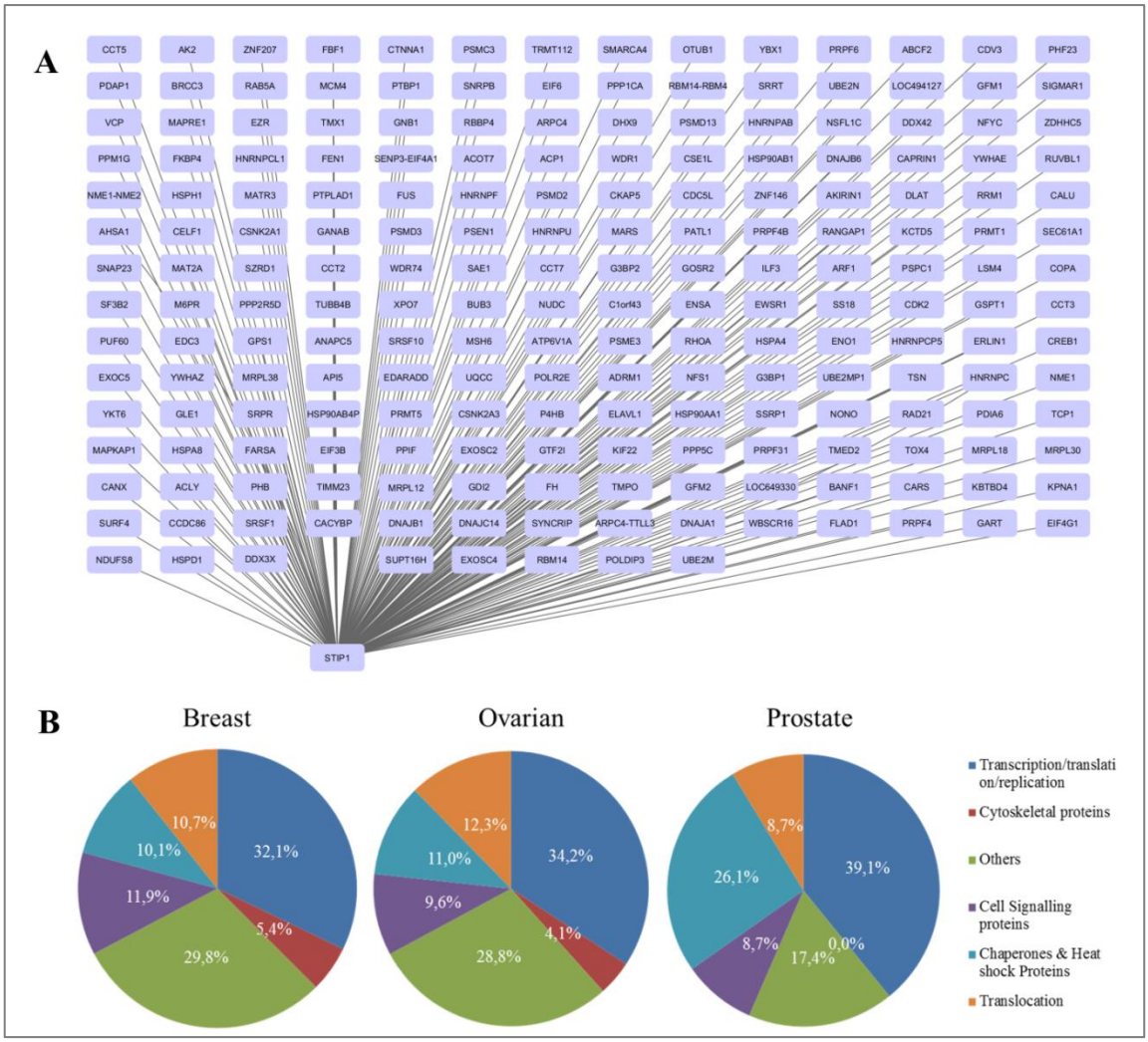


Figure 8: Identification and Categorisation of all genes co-expressed with HOP in cancers.

A) The co-expression network of HOP was compiled from data gathered from the Gemma (Zoubarev *et al.*, 2012) and C0XPRESdb (version 6) (Obayashi *et al.*, 2008) databases and visualised using Cytoscape 3.0 (Cline *et al.*, 2007). The nodes represent proteins that are connected by edges. These edges represent *in vitro* evidence for co-expression between HOP and the proteins. HOP is shown as STIP1. B) Using the Gemma database, the genes found to have evidence of co-expression with HOP particularly in human breast, ovarian and prostate cancers were categorised according to functionality. Genes were categorised as DNA replication, transcription or translation related; cell signalling proteins; cytoskeletal proteins; chaperone or heat shock proteins; translocation proteins; or other proteins. Stringency was set to a minimum requirement of 3 datasets with supporting evidence.

The analysis was then expanded to search for any genes (regardless of synteny) that had supporting *in vitro* evidence for their co-expression with HOP (Fig. 8A). A vast number of genes were identified within ovarian, prostate and breast cancer datasets and were grouped into categories related to function (Fig. 8B). HOP was found to be co-expressed with 10 other heat shock proteins or co-chaperones within breast cancers (HSPC3/HSP90AB1, HSPD1/HSP60, HSPA8/HSC70, ASHA1, HSPH2/HSPA4, HSP90AB4P, HSPH1/HSP105, DNAJB6/Mrj, DNAJB1/HSP40). The profile of genes

co-expressed with HOP in ovarian and breast cancers was similar, with the majority of genes (32.1% and 34.2%) being involved with DNA replication, transcription or translation (Fig. 8B). In both cancer types, between 10-11% of the genes were chaperones or heat shock proteins, and a similar proportion was observed (9-12%) for genes involved in translocation and cell signalling. A small percentage of genes (4-5%) were also found to encode cytoskeletal proteins. A slightly different co-expression gene profile was observed for prostate cancer.

The HSP90 Chaperone Machine Interactome Database (Hsp90Int) (Echeverría *et al.*, 2011) was used to obtain all of the known HOP protein-protein interaction data to date (Fig. 9A,B). This provided a visual representation of the network of proteins that are linked to the HOP protein (Fig. 9B) and thus generated a thorough view of the HOP interactome. As expected, molecular chaperones are the largest group of proteins that interact with HOP. However other groups of proteins included those involved in transcriptional regulation, DNA synthesis, protein synthesis, metabolism and kinases involved in signal transduction (Fig. 9A). These groups indicate proteins that are involved in the functioning of a variety of normal cellular processes. Furthermore, these proteins in turn have protein-protein interactions with a vast array of other proteins (Fig. 9B). No supporting data was found for the interaction of HOP with the syntenic genes. However the following proteins that were co-expressed with HOP in cancer were also found to interact with HOP: DNAJB1/HSP40, HSP90AA1, HSPA4, HSPA8, HSPH1, RRM1, NME2, PRPF4 (Johnson *et al.*, 1998; Scheufler *et al.*, 2000a; Ho *et al.*, 2002; Wegele *et al.*, 2003; Ewing *et al.*, 2007; Tarassov *et al.*, 2008; Mandal *et al.*, 2010).

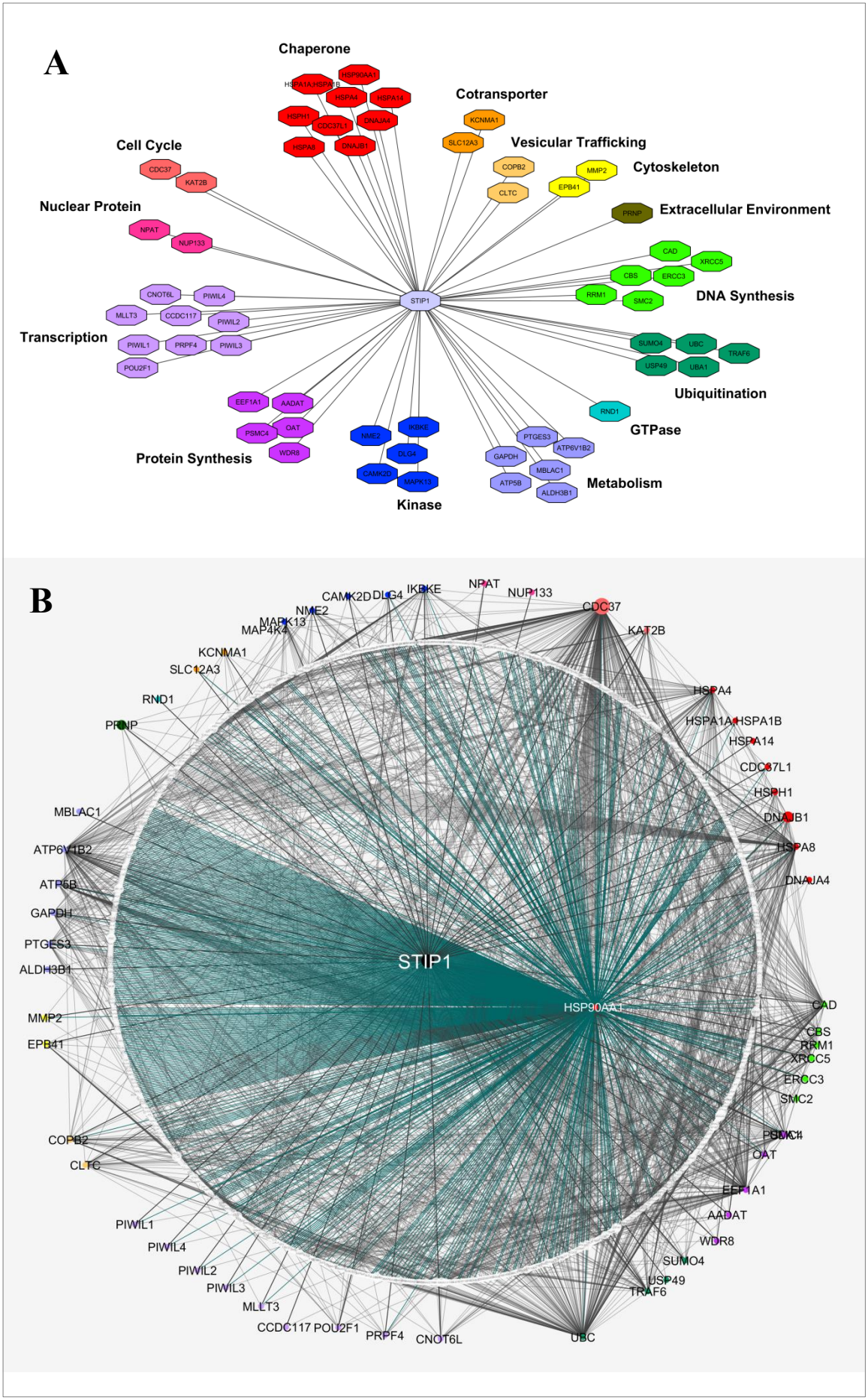

A
Chaperone
HSP90AA1
HSPA1A;HSPA1B
HSPA4
HSPA14
HSPH1
CDC37L1
DNAJA4
HSPA8
DNAJB1
Cotransporter
KCNMA1
SLC12A3
Vesicular Trafficking
COPB2
CLTC
MMP2
EPB41
Cytoskeleton
PRNP
Extracellular Environment
Cell Cycle
CDC37
KAT2B
Nuclear Protein
NPAT
NUP133
CAD
XRCC5
CBS
ERCC3
RRM1
SMC2
DNA Synthesis
CNOT6L
PIWIL4
MLLT3
CCDC117
PIWIL2
Transcription
PIWIL1
PRPF4
PIWIL3
POU2F1
STIP1
SUMO4
UBC
TRAF6
USP49
UBA1
Ubiquitination
RND1
GTPase
EEF1A1
AADAT
PSMC4
OAT
WDR8
Protein Synthesis
NME2
IKBKE
DLG4
CAMK2D
MAPK13
Kinase
PTGES3
ATP6V1B2
GAPDH
MBLAC1
ATP5B
ALDH3B1
Metabolism
B
CAMK2D
DLG4
IKBKE
NPAT
NUP133
CDC37
NME2
MAPK13
MAP4K4
KAT2B
KCNMA1
SLC12A3
HSPA4
RND1
PRNP
HSPA1A
HSPA1B
HSPA14
CDC37L1
HSPH1
DNAJB1
MBLAC1
HSPA8
ATP6V1B2
DNAJA4
ATP5B
GAPDH
PTGES3
STIP1
ALDH3B1
HSP90AA1
CAD
MMP2
CBS
RRM1
EPB41
XRCC5
ERCC3
SMC2
COPB2
CLTC
OAT
EEF1A1
PIWIL1
AADAT
PIWIL4
WDR8
SUMO4
PIWIL2
USP49
PIWIL3
TRAF6
MLLT3
CCDC117
POU2F1
PRPF4
CNOT6L
UBC

Figure 9: The regulation of HOP expression influences more than a single gene.

A) The visualisation of the individual proteins HOP has been found to directly interact with. The nodes represent proteins which are connected to HOP (shown as STIP1) via edges. The nodes are colour coded to protein function. The edges represent protein-protein interactions. B) The HOP protein-protein interaction network (HOP Interactome). The black node in the centre represents HOP (shown as STIP1). Coloured nodes represent the proteins that directly interact with HOP and are coloured coded as in (A) by protein function. White nodes represent all other proteins in the network. The size of the node corresponds to the degree of connectivity of that protein within the Interactome (the greater the degree the larger the node size). The edges connecting nodes for the protein-protein interactions between HOP and various proteins are in black, the edges connecting nodes for the protein-protein interactions between HSP90 (shown as HSP90AA1) and proteins within the HOP Interactome are in teal. All other edges are in grey and depict the protein-protein interactions between all other proteins. Data was gathered from the HSP90 Chaperone Machine Interactome Database (Echeverría *et al.*, 2011) and visualised using Cytoscape 3.0 (Cline *et al.*, 2007).

3.2.6 - Sequence analysis and identification of promoter elements upstream of the HOP gene

Despite growing evidence for a role for human HOP in cancer, relatively little is known about the mechanisms that control HOP expression. It is hence essential to look into the factors influencing the gene expression of HOP proteins. To do so, one would be required to locate and characterise the core promoter of the HOP gene *in silico* and *in vitro*. Once the promoter region has been identified, *cis*-elements could then be identified. This could be done by searching for motifs on the promoter sequence that correspond to various transcription factor binding sites (TFBS). The identification of these motifs is essential in understanding the gene expression regulation of the promoter. However, the region corresponding to the HOP promoter has not been demonstrated *in vitro*, there is no *in vivo* information to be found on the DNA sequence encoding the HOP promoter and there is a paucity of data on the factors that control HOP expression in cancer. Therefore, bioinformatic analysis of the putative HOP promoter was performed, in order to predict the regulatory DNA region responsible for controlling HOP expression, and allow for an initial investigation into the factors that alter the rate of the transcription of HOP in cancer cells.

In order to determine any conserved regions upstream of the HOP gene, the sequences corresponding to the region of the HOP gene from -500 to +100 bp across four species were aligned and compared using T-Coffee (Notredame *et al.*, 2000) (Fig.10). No specific regional conservation was observed between the human, mouse, yeast and zebrafish sequences. A further pairwise global alignment of this region by BioEdit (Hall, 1999) gave human-mouse identity as 67%, human-yeast as 43%, human-

zebrafish as 40%, mouse-zebrafish as 44%, mouse-yeast as 43% and zebrafish-yeast as 45%.

The sequences were searched for specific elements that were known to signify the presence of a promoter (Liu *et al.*, 2008; Anish *et al.*, 2009). Putative TATA boxes were identified in murine (TATATATA) and yeast (TATAAAAG) sequences at positions -102 and -30 from the transcriptional start sites (TSS), respectively (Fig.10). Both of these sequences matched the TATA consensus sequence of TATA(T/A)A(A/T)(A/G), but neither matched the more stringent TATA consensus sequence of TATA(A/T)A(A/T). The human sequence had two and the murine sequence a single GC box (encoding both GGGCGG and stimulatory protein 1 (Sp1) sites) and 6 repeats of CCAAT sequences upstream of the TSS (Fig.10). Zebrafish had a single CCAAT sequence upstream of the TSS, yeast also had a single CCAAT sequence but it was located downstream of the TSS (Fig.10). Since half of the sequences did not contain a TATA box, the sequences were scanned for CpG islands. The region of -452 to -137 (316 bp) in the human sequence met the criteria of a CpG island, as did the region of -240 to -47 (194 bp) in mouse (Fig.10). The criteria were a frequency of observed CG dinucleotides / frequency of expected CG dinucleotides that was greater than 0.6, as well as a GC % greater than 50 and a length greater than 100 bp (Li & Dahiya, 2002).

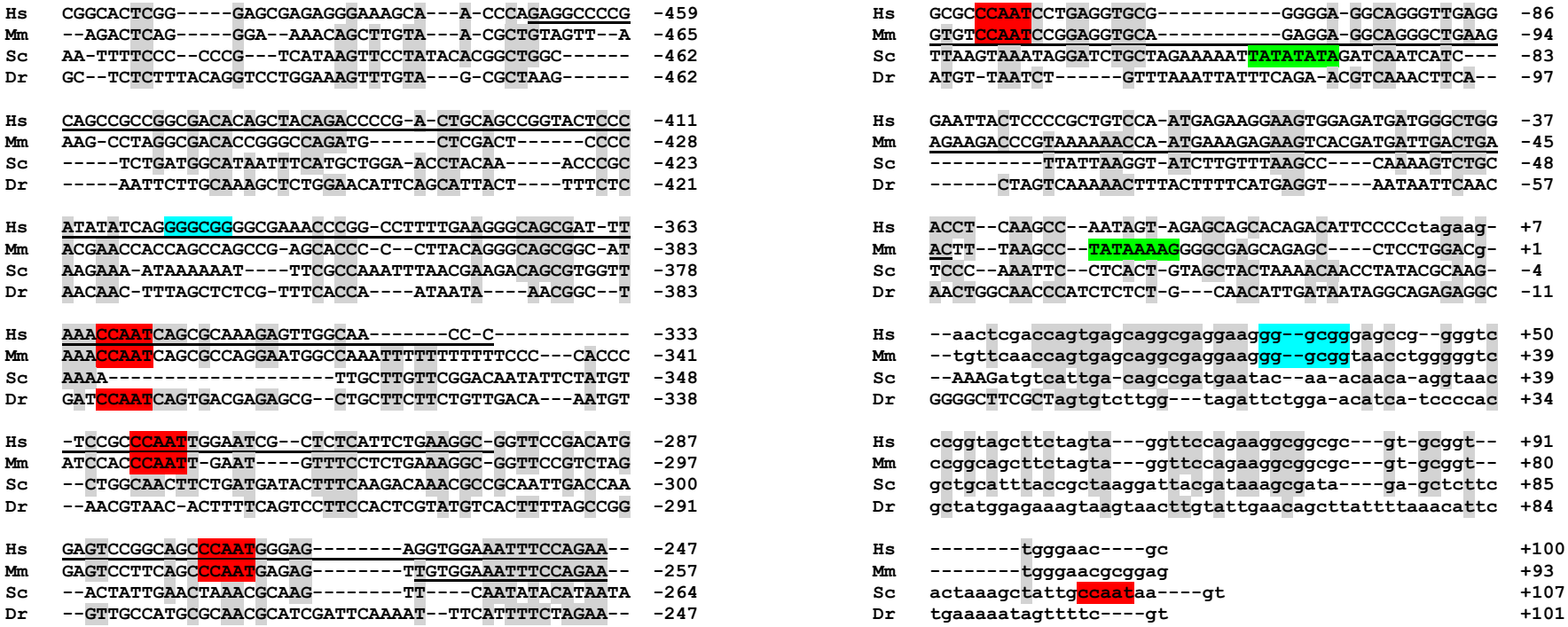


Figure 10: Sequence alignment of HOP genomic DNA to identify a putative promoter.

T-Coffee alignment of human (Hs), murine (Ms), yeast (Sc) and zebrafish (Dr) upstream sequences of the HOP genes. Sequences aligned from -500 to +100 with respect to the transcriptional start site (TSS). The putative TATA boxes were coloured green, the GC boxes (GGGCGG) were coloured cyan and the CCAAT boxes were coloured red. The TSS was indicated by the transition of uppercase to lowercase lettering. Nucleotides that were conserved between at least three orthologues were coloured grey. DNA encoding CpG islands were identified by MethPrimer (Li and Dahiya, 2002) and were underlined. BioEdit (Hall, 1999) gives human-mouse identity as 67%, human-yeast as 43%, human-zebrafish as 40%, mouse-zebrafish as 44%, mouse-yeast as 43% and zebrafish-yeast as 45%.

3.2.7 - Analysis of Transcription Factor Binding Sites in the human HOP promoter

The putative human HOP promoter sequence identified above, including 2000 bp upstream, was analysed to identify potential TFBS. Cister revealed that the region containing the highest density of *cis*-elements was found to be within 1500 bp upstream of the gene itself (data not shown). *Cis*-elements are known to be found upstream of the TSS of genes, or within the promoter itself. The increased density of *cis*-elements within the upstream region of the TSS suggested this region may have the potential for promoter activity. Additionally, the GC content of the majority of the upstream 1.5 kb region was above 50%. Hence, the analysis was focussed on this region that could potentially contain a promoter (Fig.11). Three prediction programs were utilised to investigate the -1500 to +100 bp region of the human HOP gene (Appendix A). Although minimal correlation was found across the predictions by TFSearch, PROMO and Alibaba, the high cut off thresholds used in the prediction aimed to reduce the number of false positives identified. A total of 480 putative TFBS were identified. Stimulatory protein 1 (Sp1) had the highest prediction rate, predicted to have 73 motifs; while CCAAT enhancer binding proteins (C/EBP) was the second most frequent motif at 55 predicted sites (Fig.11).

Figure 11 also shows the density of the putative *cis*-elements across the region, many of which were found in clusters. The greatest density of *cis*-elements was observed within the first 500 bp of the putative promoter. Core promoter elements such as $CEBP^{\alpha/\beta}$, NF-Y and TFIID were found in highest abundance within the first 600 bp of the putative HOP promoter (Fig.11). The mapped STAT4, AP-2, p53 and Pax-5 elements were found to be spread throughout the putative promoter. NFKB and Egr-1 were found to be restricted to the first 900 bp upstream from the TSS. Four probable HSEs for HSF1 and HSF2 were identified. Each HSF1 element was overlapped by an HSF2 element (Fig.11). Additionally, Figure 11 highlights the *cis*-elements for various TFs that are activated by oncogenic pathways found within the putative HOP promoter. Of note were the single sites were observed for c-Myc, c-Fos and c-Jun, and the five sites observed for ETS-1, which may provide a link to the increased expression of HOP in cancers.

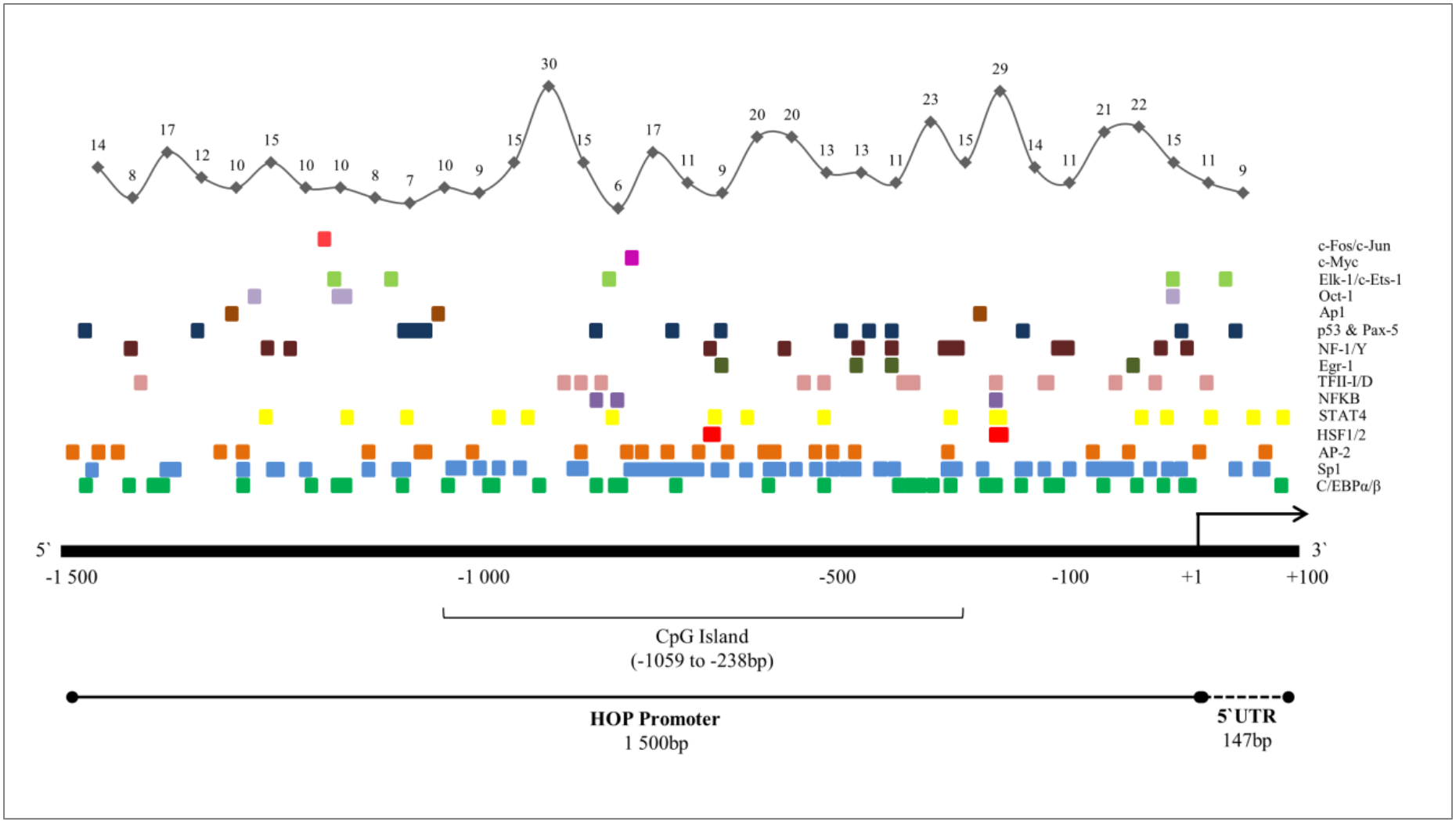


Figure 11: Putative transcription factor binding sites within the HOP promoter.

The line graph illustrates the density of the putative transcription factor binding sites within 50 bp segments of the putative HOP promoter sequence.

The distribution of selected putative transcription factor binding sites of interest was shown as boxes correlating to their position along the promoter (drawing is not to scale). Numbers indicate position relative to the transcription start site (+1). CCAAT/enhancer-binding proteins (C/EBP) shown in green, stimulatory protein 1 (Sp1) shown in blue, activator protein 2 (AP2) shown in orange, heat shock transcription factor 1 and 2 (HSF1/2) shown in red, signal transducer and activator of transcription 4 (STAT4) shown in yellow, nuclear factor kappa B (NFKB) shown in lilac, general transcription factor (TFII-I/D) shown in pink, early growth response 1 (Egr-1) shown in dark green, nuclear factor I or CCAAT-binding transcription factor (NF-1/Y) shown in maroon, p53 tumour suppressor protein (p53) and the paired box transcription factor 5 (PAX-5) shown in navy blue, Activator protein 1 (Ap1) shown in brown, octamer binding transcription factor (Oct-1) shown in lilac, member of the ETS oncogene family (Elk-1) and ETS domain transcription factor (c-Ets-1) shown in lime green, the c-Myc shown in purple, c-Fos and c-Jun shown in magenta. The CpG islands with GC content > 50% shown. Transcription factor binding sites identified by Alibaba, TFSearch and PROMO prediction tools. Note: Results for the CpG island prediction differs from that in Fig. 9 due to the size difference of the sequences analysed.

3.3 Discussion

3.3.1 - HOP is conserved across taxa and is syntenic across *H.sapiens, M.musculus* and *X.tropicalis*

The topology of a branch in a phylogenetic tree is considered correct if the bootstrap value is above 95%, and those nodes with less than 70% are not considered reliable (Hall, 2013). With this in mind, some of the branches within the tree generated in this study were true reflections of historical relationships while others were not. Furthermore, although the leftmost node within the phylogenetic tree could be presumed the common ancestor, this is incorrect as the tree is unrooted (Hall, 2013). The common ancestor of the HOP gene is unknown. However, we could determine the extent to which the homologues have changed by using the branch lengths. All branch lengths were very low implying that overall, from fungi through to humans, the HOP protein sequence has undergone minimal evolution (Hall, 2013). It must be noted though that *Saccharomyces cerevisiae* was separated from *Homo sapiens* at the first branching of the tree suggesting the least protein conservation between the two species.

Synteny is not usually found between genes that are not orthologues (Jun *et al.*, 2009). A defined region of a chromosome is considered to be syntenic to a chromosome of another species when at least 2 genes are homologous. Homology does not require that the genes be in the same orientation, on the same strand or co-linear (Nadeau, 1989; Jun *et al.*, 2009). This allows us to make the conclusion that the regions of the chromosome containing the HOP gene in *H.sapiens*, *M.musculus* and *X.tropicalis* are syntenic and that the organisms may share a common ancestor. The gradual loss of synteny observed between the remaining orthologues could be explained by deletions, rearrangements and insertions that have

occurred over time with evolution (Jun *et al.*, 2009), explaining the complete lack of synteny observed for the yeast, nematode and fruit fly orthologues. Additionally, synteny between the ortholgoues decreased as the branching between the orthologues on the phylogenetic tree increased. This further illustrated the evolutionary progress from the unicellular organisms to the mammals. The prokaryote gene order has been found to be more variable than that of eukaryotes (Jun *et al.*, 2009). For example, the gene order surrounding HSP90 in prokaryotes is different to that in eukaryotes, and the HSP90 gene in prokaryotes is different to that of eukaryotes. Prokaryotic HSP90 does not require co-chaperones for function whereas in eukaryotes HSP90 co-chaperones are essential (Li *et al.*, 2012). The consequence for this study was that prokaryotes do not contain HOP.

Synteny across numerous species can identify functional relationships between the syntenic genes (Michalak, 2008). The similar expression patterns observed between the syntenic genes in human cancers further supported this likelihood. This suggested the possibility of a functional relationship between HOP, NUDT22, FERMT3, DNAJC4, TRPT1, FLRT1, VEGFB and FKBP2. This also suggested that *cis*-elements in these gene loci may be shared by the syntenic genes, allowing for the possibility of co-regulation or co-expression. Furthermore, many of the syntenic genes were cancer related, hence it being of interest that they were in such close proximity. The first syntenic gene of interest was the DnaJ (HSP40) homologue, subfamily C, member 4 (DNAJC4). DNAJC4 (HSPF2/MCG18) is a molecular chaperone that interacts with HSP70, which is a partner protein of HOP (Sugito *et al.*, 1995; Calderwood, 2013). DNAJC4 is a type III HSP40, containing the J-domain, which is a 70 amino acid signature region that stimulates the ATPase activity of HSP70 and provides an interaction platform for HSP70 (Kampinga & Craig, 2010). The only other domain structure that DNAJC4 has is a transmembrane domain. As yet both the function and localisation of DNAJC4 within the cell is unknown (Kampinga & Craig, 2010). The second gene of interest was FK506-binding immunophilin 2 (FKBP2), as it too is a chaperone protein (Scammell *et al.*, 2003). It acts as a co-chaperone containing a TPR domain and plays a role in immunoregulation. It is able to bind HSP90 and is known to interact with dynein, suggesting that it may increase nuclear transport within the cell (Riggs *et al.*, 2007). Furthermore, FKBP2 has been found to play a role in breast and prostate cancers (Calderwood, 2013). The third gene of interest was the vascular endothelial growth factor B (VEGFB) for it is involved in regulating the growth of endothelial cells, but has been found to play an important role in inducing tumour angiogenesis and is expressed in various cancers (Cheung *et al.*, 1998;

Ruohola *et al.*, 1999; Shintani *et al.*, 2004). VEGFB has been identified in stimulating the formation of blood vessels within tumours (Shintani *et al.*, 2004).

Despite the synteny observed across *H.sapiens*, *M.musculus* and *X.tropicalis,* each species contained a different promoter. A reason for the lack of conservation observed, especially between the murine and human promoters, may be due to the fact that CpG-rich promoters are significantly less conserved than CpG-poor promoters (Wang & Hannenhalli, 2006). Furthermore, the presence of a TATA box in the murine system suggests the murine HOP may be more so a context- or tissue-specific gene, while human HOP, in part due to the large CpG island, is less so (Sandelin *et al.*, 2007). However, the conservation observed between the human and murine promoters was much greater than between human and frog or yeast. Additionally, the conservation of the GC and CCAAT boxes within the human and murine promoters suggest that although the promoters are different between the two species, they may be controlled by similar signalling networks. Thus if human xenografts are utilised in murine systems, one may overcome the problems of differing promoters between the species.

3.3.2 - There are potentially multiple isoforms of HOP in eukaryotes

In the extracellular environment of cells, HOP functions like a cytokine rather than a co-chaperone and is known to activate a variety of signalling pathways (Caetano *et al.* 2008; Arantes *et al.* 2009; Beraldo *et al.* 2013; Wang *et al.* 2010; Tsai *et al.* 2012). More specifically, HOP has also been identified to act as the receptor to the Prion protein (PrP^c). It acts independently of the HSP90 and HSP70 chaperones, and instead interacts directly with the PrP^c. This interaction between extracellular HOP and the PrP^c has a series of consequences on cell growth, survival and differentiation; mediating these effects by the activation of downstream signalling pathways (Lima *et al*. 2007; Zanata *et al*. 2002; Arantes *et al*. 2009). Little is known about how intracellular HOP is exported from the cell, although it is currently assumed that extracellular HOP is derived from the intracellular isoform. These two drastically different functions of HOP do not seem consistent with a single isoform. Our analysis predicts the presence of possible alternative transcripts of HOP leading to distinct protein isoforms. Therefore, it may be possible that there is in fact more than one isoform of HOP being produced by the cell, and that the intracellular and the extracellular versions of HOP are distinct. To date there is no experimental data beyond presence of the transcripts to

prove such a theory, yet it is unknown how the single isoform could adapt to such differing processes.

The presence of alternative HOP isoforms was also supported by comparing the exon and protein organisation of the HOP orthologues. It was observed that the distribution of the exons had no regularity. However, the distribution of the TPR domains across numerous exons in the human suggested that it could be possible for the alternate splicing of the HOP mRNA transcripts to give rise to the different HOP isoforms within human and murine species. Upon searching for possible splice variants, several had been predicted in the available databases, although most of these predicted isoforms retained the same order of exons. STIP1-010 was the most similar to the known HOP (STIP-001) except it contained an additional N terminal extension. A possible reason for such similarity may be due to the use of an alternative promoter (alternative transcription initiation) or intron retention (Srebrow & Kornblihtt, 2006). The consequence of the exclusion of exon 2 within STIP-003, which is predicted to result in a missing third TPR motif within the TPR1 domain, may affect the binding specificity of the protein for HSP70 (Onuoha *et al.*, 2008). The truncated transcripts STIP1-008 and STIP1-012 suggest a vast exon deletion in each but resulting in a maintained binding specificity for HSP70 while forfeiting binding specificity for HSP90 (Onuoha *et al.*, 2008). These splice variants are reminiscent of the HOP orthologue within *C.elegans*. Nematodes contain merely two TPR domains and 1 DP domain within the HOP gene, yet the gene still encodes a functional HOP protein. More specifically it lacks TPR1 domain and the linker region that connects it to TPR2A. Despite the missing regions, the HOP within *C.elegans* is still able to interact with HSP70 and HSP90 although only the interaction with HSP90 is strong. The consequence of this is that in *C.elegans* HOP cannot bind HSP70 and HSP90 simultaneously as HSP90 requires both interaction sites on the remaining TPR domains to regulate ATPase activity (Gaiser *et al.*, 2009).

These transcripts suggest that there may be isoforms of HOP with differing enzymatic activity or differing substrate specificity to the known HOP protein. These isoforms could have the ability to interact with a differing set of substrates and proteins entirely (Möröy & Heyd, 2007). These isoforms could also have differing subcellular localisation or be tissue specific. This raises a question of whether a particular HOP isoform could be partially or indirectly responsible for malignant transformations (Venables, 2004; Möröy & Heyd, 2007). A single point mutation in the genomic DNA is all that is required to alter the proteins

influencing the splicing mechanism of a gene, this being a probable possibility within cancer cells (Krawczak, Reiss & Cooper 1992; Cartegni, Chew & Krainer 2002; Faustino & Cooper 2003). Cells have been known to be transformed by the accumulation of the overexpression of the alternately spliced protein isoforms such as Ron or Rac1 (Srebrow & Kornblihtt, 2006), suggesting the possibility of the build-up of HOP isoforms having the potential to do the same. However, the high sequence similarity may make it difficult to identify the different transcripts using microarray or qPCR approaches.

The most direct interaction the genome has with its environment is through TFs (Thorne *et al.*, 2009). Thus an additional possibility is that distinct isoforms could be expressed in differing cell types or tumour cells due to the specific nature of the transcription machinery in those cells. Various TFs have been found to be overexpressed in cancer cells (for example: c-myc, STATs, ETS, E2F) and the disruption of the equilibrium of TFs in malignant cells influences the regulation of a gene's promoter (Ponzielli *et al.*, 2005; Seth & Watson, 2005; Tsantoulis & Gorgoulis, 2005; Thorne *et al.*, 2009; Santos & Costa-Pereira, 2011). Thus the identification of the splice variants of HOP also suggests the possibility of HOP isoforms with differing subcellular localisations. It would be of great interest to take this idea further and attempt to identify the presence of these splice variants at a protein level.

3.3.4 - The HOP gene has large interaction and co-expression networks

The Gemma gene expression database allowed for the identification of *in vitro* evidence of human HOP being co-expressed with two syntenic genes, MACRO domain containing 1 (MACROD1 or LRP16) and TRPT1. The TRPT1 protein is responsible for catalysing tRNA splicing, while MACROD1 acts as a transcriptional co-activator for the estrogen α- and androgen receptors (ERα and AR) (Spinelli *et al.*, 1998; Hopper *et al.*, 2010; Han *et al.*, 2011). Interestingly, ERα and AR are both HSP90 client proteins (Nemoto *et al.*, 1992; Bouhouche-Chatelier *et al.*, 2001; Lee *et al.*, 2002). In a proteomic study on the effect of knockdown HOP in transformed cells, when looking specifically at the HOP syntenic genes we observed that TRPT1 protein levels were unchanged, while MACROD1 and FKBP2 were both up-regulated though not significantly (personal communication). This was interesting as the co-expression data would suggest that both MACROD1 and TRPT1 (and possibly the other syntenic genes) would be deregulated upon the knockout of HOP. Additionally it has been shown that MACROD1 is overexpressed in human cancers, with specifically high levels

in endometrial carcinoma, gastric carcinoma, colorectal carcinoma and breast carcinoma (Liao *et al.*, 2006; Meng *et al.*, 2007; Li *et al.*, 2009; Xi *et al.*, 2010; Zhao *et al.*, 2010).

The co-expression profile of human HOP specifically within breast, ovarian and prostate cancers identified numerous co-expression networks of functionally related genes that had included the human HOP gene. These cancer subtypes were chosen because HOP has been found to be up-regulated in each of them (Wang *et al.*, 2010; Tan *et al.*, 2011; Willmer, 2011). To further substantiate this, it would be of interest to look at the promoter profiles of the co-expressed genes to determine if they have identical over-expressed transcription factors that regulate them; thereby determining whether those co-expressed genes may also be co-regulated. For example, 47 gene promoters were found to share the same promoter profile as the ionotropic AMPA glutamate receptor (GRIA) and of them 16 were also found to be co-expressed with GRIA (Chong *et al.*, 2007). Of further interest was that HOP was found to be co-expressed with DNAJB1/HSP40, HSPH2/HSPA4, HSPH1/HSP105, HSP90AA1, HSPA8, RRM1, NME2 and PRPF4; and HOP was also found to interact with these proteins (Echeverría *et al.*, 2011).

The generation of the HOP Interactome illustrated the magnitude of proteins that may be influenced by the up-regulation of the HOP gene in malignant cells. This is not saying that these proteins are dependent on HOP directly for stability. However HSP90 is considered a critical facilitator of cancer cell survival and oncogene addiction because it protects mutated and/or overexpressed proteins from degradation or misfolding (Bagatell & Whitesell, 2004; Sangster *et al.*, 2004; Mizrak *et al.*, 2006; Trepel *et al.*, 2010). Thus the over expression of HOP in cancers, coupled with the over expression of HSP90 in cancers, would allow for the stabilization of tumorogenic cells via the protection of many of the proteins within the HOP Interactome. This provides further evidence of the importance of HOP in cancer and the suitability of HOP as a novel drug target. Additionally, in order to accurately determine the effect of overexpressed HOP in malignant cells, we need to elucidate the molecular networks involving HOP in the cell. By doing so one could begin to identify the protein-protein interactions that may be affected to some degree by HOP inhibition. A practical example of the importance of such knowledge was shown by Fierro-Monti and colleagues (2013). Fierro-Monti *et al.* (2013) quantified the changes in the HSP90 dependent proteome upon HSP90 inhibition by geldanamycin. Through this proteomic study they identified a large cohort of previously unknown and unexpected side effects of HSP90 inhibition. They identified a

consequent down-regulation of the protein synthesis machinery as well as a remodelling of the HSP90 chaperone machine; both having potentially far-reaching consequences. Furthermore, they identified an enrichment of several oncoproteins and a parallel depletion of several tumour suppressors which were previously unidentified. This was concluded to be a pro-survival response which then suggested that there are unwanted implications for the use of geldanamycin derivatives in anti-cancer therapy (Fierro-Monti *et al.*, 2013). This works to show the importance of having a thorough understanding of the effect of the inhibition of a protein on the proteome. By proposing HOP as a potential drug target, we suggest a concurrent need to define the proteomic effect of inhibiting HOP.

3.3.5 - The HOP gene is regulated by a TATA-less promoter

For transcription to commence RNA polymerase II is required to bind in the correct orientation to core DNA elements found within a 100 bp region of the core promoter. These generally are the TATA-box or initiator element, CCAAT box and the GC-Box (Bucher, 1990; Lee & Young, 2000; Pandey & Krishnamachari, 2006). Likewise, in eukaryotes synthesis only occurs with the supplementation of certain general TFs (GTFs) (Pedersen *et al.*, 1999). The CCAAT box is one of the most common elements within promoters and is known to play an important, if not essential, role in promoter activity and the binding of CBF/NF-Y to CCAAT elements is required to recruit TFIID to the promoter (Ronchi *et al.*, 1995; Frontini *et al.*, 2002). Additionally approximately 50% of human and murine promoters are known to be associated with CpG islands, thus using the presence of CpG islands is now a common tool to identify vertebrate promoters (Li & Dahiya, 2002). A mere 10-20% of mammalian promoters have been found to contain a functional TATA box, suggesting that TATA-driven gene expression is the exception rather than the rule for eukaryote promoters (Shi & Zhou, 2006; Sandelin *et al.*, 2007). TATA boxes, instead, have been found to be restricted to sharp strong tissue-specific promoters and evolve slower than TATA-less promoters (Sandelin *et al.*, 2007). TATA-less promoters are most often found to be broad promoters characterised by a lack of TATA boxes, but with numerous Sp1 motifs, CCAAT boxes and CpG islands across the promoter region (Bjornsdottir & Myers, 2008). Additionally, broad promoters are known to contain multiple TSSs and are linked to epigenetic transcriptional control (Carninci *et al.*, 2006). The lack of a TATA box in the human HOP sequence, as well as the presence of multiple CCAAT *cis*-elements, a GC box and a CpG island, suggested it may be a TATA-less promoter. On the other hand mouse and

zebrafish may each be regulated by a TATA containing promoter. Upon further analysis of the -1500 to +100 bp region of human HOP, the presence and high abundance of Sp1 motifs within the region further supported the hypothesis that human HOP may be regulated by a TATA-less broad promoter (Carninci *et al.*, 2006).

The region of -450 to -100 bp of the human HOP was thus identified as the putative core promoter, while the region of -300 to -100 bp of the murine transcript was identified as the putative core promoter of murine HOP. However, further analysis of the presence of *cis*-elements required for basal activity was necessary to confidently assign the core promoter regions. The lack of specific regional sequence conservation observed upstream of the HOP gene across the human, mouse, yeast and zebra fish demonstrated that although the upstream sequences could share a functional correlation across species, they did not share sequence identity. The core promoter initiates basal transcription at a low rate therefore additional regulatory elements to the general TFs are required to sustain and regulate the rate of transcription in eukaryotic cells (Pedersen *et al.*, 1999; Lee & Young, 2000; Wray *et al.*, 2003). NF-Y, TFIID, AP2 and ETS are *cis*-elements commonly found within promoters (Frith *et al.*, 2001); elements which were also found upstream of the TSS of human HOP. These TFs activate transcription by recruiting basal transcription machinery to the core promoter through protein-protein interactions either directly or via adaptor proteins (Matthews, 1992; Stargell & Struhl, 1996; Ptashne & Gann, 1997; Pedersen *et al.*, 1999). Due to their presence within the upstream region of HOP, the region of the putative core promoter could be extended to the first 900 bp upstream of the TSS. However, this would need to be confirmed *in vitro*.

The clusters of TFBS indicated overlapping sites and possible competitive binding pairs between proteins. Additionally, it suggested that the elements may not all be active at once, but the functioning of the TFBS would be dependent on the physiological context of the cell at the time (Wray *et al.*, 2003). The clusters may also provide a means for the transcription factors to interact and work in conjunction with each other due to their close proximity, thereby forming composite elements allowing for further regulation of the promoter (Wray *et al.*, 2003; Chong *et al.*, 2007). The large number of *cis*-elements predicted was due to the prediction programs compromising specificity for sensitivity, thus many of the results observed would be false positives. However for the sake of this analysis that shortcoming could be over looked, for to identify a potential promoter region one needs to identify the

region upstream of the coding region of the gene of interest that contains over-represented *cis*-elements in comparison to the levels observed within the coding region itself.

The HSEs identified in the HOP promoter were of interest because HSF-1 is involved in regulating HSP gene expression. The presence of HSF-1 binding sites within the HOP promoter had already been proposed by Ruckova *et al.* (2012), and the binding of HSF-1 to the suggested HOP promoter region determined by chromatin immunoprecipitation. It was also shown that the protein levels of HSF-1 bound to the putative HOP promoter region increased in cancer cell lines upon the addition of the HSP90 inhibitor 17AAG (Ruckova *et al.*, 2012). The *in vitro* evidence from Ruckova *et al.* (2012) supports, firstly, that the region analysed in this study was in fact the HOP promoter; secondly, that the *in silico* HSF-1 sites were functional sites instead of false positives, and thirdly that the HOP promoter is a transcriptional target of HSF-1.

3.3.6 - Conclusion

Our phylogenetic analysis of HOP suggested that the gene is ancient and conserved throughout eukaryotes. This was anticipated and would suggest a conserved function. However, although the homologues are conserved, there is evidence to suggest that the structure and possibly the function of HOP in each species differs. For example, HOP is an essential gene in the mouse (Beraldo *et al*. 2013), but not in yeast (Chang *et al.* 1997). This raises the question of whether the traditional animal systems can be used as a model organism (such as yeast) when studying the function of HOP within humans. This is of vital importance when analysing HOP as a drug target in humans. This study provided a thorough analysis of the HOP gene across known orthologues so as to determine the differences between them. Murine HOP was found to be most similar to humans while our study suggests that yeast would not be the best model to use when studying the function of HOP within humans. An in depth analysis of the co-expression profile of human HOP in cancers was performed. This identified co-expression profiles for HOP that differed amongst differing cancers. Additionally, this study identified a putative promoter for HOP with a core promoter region of the first 900bp upstream of the TSS and with a proximal promoter region continuing up to at least 1 500bp upstream of the TSS. This putative promoter region became the focus of the subsequent analysis into the regulation of HOP expression.

Chapter 4

Characterisation of the HOP promoter: the regulation of HOP expression by RAS

4.1 Introduction

4.1.1 - RAS signalling in Cancer Cells

RAS genes encode for small monomeric guanine triphosphate (GTP) bound proteins that act as molecular switches, liaising between extracellular signals and intracellular signalling cascades (Guerrero *et al.*, 2000; Calvisi *et al.*, 2011; Borrego-Diaz *et al.*, 2012). Upon RAS activation, several downstream effector pathways are stimulated, such as the extracellular signal-related kinase pathway (ERK pathway), Jun amino-terminal kinase pathway (JNK pathway), p38-kinase pathway, phosphatidylinositol 3-kinase pathway (P13K pathway) and the RAS like (RAL) pathway (Borrego-Diaz *et al.*, 2012). These pathways in turn, by up-regulating downstream targets, are involved in the control of various cellular processes including cell cycle progression, apoptosis, cell migration, growth, senescence and cytoskeletal changes (Fernández-Medarde & Santos, 2011; Ho *et al.*, 2012). The balance between the crosstalk within this signalling network is essential for determining the outcome of the cellular responses within the cell (Fernández-Medarde & Santos, 2011).

RAS is found within close proximity to plasma membranes as it has fatty acid moieties itself (Borrego-Diaz *et al.*, 2012). It is characterised by two conformations (Fig.12), an inactive guanine nucleotide diphosphate (GDP)-bound and an active guanine nucleotide triphosphate (GTP)-bound state, that are controlled by guanine nucleotide exchange factors (GEFs) and RAS GTPase activating proteins (GAPs), respectively (Iwashita & Song, 2008; Calvisi *et al.*, 2011). It has been identified that approximately 33% of all human tumours express an oncogenic form of RAS, which through a desensitisation of GAPs, locks the protein into the active conformation (Calvisi *et al.*, 2011; Fernández-Medarde & Santos, 2011). These proteins become oncogenic through point mutations within the highly conserved sequence between amino acids 1-165 on codons 12 and 61 (Omerovic *et al.*, 2008; Fernández-Medarde & Santos, 2011). The three most common RAS isoforms are the HRAS, KRAS and NRAS genes, whereby the mutations render them in a permanently active conformation due to being insensitive to the RAS GAPs (Barbacid, 2008). These oncogenes were the first oncogenes to be discovered within human tumours and became the founding members of the RAS gene family (Fernández-Medarde & Santos, 2011). RAS amplification has not been linked to cancer development, rather it is oncogenic RAS mutations that have been found to almost exclusively link RAS genes to human tumour development (Fernández-Medarde & Santos, 2011).

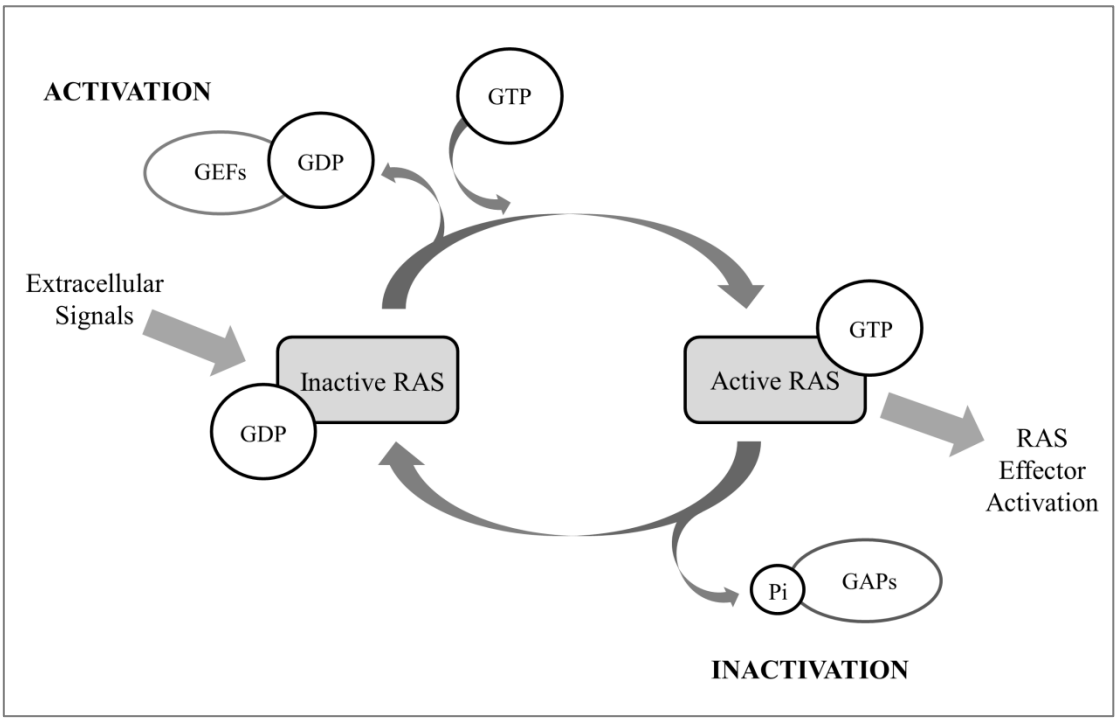


Figure 12: Regulation of RAS activation.

Schematic diagram illustrating the RAS activation cycle. RAS is in an activated conformation when in a guanine nucleotide triphosphate (GTP)-bound state, which is initiated by the exchange of guanine nucleotide diphosphate (GDP) for GTP regulated by guanine nucleotide exchange factors (GEFs). RAS is in an inactivated conformation when in a GDP-bound state, which is initiated by the exchange of GTP for GDP by RAS GTPase activating proteins (GAPs).

Various possibilities have been considered in an attempt to inhibit RAS signalling. One of the first strategies was to target the posttranslational modification responsible for the association of RAS with the plasma membrane (Casey *et al.*, 1989; Pompliano *et al.*, 1992). Farnesyltransferase (FTase) is the enzyme that catalyses the transfer of a farnesyl group from farnesyl diphosphate to the thiol of a C-terminal cysteine residue of RAS, thereby forming a thioester bond and displacing a pyrophosphate (Pompliano *et al.*, 1992). The loss of this posttranslational modification prevents the consequent *in vivo* activity of RAS. Numerous small molecules were screened and developed as potential FTase inhibitors (FTIs). However most did not pass clinical trials, as it was identified that when FTase is inhibited, RAS can serve as a substrate for geranylgeranyltransferase type I (GGTase-I). The addition of a geranylgeranyl isoprenoid group acts as an effective substitute for the farnesyl group thereby allowing membrane association (Cox *et al.*, 1992). Moving away from direct RAS inhibition, inhibitors of the downstream RAF-MEK-ERK mitogen-activated protein kinase (MAPK)

cascade were in turn considered as an indirect method of RAS inhibition. This has become an attractive target due to the non-overlapping occurrence of mutations within BRAF and RAS in cancers, suggesting equivalent roles for both mutagenic proteins in RAS-mediated oncogenesis (Dhomen & Marais, 2007).

A possible link between HOP and RAS was first proposed when an increase in HOP mRNA and protein expression was detected upon transformation of Rat1 fibroblast cells with the oncogene RAS (unpublished data; van der Spuy, 2000). Our *in silico* co-expression analysis revealed that Rho is co-expressed with HOP in human cancers. Rho proteins are a family of GTPases that are associated with the RAS signalling pathway. There are three isoforms, namely RhoA, RhoB and RhoC, that have differing but connected functions in cell migration (Karnoub *et al.*, 2004). In particular, the expression of RhoC has been found to increase metastasis in cancer cells and the expression of RhoC was found to increase as tumours become more metastatic (Ridley, 2004; Wheeler & Ridley, 2004; Wu *et al.*, 2004). Previous research from our group showed that the knock-down of HOP in cancer cells caused a reduction in the expression of RhoC (Willmer *et al.*, 2013). The previous chapter revealed *cis*-elements within the putative HOP promoter region that were activated by oncogenic pathways (c-Myc, c-Fos, c-Jun, ETS-1). Hence this data that linked HOP to RAS, together with the previous findings from our research unit, led to the investigation of the effect of the RAS signalling pathway on the regulation and production of HOP. This chapter presents an *in vitro* study of the HOP promoter as well as an analysis of the mechanisms involved in the up-regulation of HOP in cancer cells with a focus on RAS as a mechanism of HOP regulation.

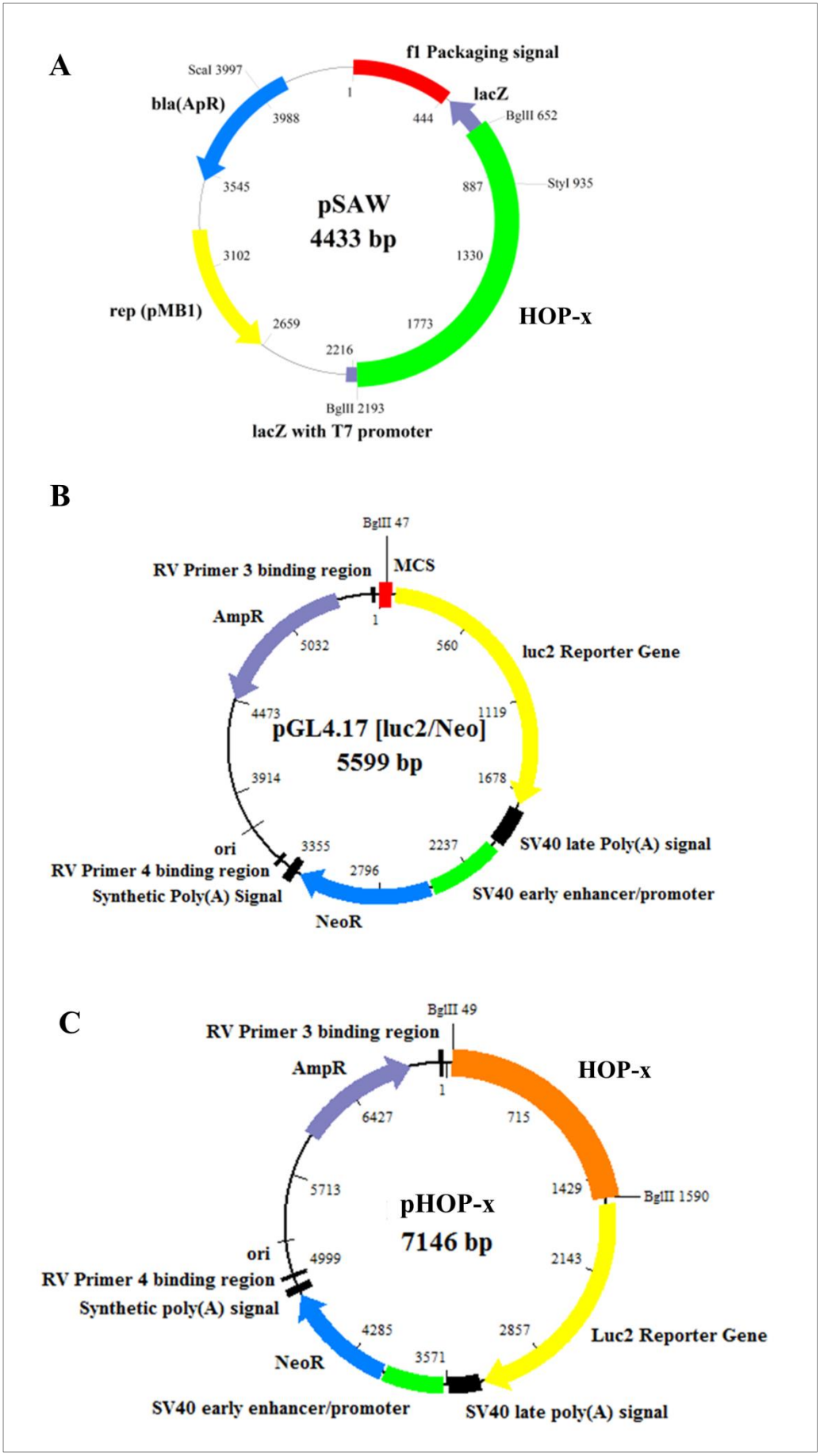
4.2 Results

4.2.1 - The upstream region of the human HOP gene encodes a functional promoter *in vitro*

To confirm that the putative HOP promoter identified in Chapter 3 was functional *in vitro*, a luciferase reporter construct containing the -1520 to +16 bp region of the human HOP gene (numbering relative to the TSS) was produced (pHOP-x) (Fig.13). This enabled an analysis of whether the region upstream of the HOP gene had promoter activity.

The pSAW plasmid was synthesised by GenScript, which was a pUC57 vector containing the 1.5 kb sequence upstream of the TSS of the HOP gene (HOP-x) (Fig.13A). pSAW was digested with *Bgl*II and *Nhe*I restriction enzymes (RE) to release HOP-x. pGL4 (the promoterless vector) was digested with the same RE. HOP-x and pGL4 were gel purified and quantified by the Nanodrop 2000. The HOP-x insert was subsequently ligated into pGL4 cloning vector (Fig. 13B,C). The ligation product was transformed into competent JM109 *E.coli* cells and the plasmids were extracted from 8 successful transformants (identified by ampicillin screening). To confirm the identity of the plasmids, the plasmids were digested with restriction enzymes and the band patterns observed were analysed (Fig. 13D). Endotoxin free plasmid extractions were performed for each plasmid required for the luciferase assays (pGL4 and pHOP-x, as well the transfection control plasmid pLV-eGFP). The plasmid identities were again confirmed by restriction enzyme digestion and DNA sequencing.

The promoterless pGL4 vector was digested with *Bgl*II and two bands were observed; the first being at 6500 bp and the second at 5599 bp (Fig.13D). The size of the plasmid was 5599 bp and hence the larger band observed was considered to be due to partial digestion of the DNA. The pSAW plasmid, containing the synthesised putative HOP promoter sequence, was digested with *Bgl*II and *Nhe*I resulting in two bands observed at 2892 and 1541 bp (Fig.13D). The sizes correlate to the size of the HOP-x insert and the backbone of the pUC57 plasmid. Lastly, pHOP-x (containing the putative HOP promoter sequence upstream of the luciferase coding sequence) was digested with *Bgl*II and *Nhe*I where four bands were observed (Fig.13D). The first two bands of approximately 7000 and 6000 bp were considered to be due to partial digestion, while the remaining two bands observed at approximately 5605 and 1541 bp were identified as the HOP-x promoter insert and the pGL4 backbone. The plasmid was confirmed by sequencing and subsequently used in reporter assays to assess the ability to drive expression of the luciferase protein.


A
f1 Packaging signal
ScaI 3997
lacZ
bla(ApR)
1
3988
444
BglII 652
3545
887
StyI 935
pSAW
4433 bp
3102
1330
rep (pMB1)
2659
1773
HOP-x
2216
BglII 2193
lacZ with T7 promoter
B
BglII 47
RV Primer 3 binding region
MCS
AmpR
1
5032
560
luc2 Reporter Gene
4473
1119
pGL4.17 [luc2/Neo]
5599 bp
3914
1678
ori
3355
2237
SV40 late Poly(A) signal
RV Primer 4 binding region
2796
Synthetic Poly(A) Signal
SV40 early enhancer/promoter
NeoR
C
BglII 49
RV Primer 3 binding region
AmpR
HOP-x
1
6427
715
5713
1429
BglII 1590
pHOP-x
7146 bp
ori
4999
2143
RV Primer 4 binding region
Synthetic poly(A) signal
4285
2857
Luc2 Reporter Gene
NeoR
3571
SV40 early enhancer/promoter
SV40 late poly(A) signal

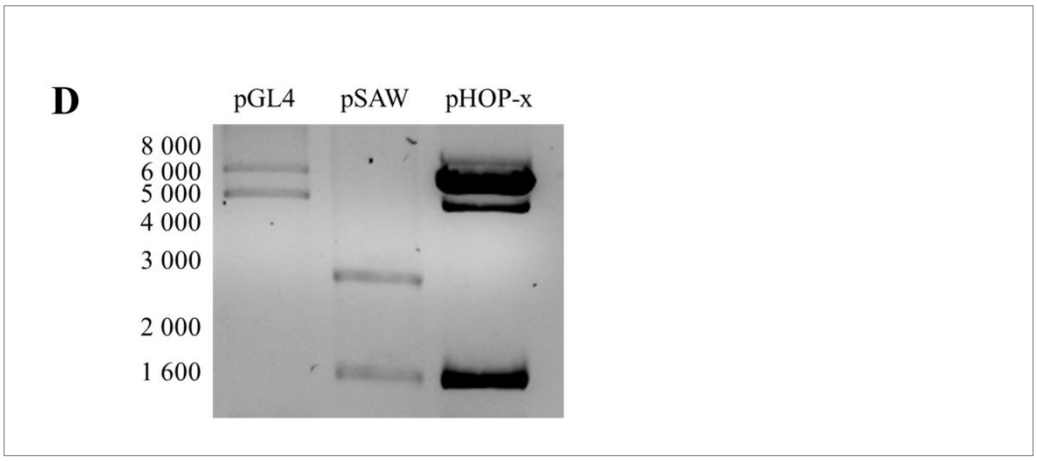


Figure 13: Generation of a Luciferase Reporter Construct for the HOP promoter.

A) Plasmid map of pSAW synthesised by GenScript. pSAW contained the 1.5 kb region upstream of the HOP gene, labelled as HOP-x. B) Plasmid map of pGL4.17, the empty luciferase reporter plasmid. C) HOP-x was cloned out of pSAW and into the pGL4 vector to produce pHOP-x, the reporter construct containing the putative HOP promoter. D) Restriction digestion confirmed the plasmids required for the luciferase assay. Lane 1 contained the contained pGL4 digested with *Bgl*II to lineraise the plasmid, lane 2 contained pSAW digested with *Bgl*II and *Nhe*I thereby releasing HOP-x from the backbone, lane 3 contained pHOP-x digested with *Bgl*II and *Nhe*I thereby releasing HOP-x from the construct. Numbering on the left hand side represent the positions of the molecular weight marker used. Bands were resolved by 1% (w/v) agarose gel electrophoresis in TAE run at 90 V for 90 minutes and visualised by ethidium bromide staining under UV light. Plasmid maps constructed using BioEdit (Hall, 1999).

The luciferase assays were performed 48 hours after transfection of the reporter plasmid (pGL4 or pHOP-x) and transfection efficiency reference plasmid (pLV-eGFP) into three cell lines; a transformed mouse embryonic fibroblast cell line (MEF-1), a human cervical cancer cell line (HeLa) and a human breast cancer cell line (Hs578T) (Fig. 14). The firefly luciferase activity was normalised against the GFP fluorescence produced by the reference plasmid, then further normalised against the control plasmid pGL4 to determine the relative promoter activity. The reporter plasmid in the murine cells produced minimal promoter activity. The reporter plasmid showed between 5 and 8 fold higher promoter activity in the cancer cell lines. The Hs578T cell line produced promoter activity 0.5 times greater than that observed within the HeLa cell line (Fig.14). These data confirmed that the HOP gene sequence (from -1520 to +16) contained a functional promoter.

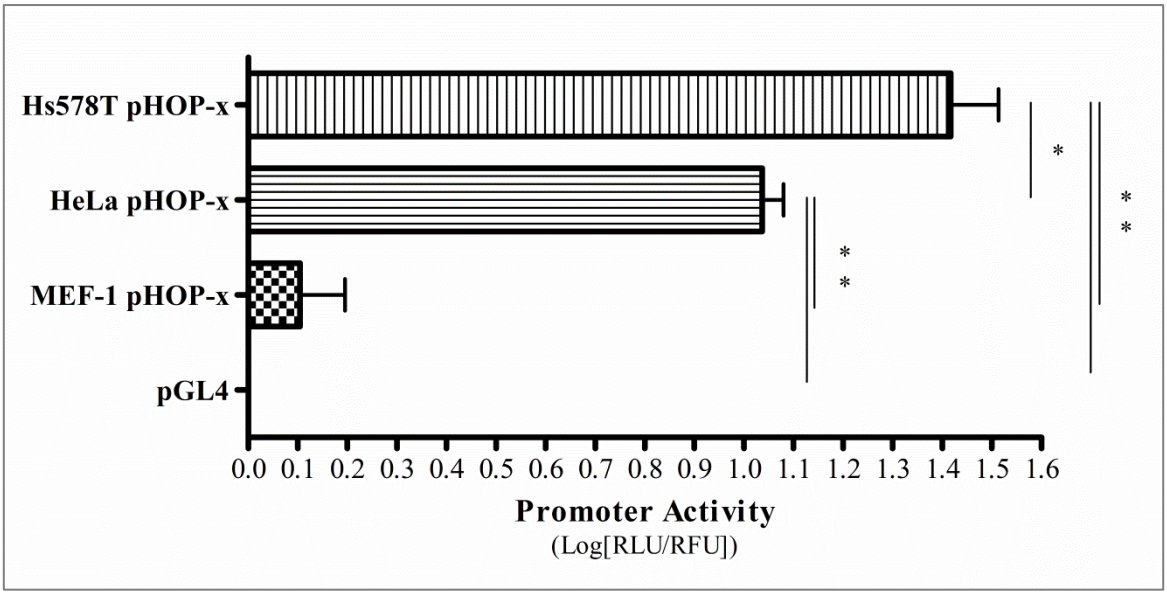


Figure 14: The upstream region of the HOP gene encodes a functional promoter.

The 1.5kb upstream region of the human HOP gene was analysed for promoter activity using a luciferase reporter construct (pHOP-x). The promoter activity was compared across two cancer cell lines and a murine transformed line; a human cervical cancer cell line (HeLa) and a human breast cancer cell line (Hs578T), and a SV40 transformed mouse embryonic fibroblast (MEF-1) cell line. Luminescence was measured in relative light units (RLU). The luciferase activity was determined by normalising the firefly luciferase activity against the fluorescence produced by the control reporter plasmid (pLV-eGFP). The results from the HOP promoter analysis were normalised against the results of the empty vector (pGL4). The results are represented as the mean values of three independent biological replicates with three technical replicates for each. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (*P < 0.01, **P<0.001).

4.2.2 – Consequences of the RAS signalling cascade on HOP expression

Gemma, as well as the prediction tool CoExSearch (Obayashi *et al.*, 2008), found Rho to be co-expressed with HOP. Gemma found Rho to be co-expressed in both ovarian and breast cancers, but not in prostate cancer. Thus a link to the RAS signalling pathway had been predicted from the *in silico* study of the HOP promoter and the co-expression analysis. In order to investigate the effect of mutated RAS on HOP production, mutated RAS was introduced into cell lines that contained no known RAS mutations. Three RAS plasmids were utilised: wildtype HRAS gene, constituently active mutated HRAS (G12V) or a dominant negative mutated HRAS (S17N) (Yasuda *et al.*, 2006). Cells would already contain the HRAS gene, however by introducing the G12V HRAS we would determine if there was any effect on HOP expression in the presence of a constitutively active RAS protein. G12V HRAS would be insensitive to the regulative signals that usually control RAS functioning. Introducing the dominant negative S17N HRAS into cells would act as a control. S17N

HRAS should not cause the effects observed in both wildtype HRAS and constitutively active HRAS as it cannot be activated (Yasuda *et al.*, 2006).

The MCF-7 and HeLa cell lines were chosen due to the absence of mutated RAS in these cell lines (Sepp-Lorenzino & Rosen, 1998). The Hs578T cell line has the Q61L RAS mutation (Eckert *et al.*, 2004). MCF-7 breast carcinoma cells were transfected with one of three RAS plasmids: wildtype HRAS, G12V HRAS or S17N HRAS. qRT-PCR was performed on the RNA extracted from each MCF-7 sample and the relative quantity of HOP mRNA present in each determined (Fig. 15A). Each of the RAS plasmids increased the levels of HOP mRNA compared to MCF-7 cells without the mutated RAS (Fig. 15A), however the increases were not significant. HRAS G12V induced the greatest increase in expression of HOP mRNA with a 2.5 fold difference, while HRAS S17N induced no change in HOP mRNA levels. Wildtype HRAS caused a mean increase less than 1.5 fold.

HeLa and Hs578T cell lines were subsequently transfected with the RAS plasmids and the effect on the HOP promoter activity was analysed (Fig. 15B). Additonally, cell lysates were produced from each transfection and the levels of HOP protein within each determined using SDS-PAGE and Western Blot (Fig. 15C,D). Mutated HRAS had no effect on the HOP promoter activity within the HeLa cells when compared to that of the HOP promoter alone (Fig. 15B). However HRAS caused a significant increase in the HOP promoter activity of up to 6 fold within the Hs578T cells when comparing to the HOP promoter alone. HRAS caused a 5 fold increase in the HOP promoter activity, G12V HRAS caused a 6 fold increase and S17N HRAS caused a 3 fold increase; all of which were significant. However, surprisingly, the protein levels of HOP in both HeLa and Hs578T cells transfected with HRAS showed no significant change relative to that of cells with wildtype RAS (Fig. 15C,D). Thus HRAS caused an up-regulation the promoter activity of Hs578T cells but not the HOP mRNA and protein expression levels.

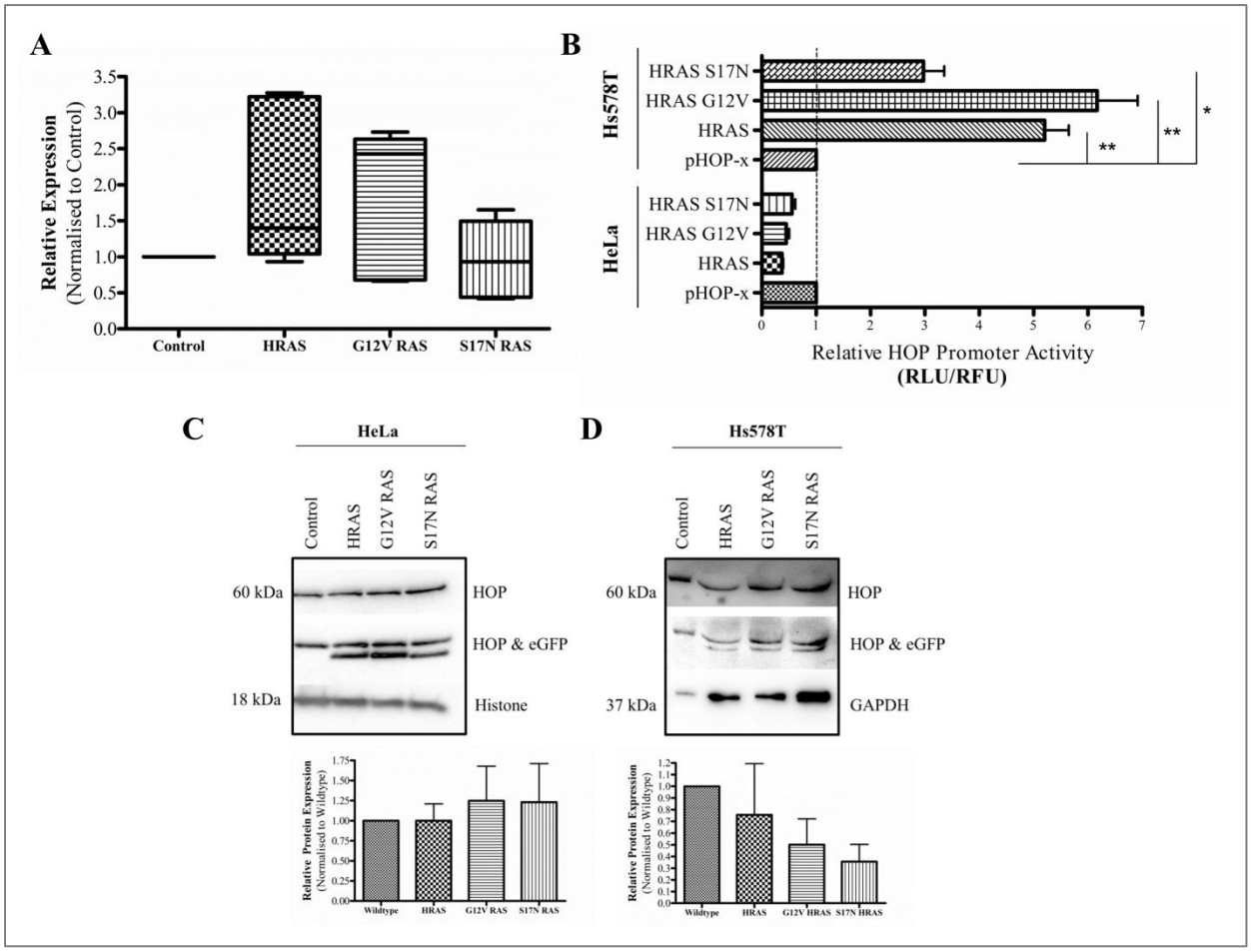


Figure 15: Mutated RAS increased the levels of HOP.

The effect of mutated RAS on HOP mRNA levels was analysed (A). MCF-7, Hs578T and HeLa cells were transfected with wildtype HRAS, constituently active HRAS (G12V) or dominant negative HRAS (S17N). RNA was extracted from MCF-7 cells after a 48 hr incubation and used for real-time reverse-transcription polymerase chain reaction (qRT-PCR). The Ct value of HOP was measured in all samples, and normalised against the Ct value of GAPDH to give a relative quantification of HOP (ΔCt). The ΔCt value for each control sample was subtracted from the ΔCt value of the corresponding inhibitor sample to determine the ΔΔCt value, the normalised expression. The fold difference of HOP in each sample was then determined from the $2^{-\Delta\Delta Ct\ value}$. The results are represented in a box and whiskers plot as the mean values of three independent experiments with duplicates for each sample. B) The effect of mutated RAS on HOP promoter activity was analysed. HeLa and Hs578T cells were co-transfected with the RAS plasmids and pHOP-x. The luciferase activity of the pHOP-x was measured after 48hrs. Luminescence was measured in relative light units (RLU). The luciferase activity was determined by normalising the firefly luciferase activity against the fluorescence produced by the eGFP gene on each of the RAS plasmids. The results from the HOP promoter analysis were then normalised against the results of pHOP-x and HRAS. The results are represented as the mean values of three independent biological replicates with three technical replicates for each. The effect of mutated RAS on protein levels was analysed. HeLa cells (C) and Hs578T cells (D) were transfected with the RAS plasmids. Cell lysates were produced after 48hrs for each sample and used to perform Western Blot analyses (using 100 µg protein). RAS was detected using GFP. Densitometry analysis was performed for each sample using the ImageJ software. For each sample the expression of HOP was normalised against a GAPDH loading control, each inhibitor sample was then normalised against the corresponding control sample (HeLa) or the HRAS alone (Hs578T) to determine the fold difference of HOP expression. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (*P < 0.05, **P<0.01).

We had observed mutated RAS was having an effect on the HOP promoter but we were yet to elucidate why this was so. RAS signalling is mediated by numerous downstream signalling pathways, namely the RAF/MEK/ERK (or MAPK) pathway, the PI3-kinase/AKT/mTOR pathway, PLC_{ε}, RAC and Ral GTPase pathways (Fig.16) (Stephen *et al.*, 2014). The MAPK signalling cascade was focussed on due to the identification of four putative *cis*-elements in the HOP promoter. The TFs which bind to those *cis*-elements were downstream targets of the MAPK signalling pathway (ETS-1, c-Jun, c-Fos, c-Myc, C/EBP^{β}) (Fig.16) (Oikawa & Yamada, 2003; Roskoski, 2012). Additionally, JNK was also known to activate c-Jun and c-Fos TFs (Davis, 1999). Thus in an attempt to determine which RAS pathways were responsible for regulating the HOP promoter activity, Hs578T cells were treated with one of four inhibitors of the MAPK signalling cascade: *α*-hydroxy Farnesyl Phosphonic Acid (*α*HFPA) (RAS inhibitor), Sorafenib (RAF inhibitor), U0126 (MEK inhibitor) or 3-(2-Aminoethyl)-5-((4-ethoxyphenyl)methylene)-2,4-thiazolidinedione (ERK2i, ERK2 inhibitor) (Fig.16).

*α*HFPA blocks the association of RAS with the plasma membrane thereby inhibiting the processing of RAS and preventing it from being biologically activated (Pompliano *et al.*, 1992; Gibbs *et al.*, 1993). It does this by inhibiting the farnesyl-protein transferase which is responsible for the post-translational modification of RAS by farnesylation (Pompliano *et al.*, 1992; Gibbs *et al.*, 1993). By blocking the signalling of RAS, we would be able to observe whether HOP expression is dependent on the RAS signalling cascade (Fig.16). Sorafenib supresses the activity of RAF thereby inhibiting the phosphorylation of MEK1/2, and subsequent activation of ERK1/2 (Wilhelm *et al.*, 2004; Liu *et al.*, 2006). By supressing the activity of RAF, we would be able to begin to discriminate which RAS signalling pathways were involved in regulating HOP expression (Fig.16). If HOP expression decreased upon RAF inhibition, it would signify that the MAPK cascade was involved in the regulation of HOP expression. If HOP expression did not decrease, it would signify that another RAS signalling pathway was involved instead (such as PI3K). U0126 selectively inhibits the kinase activity of MEK1/2 (Favata *et al.*, 1998). This inhibitor, like RAF inhibition, would allow for the determination of the involvement of the MAPK signalling cascade in the regulation of HOP expression. The ERK2 inhibitor (ERK2i) inhibits the docking domain of ERK2 and thereby prevents ERK2 from interacting with substrates (Hancock *et al.*, 2005). This inhibitor would enable us to discriminate between the roles of ERK1 and ERK2 in the

regulation of HOP expression. If HOP expression decreased it would signify that ERK2 is possibly responsible for the activation of the TFs that could regulate HOP expression (Fig.16)

A WST-1 cell cytotoxicity assay was performed for each inhibitor to determine suitable EC_{50} values for each when working with Hs578T cells (Fig.17). Sorafenib, U0126 and ERK2i decreased the cell viability of Hs578T cells resulting in IC_{50} values of 3.6 µM, 37.7 µM and 60 µM respectively. The data generated from Sorafenib (Fig. 17B) treatment was considered the most accurate due to the high R^2 value of 0.82 and very low standard deviations. Data obtained for U0126 and ERK2i treatments (Fig. 17C,D) had lower R^2 values of 0.65 and 0.77 respectively. U0126 data contained standard deviations of up to 75% while ERK2i data contained standard deviations of up to 50%. This suggested that the accuracy of the data obtained for U0126 and ERK2i treatments was less than that obtained for the Sorafenib treatment. αHFPA, with a negative R^2 value of -10.66, was ineffective as an inhibitor in this cell line (Fig. 17A). The Q61L HRAS mutation found within Hs578T cells alters the binding capacity of HRAS, causing it to bind to GGPTase I instead of FTPase (Pompliano *et al.*, 1992; Gibbs *et al.*, 1993). Since αHFPA is a FTI it would not affect the binding of Q61L HRAS to GGPTase I. αHFPA was thus used as a negative control.

Upon determining the subtoxic concentrations of each inhibitor the effect of these three inhibitors on HOP expression in Hs578T cells was determined through qRT-PCR and western blot analysis (Fig.18). Hs578T cells were treated with either 30 µM αHFPA, 3.6 µM Sorafenib or 37.7 µM U0126. Although no significant differences were observed in the levels of HOP mRNA isolated from inhibited cells versus the controls, both Sorafenib and U0126 decreased the levels of HOP mRNA relative to the control (Fig.18A). Treatment of cells with αHFPA resulted in varied levels of HOP mRNA. Sorafenib also reduced the levels of HOP protein by half that of the control (Fig. 18B). Results from the U0126 treatments were not reproducible, but the average of the HOP mRNA observed was greater than the control expression.

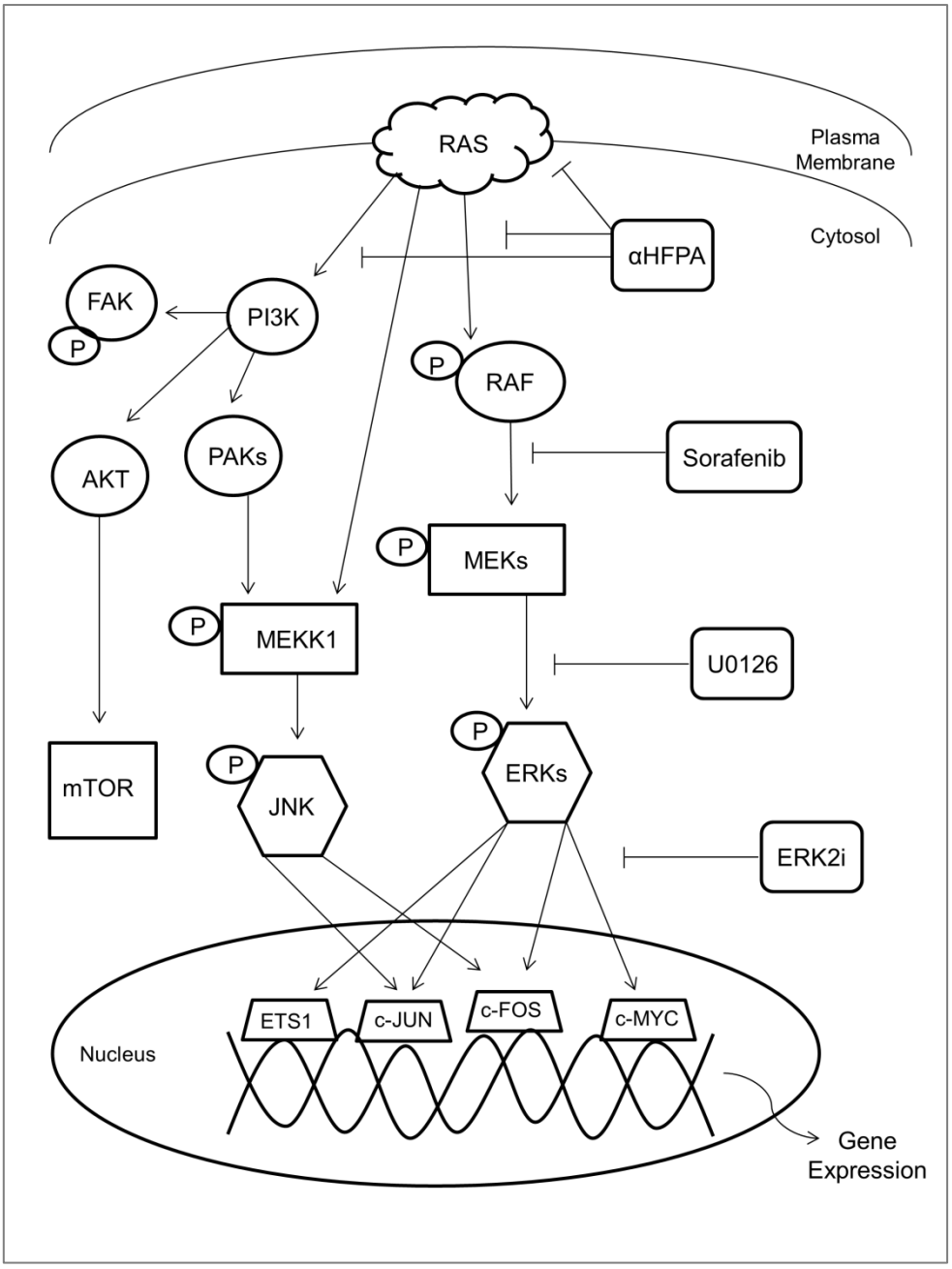


Figure 16: Inhibition of the RAS signalling pathways.

Schematic diagram illustrating the positions of inhibition of the four inhibitors along the mitogen-activated protein kinases (MAPK) / extracellular signal-regulated kinases (ERK 1 and ERK 2) signalling pathway and the phosphoinositide 3-kinase (PI3K) signalling pathway. *α*-hydroxy Farnesyl Phosphonic Acid (*α* HFPA) is a RAS inhibitor, Sorafenib is a RAF inhibitor, U0126 a MEK inhibitor and 3-(2-Aminoethyl)-5-((4-et hoxyphenyl)methylene)-2,4-thiazolidinedione (ERK2i) is an ERK2 inhibitor.

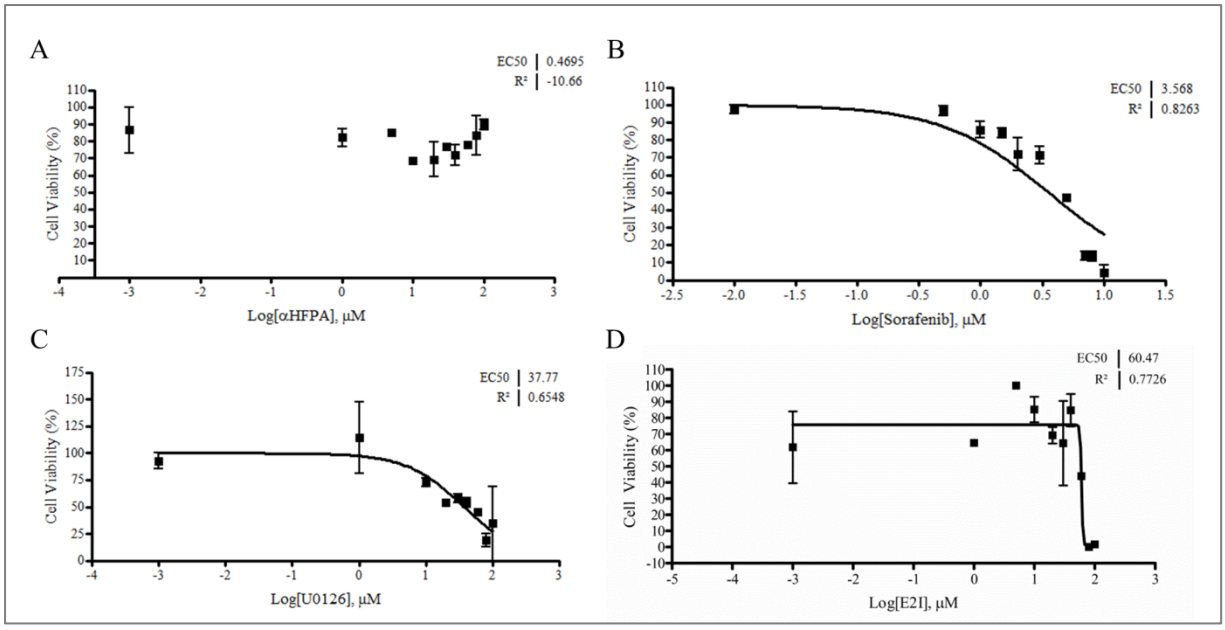


Figure 17: Chemosensitivity of Hs578T cells to RAS signalling inhibitors.

The cytotoxicity of RAS signalling inhibitors on the mammalian breast epithelial cell line Hs578T was tested using the cell proliferation reagent WST-1 and an incubation period of 96 hrs. A) *α*-hydroxy Farnesyl Phosphonic (αHFPA) Acid, with a concentration range of 0.001 µM to 100 µM, B) Sorafenib with a concentration range of 0.001 µM to 10 µM, C) U0126 with a concentration range of 0.002 to 100 µM. D) 3-(2-Aminoethyl)-5-((4-ethoxyphenyl)methylene)-2,4-thiazolidinedione (ERK2i) with a concentration range of 0.01 to 100 µM.

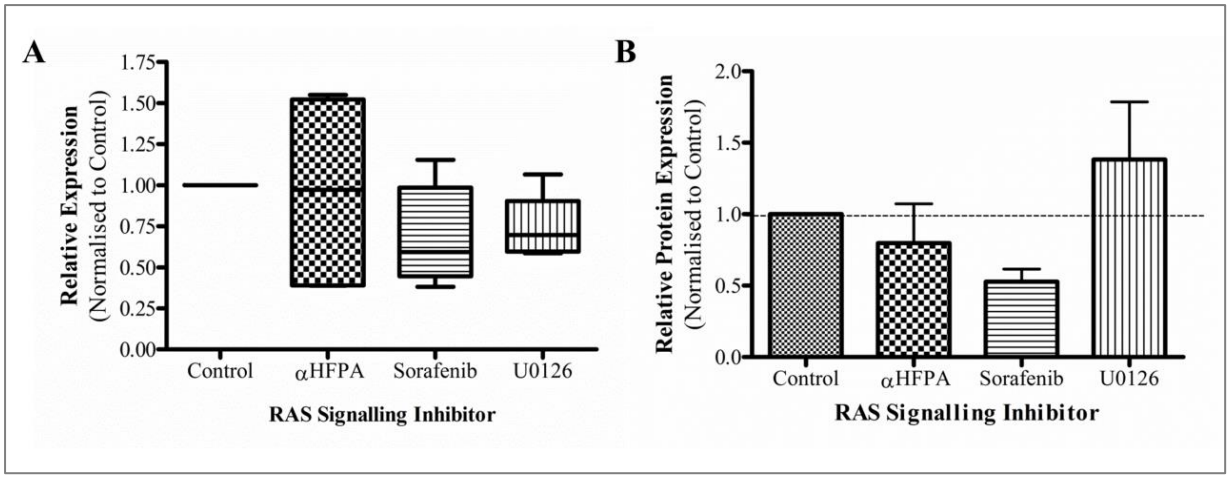


Figure 18: Inhibition of RAS signalling effects the expression of HOP.

Hs578T cells were treated with either 30 µM αHFPA, 3.6 µM Sorafenib, 37.7 µM U0126 or the cooresponding dimethyl sulfoxide (DMSO) or Ethanol vehicle controls and incubated for 48 hrs. A) RNA was extracted from Hs578T cells and used for real-time reverse-transcription polymerase chain reaction (qRT-PCR). The Ct value of HOP was measured in all samples, and normalised against the Ct value of GAPDH to give a relative quantification of HOP (ΔCt). The ΔCt value for each control sample was subtracted from the ΔCt value of the corresponding inhibitor sample to determine the ΔΔCt value, the normalised expression. The fold difference of HOP in each sample was then determined from the $2^{-\Delta\Delta Ct}$ value. B) Cell lysates were produced for each sample and used to perform Western Blot analyses (using 100 µg protein). Densitometry analysis was performed for each sample using the ImageJ software. For each sample the expression of HOP was normalised against a histone control, each inhibitor sample was then normalised against the corresponding control sample to determine the fold difference of HOP expression. The results are represented as the mean values of three independent experiments with duplicates for each sample. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance ($P < 0.05$).

4.2.3 - Consequences of the RAF signalling on HOP production

Since the RAF inhibitor, Sorafenib, was the only drug that caused a substantial decrease in the levels of HOP protein, we decided to focus on the effect of RAF signalling on the expression of HOP. To determine the effect of the downstream signalling effectors of RAF upon RAF inhibition, Hs578T breast cancer cells were treated with the RAF inhibitor Sorafenib (Fig. 16). The phosphorylation of the signalling molecules within the downstream cascades was used to indicate their activation, and thus the loss of phosphorylation would indicate reduced activation of the proteins. The change in the levels of phosphorylation of RAF activated signalling molecules was determined by Western blot analyses for pFAK, pERK1/2 and pJNK (Fig. 19A), and by changes in subcellular localisation for pFAK and ppERK1/2 using confocal microscopy (Fig. 19B). Our antibody for pJNK was not compatible with fluorescent staining and we were unable to detect the levels of p38 (data not shown).

The phosphorylation of RAF downstream signalling molecules was deactivated upon treatment with Sorafenib. The pERK1/2 antibody detected p44 (ERK1) and p42 (ERK2) MAP kinases when phosphorylated. For FAK, the 46 kDa band of JNK, and ERK1/2 there was a decrease in phosphorylation compared with their control counterparts upon RAF inhibition (Fig. 19A). Phosphorylated ERK2 was detected in the control sample in greater quantities than phosphorylated ERK1, identified by the band at molecular weight 42 kDa with a greater intensity than the intensity of the band observed at molecular weight 44 kDa (Marais *et al.*, 1993; Kortenjann *et al.*, 1994). The reduction in phosphorylated ERK2 upon Sorafenib treatment was greater than that observed for phosphorylated ERK1. The pAKT and RhoC signals were not decreased by the inhibition of RAF. Fluorescent microscopy was utilised to visualise individual cells. Both pFAK and pERK1/2 were observed to be mainly within the nuclei of the cells (Fig. 19B). Upon Sorafenib treatment, there was a gradual loss of the pERK1/2 protein within the nuclei as the proteins dispersed into the cytoplasm. This trend continued at 24 hrs for pERK1/2. A total loss of pFAK was observed within the cells by 24hrs of Sorafenib treatment.

Since Sorafenib inhibited the phosphorylation of various MAPK signalling effectors, we next looked at the effect of Sorafenib on HOP protein levels. The levels of HOP were also decreased upon treatment with the RAF inhibitor (Fig. 19C), as identified previously in Figure 18. A time course study of the effect of Sorafenib on HOP protein levels was performed (Fig. 19D). Hs578T cells were treated with Sorafenib over varying time periods and the consequent levels of HOP protein detected using Western Blot analyses. The densitometry for the triplicates is shown. The levels of HOP within Hs578T cells treated with Sorafenib decreased at a 48 hr incubation period with the inhibitor, although the cells expressed an increased level of HOP relative to the untreated cells at time points 12, 24 and 36 hrs (Fig. 19D). Between time points 36 and 48 hrs the levels of HOP dropped by half relative to the untreated cells. Lastly, we considered inhibiting the last effector within the MAPK signalling pathway, ERK2 (Fig. 16). Hs578T cells were thus treated with 60 μM ERK2i and the consequent protein levels of HOP determined (Fig. 19E). Western Blot analyses revealed a significant reduction in the protein levels of HOP in Hs578T cells treated with ERK2i compared to those without the inhibitor treatment.

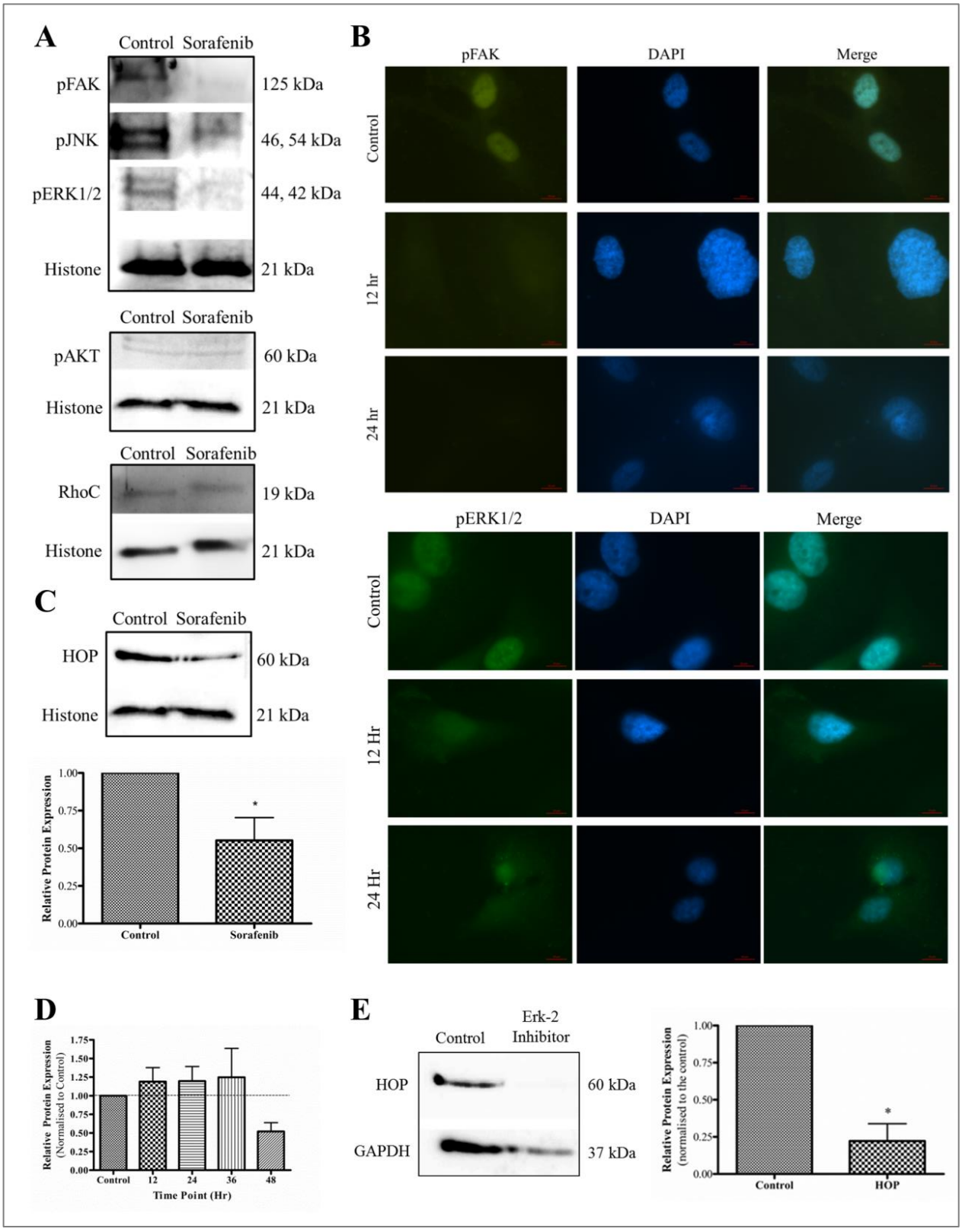


Figure 19: HOP is down-regulated by inhibition of the MAPK pathway.

Hs578T cells were treated with 3.6 µM Sorafenib or dimethyl sulfoxide (DMSO) acting as the vehicle control. Cell lysates were produced for each sample after a 48 hr incbation at 37 ^{o}C and used to perform Western Blot analyses (using 100 µg protein) (A). Various signalling molecules within the RAS signalling pathway were probed: phosphorylated Akt (pAKT), the signalling G protein RhoC, phosphorylated focal adhesion kinase (pFAK), phosphorylated c-Jun N-terminal kinase (pJNK), and phosphorylated mitogen-activated protein (MAP)

kinases ERK1 and ERK2 (pERK1/2). Histone H3 and GAPDH were used as a loading control for total protein. The images represent experiments done in triplicate. B) The subcellular localisation of pFAK and ppMAPK (pERK1/pERK2) in Hs578T cells was determined using indirect immunofluorescence staining and fluorescence microscopy. Hs578T cells were seeded overnight onto glass coverslips and treated with 3.6 μM Sorafenib or dimethyl sulfoxide (DMSO) for 12 or 24 hours. Cells were incubated with pFAK or ppMAPK primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss AxioVert AI FL-LED inverted fluorescence microscope with a 100x objective and were analysed by Zen Lite 2012. All samples were captured with equivalent laser settings. The scale bars indicate a length of 10 μm. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown. C) The cell lysates produced from Hs578T cells treated with 3.6 μM Sorafenib or dimethyl sulfoxide (DMSO) were used in Western blot analyses and HSP70/HSP90 organising protein (HOP) was probed. Densitometry analysis was performed for each sample using the ImageJ software. For each sample the expression of HOP was normalised against a Histone loading control, each inhibitor treated sample was then normalised against the corresponding DMSO control sample to determine the fold difference of HOP expression. The results were represented as the mean values of three independent experiments. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (*$P < 0.05$). D) Time course study on the effect of Sorafenib on the levels of HOP protein. Hs578T cells were treated with 3.6 μM Sorafenib or dimethyl sulfoxide (DMSO). Cell lysates were produced for each sample after increasing incubation times at 37 ^{o}C and used to perform Western blot analyses (using 100 μg protein) and densitometry. GAPDH was used as a loading control for total protein. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance ($P < 0.05$). E) Hs578T cells were treated with 60 μM ERK2i or dimethyl sulfoxide (DMSO). Cell lysates were produced for each sample after a 48 hr incbation at 37 ^{o}C and used to perform Western blot analyses (using 100 μg protein). GAPDH was used as a loading control for total protein. Densitometry analysis was performed for each sample. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (*$P < 0.02$).

*For an electronic copy of the fluorescence images see Electronic Appendix B.

4.2.4 – ETS-1 and C/EBP$^{\beta}$ *cis*-elements are linked to HOP regulation

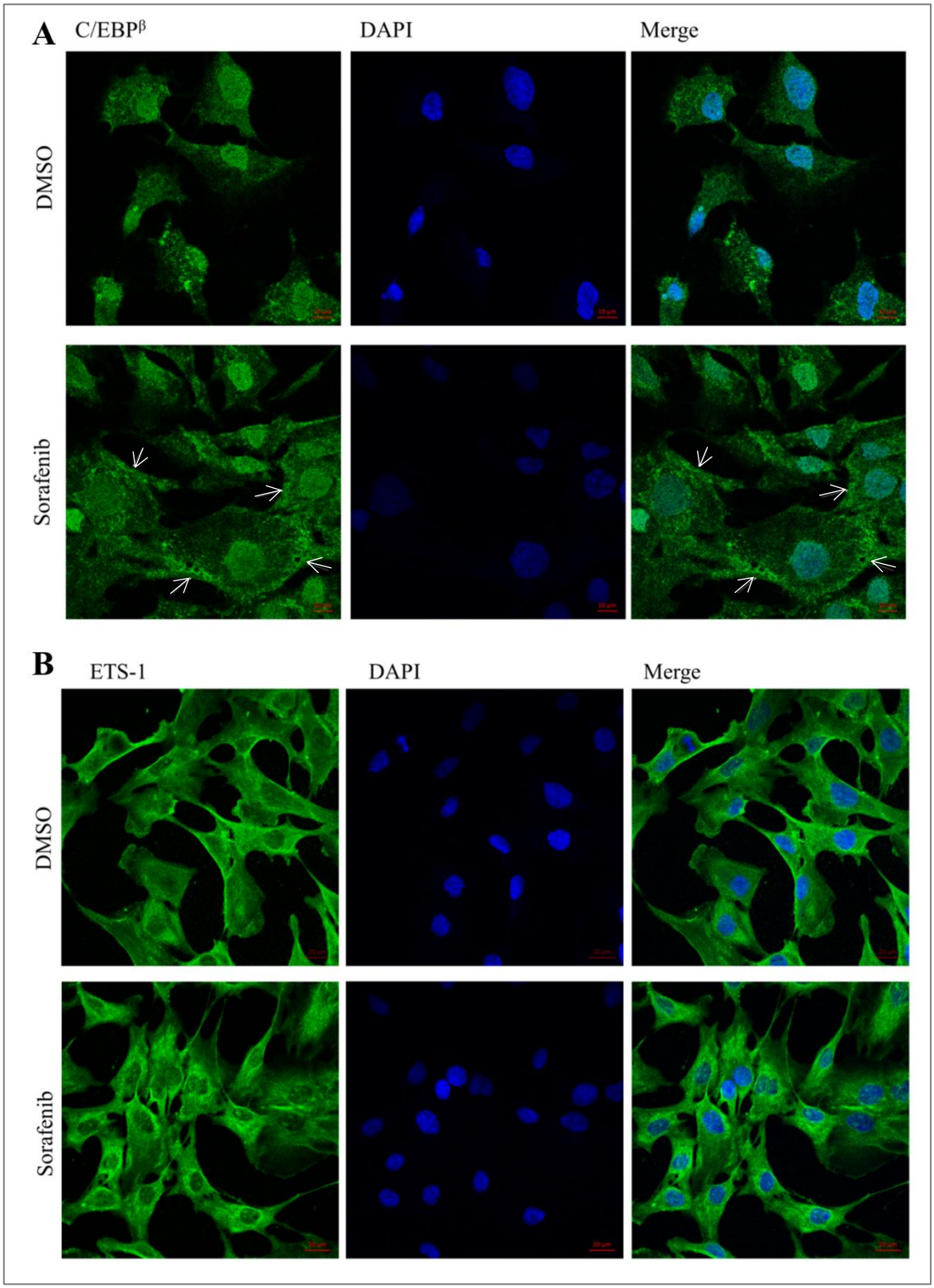
Next, to determine which downstream targets of the MAPK signalling pathway may have been involved in the regulation of HOP expression, the levels and localisation of two key TFs were analysed upon Sorafenib treatment. *Cis*-elements identified in the *in silico* HOP promoter analysis that were targets of TFs involved in the MAPK signalling pathway were ETS-1 and C/EBP$^{\beta}$ (Fig 16). Confocal microscopy was utilised to visualise the protein levels and subcellular localisation of C/EBP$^{\beta}$ and ETS-1 within Hs578T cells with and without Sorafenib treatment. The change in the protein levels of the TFs was determined by Western Blot analyses (Fig.20).

Under normal conditions, C/EBP$^{\beta}$ was found with a punctate staining dispersed throughout the cell but with a strong presence in the nucleus. Upon Sorafenib treatment the protein was identified specifically in clusters along the periphery of the cell (Fig. 20A, clusters identified by white arrows), still maintaining the punctate staining pattern, with a reduction in protein levels observed in the nuclei. To confirm the difference observed in fluorescence intensity in the treated versus the control cells, an analysis of the fluorescent intensity (mean grey values) of the C/EBP$^{\beta}$ fluorescence in the cytoplasm compared to the nuclei was performed for 10

cells from each treatment. The overall fluorescence intensity decreased in the Sorafenib treated cells, with a statistically significant reduction of fluorescence intensity within the nuclei of Sorafenib treated cells (Fig. 20C).

ETS-1 was also found dispersed throughout the cell under control conditions, with a strong presence in narrow regions along the cell periphery (Fig. 20B). Upon Sorafenib treatment, the peripheral staining was lost and a pronounced reduction of protein was observed within the nucleus. Overall staining was more punctate than under normal conditions, especially within the nucleus. Additionally, the protein was observed in particularly high levels immediately surrounding the nucleus. An analysis of the fluorescence intensity identified a reduction in protein levels in the cytoplasm upon Sorafenib treatment and the ratio of ETS-1 within the cytoplasm versus the nuclei was reduced (Fig. 20D). This suggested that ETS-1 protein moved from the cytoplasm into the nucleus. Additionally, both C/EBP$^{\beta}$ and ETS-1 total protein levels were reduced upon Sorafenib treatment (Fig. 20E).

The last effectors within the MAPK signalling pathway to be considered was ERK1/2 (Fig. 16). ERK2 is the effector that directly activates ETS-1 and C/EBP$^{\beta}$ TFs via phosphorylation (Yang *et al.*, 1996; Hanlon *et al.*, 2001). Hence, ERK2 was inhibited in order to discriminate between the role of ERK1 and ERK2 in HOP expression. It may also prevent cross-talk between the MAPK signalling cascade with other effectors that activate ERK2. Hs578T cells were treated with ERK2i and confocal microscopy was utilised to determine the subcellular localisation of ETS-1 upon ERK2 inhibiton compared to normal conditions (Fig.21). ETS-1 was found to be dispersed almost evenly throughout the cell under both control and treated conditions (Fig. 21A). However, a significant reduction in the ETS-1 fluorescence levels were observed upon ERK2i treatment (Fig. 21B). The ratio of ETS-1 fluorescence detected within the cytoplasm compared to within the nuclei of cells was observed to stay the same between control and ERK2i treatments.


A
C/EBPβ
DAPI
Merge
DMSO
Sorafenib
B
ETS-1
DAPI
Merge
DMSO
Sorafenib

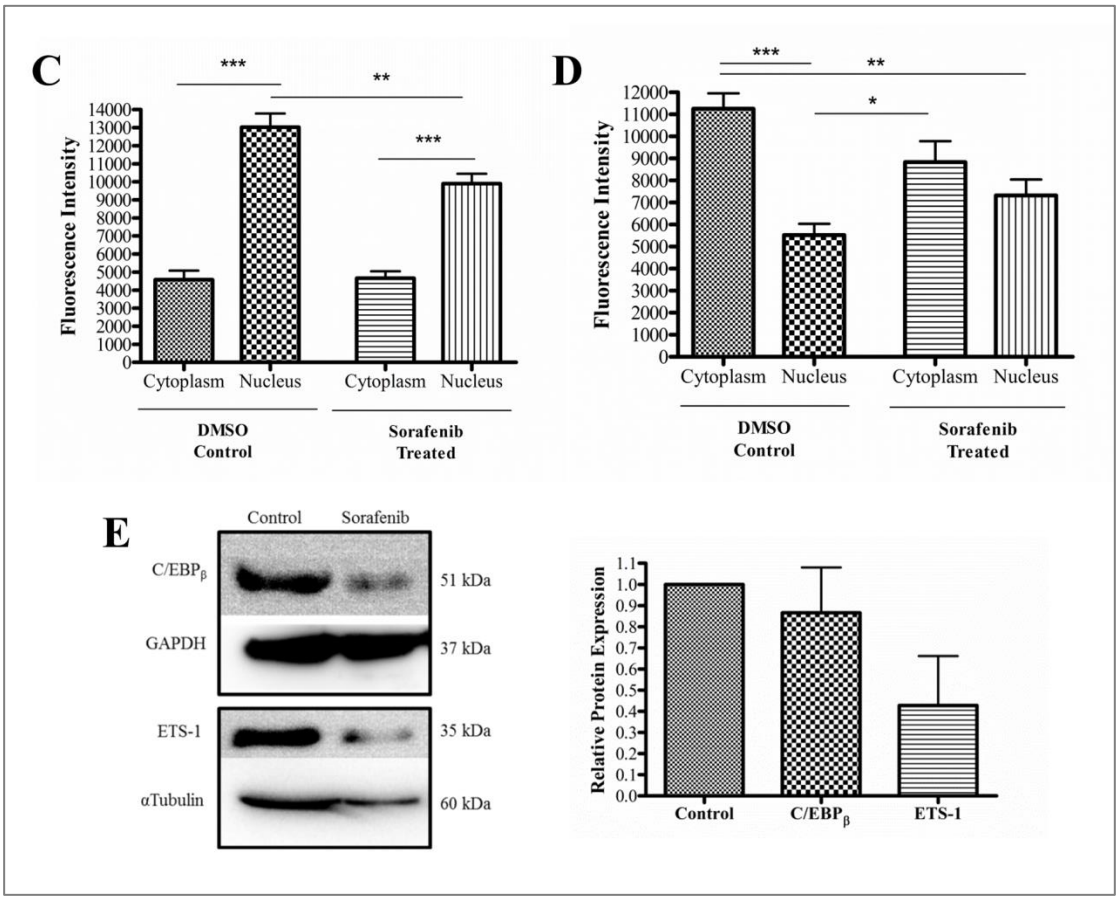


Figure 20: C/EBP$^{\beta}$ and ETS-1 protein levels are down-regulated by RAF Inhibition.

The subcellular localisation of (A) C/EBP$^{\beta}$ and (B) ETS-1 in Hs578T cells was determined using indirect immunofluorescence staining and confocal microscopy. Hs578T cells were seeded overnight onto glass coverslips and treated with 3.6 µM Sorafenib or dimethyl sulfoxide (DMSO) for 24 hours. Cells were incubated with C/EBP$^{\beta}$ or ETS-1 primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss LSM 780 Confocal Microscope with 60x and 100x objectives and were analysed by Zen Lite 2012. All samples were captured with equivalent laser settings. The scale bars represent a length of 10 µm or 20 µm as indicated. White arrows indicate C/EBP$^{\beta}$ staining in peripheral clusters within the Hs578T cells. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown. The fluorescence intensity of (C) C/EBP$^{\beta}$ and (D) ETS-1 observed within the cytoplasm and nuclei of the Hs578T cells was determined by comparing the fluorescence intensity (mean grey values) within each and dividing the total by the distance of the subcellular compartment. The results were represented as the mean values of 10 Hs578T cells for each treatment. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (*P<0.5, **P<0.01, ***P<0.001). (E) Cell lysates were produced for each sample after a 48 hr incbation at 37 ^{o}C and used to perform Western Blot analyses (using 100 µg protein), the CCAAT/enhancer-binding protein β (C/EBP$^{\beta}$) and ETS domain transcription factor (Ets-1) were probed. The images represent experiments done in triplicate. Densitometry analysis was performed for each sample. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance.

*For an electronic copy of the confocal images see Electronic Appendix B.

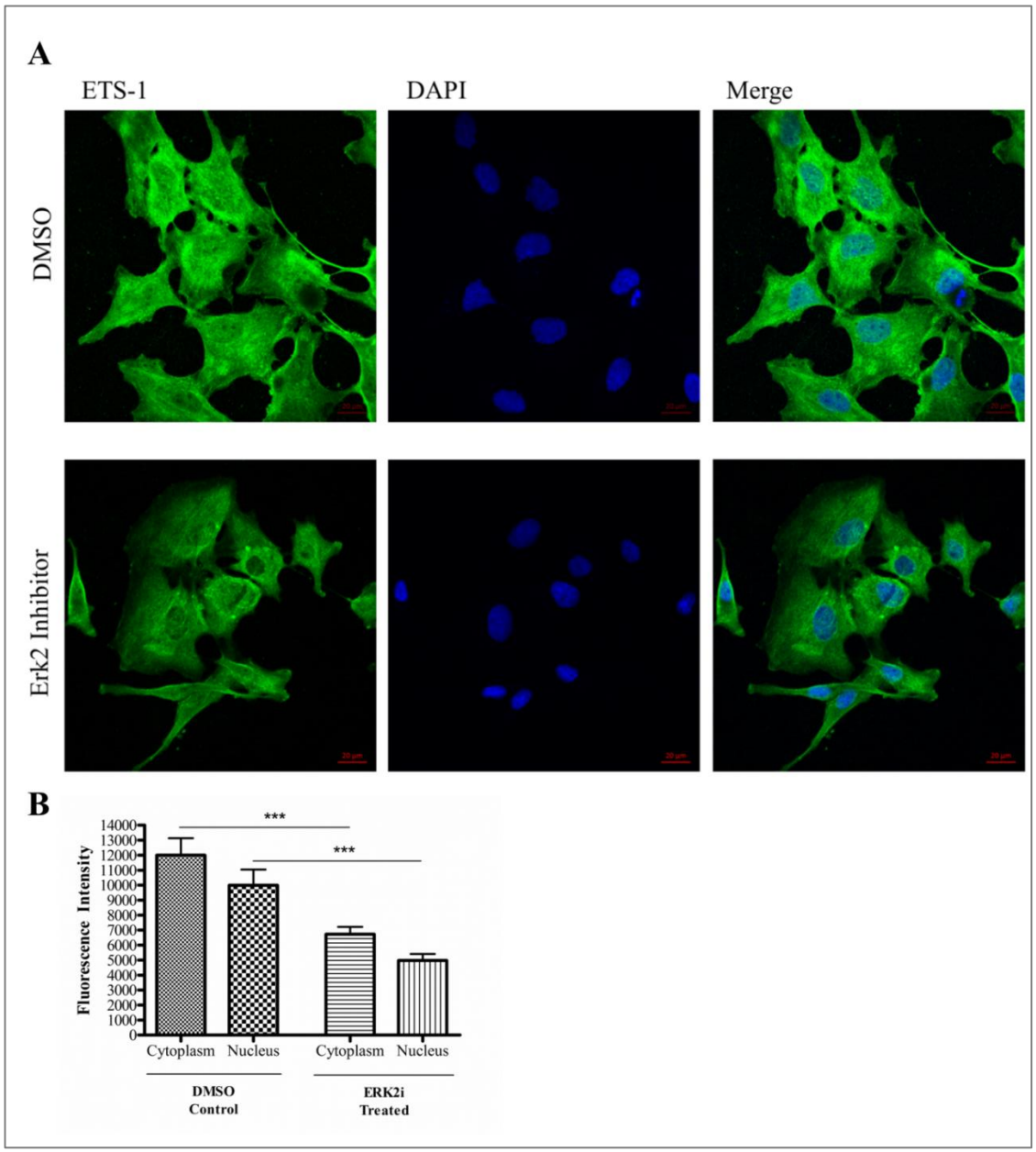


Figure 21: ETS-1 protein levels are down-regulated by ERK2i.

The subcellular localisation of ETS-1 in Hs578T cells was determined using indirect immunofluorescence staining and confocal microscopy (A). Hs578T cells were seeded overnight onto glass coverslips and treated with 60 µM ERK2i or dimethyl sulfoxide (DMSO) for 24 hours. Cells were incubated with ETS-1 primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss LSM 780 Confocal Microscope with a 100x objective and were analysed by Zen Lite 2012. All samples were captured with equivalent laser settings. The scale bars indicate a length of 20 µm. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown. (B) The fluorescence intensity of ETS-1 observed within the cytoplasm and nuclei of the Hs578T cells was determined by comparing the mean grey values within each and dividing the total by the distance of the subcellular compartment. The results were represented as the mean values of 10 Hs578T cells for each treatment. The bars represent the standard error. One way ANOVA and Tukey's Multiple Comparison Tests were performed to determine statistical significance (***$P<0.001$).

*For an electronic copy of the confocal images see Electronic Appendix B.

4.3 Discussion

The data presented in this study provided the *in vitro* evidence that the 1.5 kb region upstream of the HOP gene TSS had promoter activity. Our inhibition of RAS signalling did reduce the levels of HOP protein observed specifically through the MAPK signalling cascade, however we did not conclusively demonstrate that this was due to changes of the promoter activity of HOP. Additionally, the reduction of HOP levels observed upon the inhibition of the MAPK signalling pathway may be independent of the HOP promoter activity and may have been due to the degradation of HOP protein.

4.3.1 – The HOP promoter is differentially activated

The luciferase assays performed confirmed that the promoter region of HOP predicted *in silico* was able to drive transcription of the luciferase coding region *in vitro.* This confirmed that the -1520 to +16 bp region of the human HOP gene was in fact the HOP promoter. To determine which region within the promoter is the core promoter, sequential truncations of the HOP-x sequence would need to be produced and ligated back into the pGL4 vector. Analysis of whether these truncated sequences of the HOP promoter could drive transcription of the luciferase coding region, would ascertain which region is responsible for the core promoter activity. The promoter activity of HOP in the non-cancerous cell line MEF-1 was minimal, whereas the HOP promoter activity within the cancer cell lines was much greater. This suggested that just as protein levels of HOP have been found to be up-regulated in numerous cancers (Kubota *et al.*, 2010; Walsh *et al.*, 2011; Willmer, 2011), so too is the HOP promoter activity up-regulated in cancer cell lines. Furthermore this suggested that within cancer cells the HOP promoter may be in a constant state of activation. This activation may be due to epigenetic factors, cellular stresses or deregulated signal cascades. To confirm this conclusion one would need to determine the HOP promoter activity levels in normal human cell lines.

Alternatively, the lack of HOP promoter activity in MEF-1 cells may be due to the fact that these cells are murine and our promoter sequence is from the human HOP gene. The HOP promoter in murine cells may be regulated differently to the HOP promoter within human cells. It was observed in the bioinformatics comparison of the upstream regions of the HOP gene in the orthologues that the murine and human upstream regions were different, with low sequence identity. The human promoter sequence was regulated by a TATA-less, broad

promoter while the murine HOP promoter contained a TATA box *cis*-element and thus was most likely a TATA regulated promoter. Additionally, the sequence identity between the promoters was 67% suggesting that the promoters would be regulated by different *cis*-elements. This would lead one to question the viability of using mouse models to study the mechanism of HOP functioning with the purpose of extending the scope of the analysis to humans. However, it did provide a useful control for the luciferase assay study to demonstrate that promoter activity from this construct was context dependent.

There may be numerous possible reasons as to why the HOP promoter activity was higher in the Hs578T cell line compared to the HeLa cell line. Genes are known to be differentially transcribed according to environmental conditions, in differing tissue types, in differing sexes and according to each stage of the life cycle (Wray *et al.*, 2003). Additionally, the activation of TFs is context dependent, thus the cellular environment within each cell type may be sufficiently different to provide differing TF ensembles within each the nuclei of the two tissue types (Wray *et al.*, 2003). This would support our observation of differing tissue types (HeLa cells being cervical and Hs578T cells being breast) having different HOP promoter activity levels. Of particular interest in this study was that the Hs578T cell line contained a mutated RAS gene whereas the HeLa cell line had no known RAS mutation. The Hs578T cell line contains the Q61L HRAS mutation that is prevalent in tumours; the mutation renders HRAS in a constitutively active state (Eckert *et al.*, 2004). Furthermore, oncogenic and wildtype RAS has been found to have differential roles in signal transduction (Young *et al.*, 2013). And various genes related to the MAPK signalling pathway have been identified as up-regulated in triple negative breast cancers (TNBC) (Creighton *et al.*, 2006; Hoeflich *et al.*, 2009; Mirzoeva *et al.*, 2009), similar to the Hs578T cell line.

4.3.2 – Mutated RAS activates the HOP promoter but did not significantly alter the mRNA and protein levels

Introducing HRAS into cells that otherwise had no known RAS mutations had no effect on the HOP promoter activity or on the production of HOP protein in cancer cells. However, upon introducing HRAS into a cell line that already had a mutated version of RAS (Hs578T), the activity of the HOP promoter increased significantly. A study by Petanidis *et. al*. (2013) revealed that upon introducing HRAS into the MCF-7 cell line there were no significant changes in cell viability or cell proliferation, merely an increase in the expression of the HRAS protein. A study by Omerovic *et. al.* (2008) showed that HeLa cells (transfected with

HRAS G12V) did not activate the expressed HRAS protein until stimulated with human growth factors (specifically hepatocyte growth factor [HGF] or epidermal growth factor [EGF]). It has been proposed once before that a definable amount (as yet unquantified) of activated RAS is required before RAS-dependent signalling processes can occur (Hamilton & Wolfman, 1998). Such observations from Petanidis *et. al*. (2013), Omerovic *et. al.* (2008) and Hamilton and Wolfman (1998) along with the data presented here suggest that introducing a RAS mutation into cell lines that do not already have a mutated form of RAS is not enough on its own to produce the effects of an active oncogenic RAS protein. Thus the HeLa cells containing the transfected HRAS plasmids were not able to activate the HRAS proteins in a sufficient quantity to see the effect of the oncogenic RAS, due to a lack of the extracellular signals required for induction. Whereas, introducing an additional mutated form of RAS (HRAS G12V) into a cell line that already contained a different RAS mutation (HRAS Q61L), seemed to amplify the effect of the oncogenic RAS within the cells. Since the HRAS Q61L mutant is in a constitutively active state, it does not require the RAS activators that accumulate within the cell to activate it. Thus, the introduction of further wildtype HRAS into the cells may have allowed for the utilisation of the RAS activators and caused a further increase in MAPK signalling.

The HRAS S17N gene expressed a dominant negative protein, which in theory should have blocked HRAS functioning. There are two schools of thought on the mechansim by which the HRAS S17N mutant inhibits RAS signalling. The first is that the mutant sequesters the RAS activators and GEFs within the cell and thereby blocks the activation of endogenous RAS. The alternative argument is that the S17N mutant has a low affinity for GTP in addition to the inability of GTP to activate the S17N mutant, thereby preventing the mutant from binding and activating downstream effectors (Farnsworth & Feig, 1991; Stacey *et al.*, 1991; Nassar *et al.*, 2010). However, since the Hs578T cell line already contained a constitutively active mutant, HRAS Q61L, the inhibitory effect of the HRAS S17N mutant could have been undermined to a certain extent. This does not explain how HRAS S17N increased the HOP promoter activity in the presence of the HRAS Q61L mutant. Although the effect of the HRAS S17N on HOP promoter activity decreased by approximately half of that of HRAS (wildtype and G12V), it still caused a 3 fold increase in promoter activity. This suggested that, specifically within the Hs578T cell line, RAS signalling cascades may have been activated via an alternative mechanism that was not hindered by the HRAS S17N mutation. A possible pathway could be the Protein Kinase C (PKC) –RAF-1 – ERK signalling pathway.

Active PKC has been shown to activate RAS, as well as activate RAF-1 in a RAS independent manner (Sözeri *et al.*, 1992; Cacace *et al.*, 1996; Hirai, 1996; Kawakami *et al.*, 2003). PKC mediated activation of RAF via phosphorylation occurs in the presence of human growth factors (HGF and EGF), which are also known to cause the activation of RAS (Corbit *et al.*, 2003).

Elevated promoter activity does not necessarily mean there will be an elevated rate of translation (Hill & Treisman, 1995; Kapeli & Hurlin, 2011). In this study, the 6 fold increase in the HOP promoter activity within Hs578T cells, in the presence of the HRAS mutation, did not correlate to a subtantial increase in HOP protein levels. Hence the protein levels observed may be due to slow rates of protein translation. From our data, we could then conclude that RAS was influencing the exogenous HOP promoter activity but it did not influence the levels of endogenous HOP mRNA and protein synthesis, despite the fact that inhibiting signalling pathways downstream from RAS reduced the levels of endogenous HOP protein. Could this be due to the saturation of the transcriptional and/or translational machinery in the cells containing HRAS Q61L (Studier & Moffatt, 1986)? Constitutively active HRAS would be causing the up-regulation of numerous promoters throughout the cell, however there is a limited supply of transcriptional and translational machinery available. Or the rate limiting step may be the posttranslational modifications required for mature protein synthesis.

To test this proposition we inhibited RAS signalling instead of inducing it in the cells that already contained constitutively active RAS. However the RAS inhibitor we had available, αHFPA, was not toxic to the Hs578T cells. The Q61L mutation involves the last amino acid in the CAAX peptide within the HRAS sequence. The CAAX residues are the substrate for farnesylation which facilitates the membrane association of RAS. The C represents a cysteine residue while the A represents aliphatic residues and the X represents any amino acid. The substitution of the X has great consequences on what type of transferase the peptide will bind to. By substituting the glutamine to a leucine, the peptide will bind to GGPTase I rather than FPTase (Pompliano *et al.*, 1992; Gibbs *et al.*, 1993). Since αHFPA was a FTI it would not affect the mutated HRAS (Q61L) found within Hs578T cells and explained the lack of toxicity of the compound in the cell line. Although this explained the lack of toxicity to the cell line, at first, it did not explain the slight reduction in the mRNA and proteins levels of HOP observed when Hs578T cells were treated with αHFPA. It is important to note that RAS is not the only protein that is a substrate for FPTases. Therefore, while αHFPA may have had

no effect on HRAS within the cells, it would still have inhibited the farnesylation of numerous other prenylated substrates within the cell (Baines *et al.*, 2011). Most prenylated proteins are those involved in signal transduction cascades with downstream targets involved in normal cellular functions (Li & Sparano, 2003). This could have had inhibitory consequences for the cell. For example, the Rheb small GTPase is a substrate for FTPase, which requires farnesylation before it can activate the mammalian target of rapamycin (mTOR). mTOR in turn is a downstream signalling molecule within the phosphoinositide 3-kinase (PI3K) pathway (PI3K-AKT-mTOR), the deregulation of which is a common characteristic of cancer (Basso *et al.*, 2005). This is an example of the inhibition of the functioning of just one protein via αHFPA and suggests a possible reason for the observation of the reduction HOP expression. However, these effects could probably be considered minor given that the αHFPA was not toxic to the cells.

Since HRAS could not be directly inhibited using αHFPA in the Hs578T line, inhibitors of the RAF-MEK-ERK mitogen-activated protein kinase (MAPK) cascade were considered an indirect method of RAS inhibition. Inhibitors of two intermediates within the pathway were utilised – RAF1 and MEK1/2. MEK1 and MEK2 are the only known catalytic substrates of RAF1, whereas ERK1 and ERK2 are the only known substrates of MEK1/2 (Baines *et al.*, 2011). However one cannot over simplify this pathway by linearizing it as ERK1/2 is a substrate for several other kinases. This would explain why inhibiting MEK1/2 with U0126 did not give reproducible results, and why Sorafenib treatment didn't provide a complete reduction in the levels of HOP. By suppressing the MAPK signalling pathway one does not wholly block ERK1/2 activation by mutated RAS, additional pathways that phosphorylate ERK1/2 may have compensated for the lack of RAF and MEK1/2 signalling (Baines *et al.*, 2011). The activation of alternative signalling pathways has been observed twice in clincial trials for the RAF inhibitor PLX4032 (Johannessen *et al.*, 2010; Nazarian *et al.*, 2010). This suggested that HOP protein expression could be regulated by more than a single signalling pathway (Favata *et al.*, 1998).

The consequence of Sorafenib treatment on HOP production suggested that the greatest inhibiting effect on HOP protein levels was observed after 36 hrs. However, in order to completely nullify the effect of mutated RAS in the Hs578T cell line, we should have used a combination of inhibitors that would have targeted both the MAPK and PI3K-AKT signalling pathways. Such combination therapies have shown greater effective inhibition of mutated

RAS in human melanoma cells, hepatocellular carcinoma and non-small cell lung cancer (Molhoek *et al.*, 2005; Legrier *et al.*, 2007; Wang *et al.*, 2008; Saini *et al.*, 2013). Due to the dependence of cells on these pathways for survival, one would require to determine effective subtoxic concentrations of each inhibitor when used in combination. Due to the high toxicity of a combinatorial approach one would expect large scale side effects, some of which would not be observed *in vitro* but only *in vivo*. Thus although initial clinical benefit of dual inhibition has been shown (phase I clinical trials), phase II data is yet to be released and will be a better representative of the feasibility of dual inhibition of these pathways *in vivo* (Saini *et al.*, 2013).

4.3.3 - RAF and ERK2 regulate HOP protein levels through the MAPK signalling cascade

RAF inhibition decreased the activation of ERK1 and ERK2 within Hs578T cells but not to the same degree; ERK2 was deactivated to a greater degree than ERK1 (Marais *et al.*, 1993; Kortenjann *et al.*, 1994). ERK2 has been found to have higher protein levels than ERK1 in the Hs578T cancer cell line when compared to that of less aggressive cancer cell lines such as MCF-7 (Marais *et al.*, 1993; Kortenjann *et al.*, 1994; Chrestensen *et al.*, 2007), which correlates to the difference in phosphorylated ERK1/2 levels observed in our experiments. This suggested that ERK1 and ERK2 may have had functional differences within Hs578T cells.

FAK phosphorylation is known to be induced by two different signalling pathways, the one being PI3K dependent and the other ERK pathway dependent (Hunger-Glaser *et al.*, 2004). ERK1/2 phosphorylates FAK on Ser-910 while Tyr-397 is phosphorylated by PI3-kinase (Hunger-Glaser *et al.*, 2003). A direct link between ERK and FAK was observed in our study whereby the detection patterns of the signalling molecules upon the treatment of Sorafenib were similar. It suggested that the phosphorylation of pERK1/2 was deactivated in the presence of Sorafenib, which in turn may have caused the deactivation of the phosphorylation of FAK (Hunger-Glaser *et al.*, 2003, 2004; Sawhney *et al.*, 2006; Zheng *et al.*, 2009). Phosphorylation of AKT did not decrease upon Sorafenib treatment suggesting that the MAPK signalling pathway, and not the P13K-AKT signalling pathway, was linked to the regulation of the HOP levels. A study by Eckert *et al.* (2004) supports this proposition as they found that, particularly within the Hs578T cell line, AKT was activated by a RAS independent mechanism. The Jun amino-terminal kinase, or JNK protein kinases (JNK1,

JNK2, JNK3), are expressed as 46 kDa and 55 kDa proteins, and are all a part of a parallel MAPK signalling cascade (Davis, 1999). Although JNK is not directly activated by RAS-RAF, JNK is activated by phosphorylation via two MAP kinase kinases (MKK4 and MKK7) (Ip & Davis, 1998). MKK4 and MKK7 are in turn activated via phosphorylation by MEKK1 and MLK3. MEKK1 is activated directly by RAS or by PAK1/2 in the PI3K pathway. However, since the AKT signalling did not seem to be affected by RAF inhibition, it suggests that the activation of MEKK1 and/or MLK3, or MKK4 and/or MKK7 may have been inhibited instead.

Phosphorylation of FAK and ERK1/2 represent the activation of these signalling molecules within the cell (Khokhlatchev *et al.*, 1998). Under resting conditions wherein the proteins are inactivated, FAK and ERK1/2 can be observed anchored within the cytoplasm of cells (Shaul & Seger, 2007). Upon activation the majority of the phosphorylated signalling molecules are released from the anchors and translocate to the nuclues where they in turn activate various TFs. The lack of pFAK and the reduction of pERK1/2 observed in the nuclei of Hs578T cells when treated with Sorafenib further indicates the loss of protein phosphorylation and hence the activation of the two signalling molecules (Hamilton & Wolfman, 1998). Thus the analysis revealed that each protein was activated specifically by the MAPK pathway.

The reduction of HOP proteins levels corresponded to the reduction in phosphorylated ERK1/2, FAK and JNK upon RAF inhibition with Sorafenib. This suggested a possible link between the activation of these effectors with the regulation of HOP expression. Furthermore, ERK2 inhibition resulted in a reduction in HOP protein levels. This suggested that it was ERK 2, instead of ERK1, that was involved in the regulation of HOP expression.

4.3.4 - C/EBP$^{\beta}$ and ETS-1 are associated with the transcriptional regulation of HOP

Oncogenic RAS has already been found to cause the activation of C/EBP$^{\beta}$ via phosphorylation in fibroblasts, erythroblasts and p19 embryonic carcinoma cells (Nakajima *et al.*, 1993; Kowenz-Leutz *et al.*, 1994). More specifically, ERK2 has been determined to phosphorylate the transcription factor C/EBP$^{\beta}$ via a T188 phosphorylation site (Hanlon *et al.*, 2001; Zhu *et al.*, 2002). Oncogenic HRAS activation of C/EBP$^{\beta}$ is dependent on T188 phosphorylation and thus the endogenous C/EBP$^{\beta}$ protein has been classified as a downstream mediator of oncogenic HRAS signalling (Zhu *et al.*, 2002). This was interesting due to the large number of putative CCAAT elements identified within the HOP promoter

region. Our results supported the findings of Zhu *et. al*. (2002) in that we observed in the Hs578T cell line that C/EBP$^{\beta}$ protein levels decreased upon RAF inhibition. The consequent reduction of the TF observed in the nucleus suggests that the deactivation of ERK2 by RAF inhibition blocked the phosphorylation of C/EBP$^{\beta}$. Since the reduction of C/EBP$^{\beta}$ corresponded to the reduction of HOP observed upon the inhibition of RAF, this led to the proposal that the MAPK signalling pathway may regulate HOP expression through C/EBP$^{\beta}$ *cis*-elements within the HOP promoter.

ETS-1 is a well characterised TF and was identified as a nuclear target of the RAS signalling pathway in 1996 by Yang *et al*. The function of ETS-1 relies on the phosphorylation of the protein, the consequence of which is the ability of the protein for DNA binding, protein-protein interactions, and transcriptional activation (Seth & Watson, 2005). Phosphorylation of ETS-1 at Thr-38 occurs via ERK2 (Callaway *et al.*, 2010). Being one of the last effectors at the end of the RAS signalling pathways, the functioning of ETS-1 is also altered upon the deregulation of RAS in cancers. An example of such deregulation was identified in breast cancer where ETS-1 was found to be up-regulated within the Hs578T cell line (Behrens *et al.*, 2001). A similar trend of elevated ETS-1 levels has been observed in colon, thyroid, lung and pancreatic cancers; with these elevated levels corresponding to tumour progression (Simpson *et al.*, 1997; Ito *et al.*, 1998, 2002; Nakayama *et al.*, 1999).

ETS-1 interacts with various TFs such as AP-1, PAX-5 and NF-Kβ and in doing so co-regulates the expression of a variety of genes (Li *et al.*, 2000). The putative *cis*-elements for each of those TFs were identified within the HOP promoter. This led to the investigation of the effect of ETS-1 on the expression of HOP in cancer cells. Upon ERK2 inhibition, a significant reduction of ETS-1 was identified within the nuclei of the Hs578T cells. This suggested a loss of phosphorylation of ETS-1 would have occurred as well, which would in turn explain a loss of function for the protein as a TF, the ability of the protein to bind to DNA and to interact with other TFs (Oikawa & Yamada, 2003). The inhibition of ETS-1 in both Sorafenib and ERK2i treated Hs578T cells correlated with the reduction of HOP protein levels in these cells. This might suggest that the RAS pathway regulated HOP expression through the ETS-1 *cis*-elements within the HOP promoter as well and that ETS-1 may be one of the regulating factors causing the observed increase in HOP levels in cancer cells.

However, to truly confirm whether HOP is a target gene for ETS-1 and C/EBP$^{\beta}$ proteins, chromatin immunoprecipitation analysis of the putative *cis*-elements within the HOP promoter would be required to determine whether ETS-1 and C/EBP$^{\beta}$ functionally binds any of those sites (Mukhopadhyay *et al.*, 2008). An attempt was made to perform an initial analysis into the function of ETS-1 *cis*-elements within the HOP promoter using truncations of the HOP promoter reporter construct. The designed truncation excluded two of the predicted ETS-1 *cis*-elements at positions -1102 bp and -1044 bp while including another two predicted ETS-1 *cis*-elements at positions -690 bp and -62 bp (Appendix C). The truncation was to be ligated into the pGL4 vector and transfected into Hs578T cells. Luciferase assays would have utilised to determine the level of promoter activity of the truncated promoter versus the full length promoter. A reduced level of promoter activity in the truncated promoter, compared to that of the full length, would have indicated *in vitro* evidence for the position of the core promoter of HOP, as well as suggested the necessity of the ETS-1 *cis*-elements for promoter activity. Additionally, if the transfected cells had been treated with Sorafenib prior to the luciferase assay, a change or lack thereof in the levels of the promoter activity for the truncated and full length promoters would have confirmed whether the ETS-1 elements were active or not and whether they were regulated by the MAPK signalling pathway. The aforementioned truncation was synthesised and inserted into a pUC57 vector by GenScript (USA). The insert was successfully digested out of the pUC57 vector using *Bgl*II restriction enzymes but the consequent blunt end cloning of the insert into a pGL4 vector was unsuccessful. Of the 25 transformants screened for the correct direction of the insert, only 1 clone contained the insert and it was in the wrong direction (Appendix C). A possible reason for this difficulty could have been due to the high GC content of the truncation. These experiments are ongoing and will be used to support our conclusions on the role of ETS-1 in HOP promoter activity in the future.

4.3.5 – Conclusion

The data presented here indicates that upon the inhibition of the MAPK cascade HOP protein levels are decreased. However, our data does not conclusively suggest that this is due to a decrease in the promoter activity of HOP. An alternative possibility is that the blocking of the MAPK signalling pathway may have been independent of the activity of HOP promoter and may have led to the degradation of HOP protein. To test this hypothesis, cells would be

treated with a MAPK pathway inhibitor (αHFPA, Sorafenib, or U1026). Next these cells would be treated with or without the proteasome inhibitor MG132. Treatment with MG132 would prevent the degradation of any protein synthesised after treatment of the cells (Lee & Goldberg, 1998). Thus, if we observed greater HOP protein levels (or ubiquitin bound HOP) in the cells treated with MG132 compared to those cells without MG132 treatment, it would indicate that inhibition of the MAPK cascade led to the degradation of HOP protein.

In conclusion, our data suggest that in cancer cell lines containing mutated and contitutively active HRAS, the HOP promoter may be regulated by a mechanism involving but not exclusive to the MAPK signalling pathway, potentially through the ETS-1 and C/EBP$^{\beta}$ *cis*-elements within the HOP promoter. Our data further serves as a reminder that RAS signalling cascades are not simply linear, that there is vast crosstalk across the signalling pathways that can both enhance or block RAS-mediated signalling.

Chapter 5

Conclusion

This research has allowed us to decipher a small part of the biological systems governing HOP expression as well as elucidate the importance of HOP in cancer. The putative HOP promoter region was identified *in silico* and *in vitro*. In cancer cell lines containing mutated RAS (Hs578T), HOP was up-regulated via a mechanism involving the MAPK signalling pathway, possibly involving the Ets-1 and C/EBP$^{\beta}$ *cis*-elements, within the HOP promoter (Fig.22). These findings suggest for the first time that HOP expression in cancer may be regulated by RAS activation of the HOP promoter and further advocates HOP as a novel drug target for anticancer therapeutics. At present we assume that these data represent the known HOP. However, depending on the nature of the probe used to identify them and the similarity of the sequence between the regions that are conserved, it is possible that some of the expression analyses may in fact reflect the alternative isoforms of HOP predicted in our *in silico* analysis. Additionally, we provided *in silico* evidence for murine systems being the best available model organism for the *in vivo* study of HOP functioning in human disease, due to the similarity observed between the two in synteny, gene structure and co-expression analyses. A comparison of HOP within model organisms had yet to be presented to the scientific community. From here, both CHIP analyses of the ETS-1 *cis*-elements as well as HOP promoter truncations need to be performed to confirm our presented data. Further luciferase assays should be performed to determine the HOP promoter activity in the presence of Sorfenib, U0126 and ERK2i so as to further elucidate the mechanism by which the MAPK signalling cascade regulates HOP expression. An *in vitro* analysis of the HOP splice variants should be performed to elucidate the existence of HOP isoforms at the protein level. If the existence of such isoforms is confirmed, it would be of importance and interest to characterise those isoforms and determine their function within the cell.

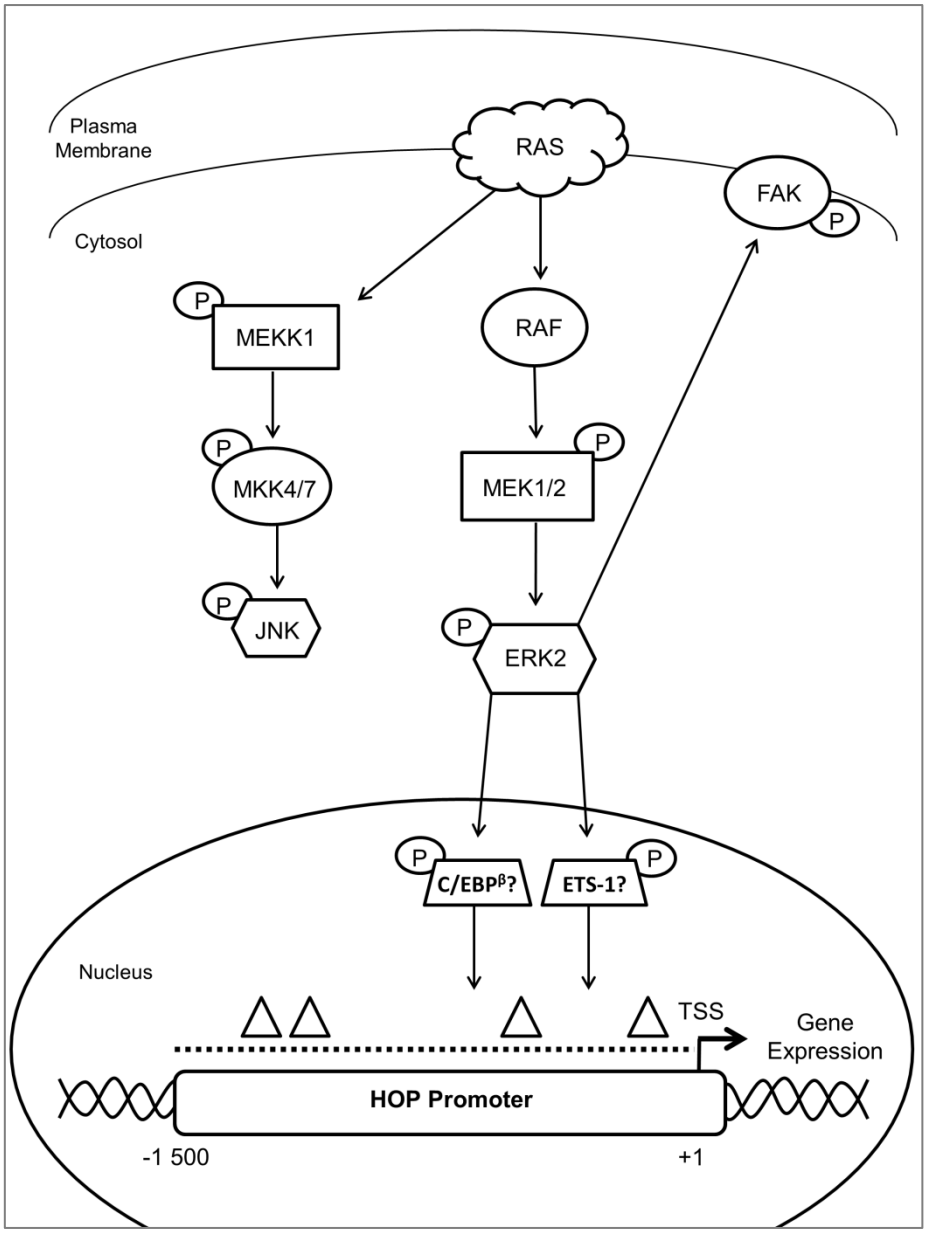


Figure 22: A proposed model for the regulation of HOP in cells containing mutated RAS.

The MAPK signalling pathway activates the transcription factors C/EBP$^{\beta}$ and ETS-1; which in turn possibly regulate the HOP gene expression in cancer cells. Arrows represent the direction of activation. P represents phosphorylation of effectors or transcription factors. TSS represents the transcription start site of the HOP gene. The numbers represent the positions within the HOP promoter relative to the TSS. Triangles represent putative ETS-1 *cis*-elements and the dots represent the putative C/EBP$^{\beta}$ *cis*-elements.

Chapter 6

References

Akerfelt, M., Morimoto, R.I. & Sistonen, L. (2010) Heat shock factors: integrators of cell stress, development and lifespan. *Nature reviews. Molecular cell biology*, 11, 545–55.

Akiyama, Y. (1995) TFSEARCH: Searching Transcription Factor Binding Sites. (Web Program)

Alhopuro, P., Sammalkorpi, H., Niittymäki, I., Biström, M., Raitila, A., Saharinen, J., *et al.* (2012) Candidate driver genes in microsatellite-unstable colorectal cancer. *International journal of cancer. Journal international du cancer*, 130, 1558–66.

Alvira, S., Cuéllar, J., Röhl, A., Yamamoto, S., Itoh, H., Alfonso, C., *et al.* (2014) Structural characterization of the substrate transfer mechanism in Hsp70/Hsp90 folding machinery mediated by Hop. *Nature communications*, 5, 5484.

Andersen, J.B., Spee, B., Blechacz, B.R., Avital, I., Komuta, M., Barbour, A., *et al.* (2012) Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. *Gastroenterology*, 142, 1021–1031.

Anglesio, M.S., Arnold, J.M., George, J., Tinker, A. V, Tothill, R., Waddell, N., *et al.* (2008) Mutation of ERBB2 provides a novel alternative mechanism for the ubiquitous activation of RAS-MAPK in ovarian serous low malignant potential tumors. *Molecular cancer research : MCR*, 6, 1678–90.

Anish, R., Hossain, M.B., Jacobson, R.H. & Takada, S. (2009) Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. *PloS one*, 4, e5103.

Ankeny, R.A. & Leonelli, S. (2011) What's so special about model organisms? *Studies in History and Philosophy of Science Part A*, 42, 313–323.

Arantes, C., Nomizo, R., Lopes, M.H., Hajj, G.N.M., Lima, F.R.S. & Martins, V.R. (2009) Prion protein and its ligand stress inducible protein 1 regulate astrocyte development. *Glia*, 57, 1439–49.

Armengaud, J., Trapp, J., Pible, O., Geffard, O., Chaumot, A. & Hartmann, E.M. (2014) Non-model organisms, a species endangered by proteogenomics. *Journal of proteomics*, 105, 5–18.

Bagatell, R. & Whitesell, L. (2004) Altered Hsp90 function in cancer: a unique therapeutic opportunity. *Molecular cancer therapeutics*, 3, 1021–30.

Baines, A.T., Xu, D. & Der, C.J. (2011) Inhibition of Ras for cancer treatment : the search continues. *Future Medical Chemistry*, 3, 1787–1808.

Baitsch, L., Baumgaertner, P., Devêvre, E., Raghav, S.K., Legat, A., Barba, L., *et al.* (2011) Exhaustion of tumor-specific $CD8^+$ T cells in metastases from melanoma patients. *The Journal of clinical investigation*, 121, 2350–60.

Barbacid, M. (2008) Ras oncogenes: their role in neoplasia. *European Journal of Clinical Investigation*, 20, 225–235.

Basso, A.D., Mirza, A., Liu, G., Long, B.J., Bishop, W.R. & Kirschmeier, P. (2005) The farnesyl transferase inhibitor (FTI) SCH66336 (lonafarnib) inhibits Rheb farnesylation and mTOR signaling. Role in FTI enhancement of taxane and tamoxifen anti-tumor activity. *The Journal of biological chemistry*, 280, 31101–8.

Behrens, P., Rothe, M., Wellmann, A., Krischler, J. & Wernert, N. (2001) The Ets-1 transcription factor is up-regulated together with MMP 1 and MMP 9 in the stroma of pre-invasive breast cancer. *The Journal of pathology*, 194, 43–50.

Beissinger, M. & Buchner, J. (1998) How chaperones fold proteins. *Biological chemistry*, 379, 245–59.

Beraldo, F.H., Soares, I.N., Goncalves, D.F., Fan, J., Thomas, A. a, Santos, T.G., *et al.* (2013) Stress-inducible phosphoprotein 1 has unique cochaperone activity during development and regulates cellular response to ischemia via the prion protein. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 27, 3594–607.

Bjornsdottir, G. & Myers, L.C. (2008) Minimal components of the RNA polymerase II transcription apparatus determine the consensus TATA box. *Nucleic acids research*, 36, 2906–16.

Borrego-Diaz, E., Terai, K., Lialyte, K., Wise, A.L., Esfandyari, T., Behbod, F., *et al.* (2012) Overactivation of Ras signaling pathway in CD133+ MPNST cells. *Journal of neuro-oncology*, 108, 423–34.

Bouhouche-Chatelier, L., Chadli, A. & Catelli, M.G. (2001) The N-terminal adenosine triphosphate binding domain of Hsp90 is necessary and sufficient for interaction with estrogen receptor. *Cell stress & chaperones*, 6, 297–305.

Brocchieri, L., Conway de Macario, E. & Macario, A.J.L. (2008) hsp70 genes in the human genome: Conservation and differentiation patterns predict a wide array of overlapping and specialized functions. *BMC evolutionary biology*, 8, 19.

Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *Journal of molecular biology*, 212, 563–578.

Buchner, J. (1996) Supervising molecular the fold : functional pri ciples of. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 10, 10–19.

Bukau, B., Deuerling, E., Pfund, C. & Craig, E.A. (2000) Getting newly synthesized proteins into shape. *Cell*, 101, 119–22.

Cacace, A.M., Ueffing, M., Philipp, A., Han, E.K., Kolch, W. & Weinstein, I.B. (1996) PKC epsilon functions as an oncogene by enhancing activation of the Raf kinase. *Oncogene*, 13, 2517–26.

Cairns, R.A., Harris, I.S. & Mak, T.W. (2011) Regulation of cancer cell metabolism. *Nature reviews. Cancer*, 11, 85–95.

Calderwood, S.K. (2013) Molecular Cochaperones: Tumor Growth and Cancer Treatment. *Scientifica*, 2013, 1–13.

Callaway, K., Waas, W.F., Rainey, M.A., Ren, P. & Dalby, K.N. (2010) Phosphorylation of the transcription factor Ets-1 by ERK2: rapid dissociation of ADP and phospho-Ets-1. *Biochemistry*, 49, 3619–30.

Calvisi, D.F., Ladu, S., Conner, E. a, Seo, D., Hsieh, J.-T., Factor, V.M., *et al.* (2011) Inactivation of Ras GTPase-activating proteins promotes unrestrained activity of wild-type Ras in human liver cancer. *Journal of hepatology*, 54, 311–9.

Caplan, A.J. (2003) What is a co-chaperone? *Cell stress & chaperones*, 8, 105–107.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., *et al.* (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nature genetics*, 38, 626–35.

Carrigan, P.E., Nelson, G.M., Roberts, P.J., Stoffer, J., Riggs, D.L. & Smith, D.F. (2004) Multiple domains of the co-chaperone Hop are important for Hsp70 binding. *The Journal of biological chemistry*, 279, 16185–93.

Cartegni, L., Chew, S.L. & Krainer, A.R. (2002) Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature reviews. Genetics*, 285–298.

Casey, P.J., Solski, P.A., Der, C.J. & Buss, J.E. (1989) p21ras is modified by a farnesyl isoprenoid. *Proceedings of the National Academy of Sciences*, 86, 8323–8327.

Chang, H.C., Nathan, D.F. & Lindquist, S. (1997) In vivo analysis of the Hsp90 cochaperone Sti1 (p60). *Molecular and cellular biology*, 17, 318–25.

Cheung, N., Wong, M.P., Yuen, S.T., Leung, S.Y. & Chung, L.P. (1998) Tissue-specific expression pattern of vascular endothelial growth factor isoforms in the malignant transformation of lung and colon. *Human pathology*, 29, 910–4.

Chi, J.-T., Wang, Z., Nuyten, D.S.A., Rodriguez, E.H., Schaner, M.E., Salim, A., *et al.* (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS medicine*, 3, e47.

Cho, J.Y., Lim, J.Y., Cheong, J.H., Park, Y.-Y., Yoon, S.-L., Kim, S.M., *et al.* (2011) Gene expression signature-based prognostic risk score in gastric cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 17, 1850–7.

Chong, A., Zhang, Z., Choi, K.P., Choudhary, V., Djamgoz, M.B.A., Zhang, G., *et al.* (2007) Promoter profiling and coexpression data analysis identifies 24 novel genes that are coregulated with AMPA receptor genes, GRIAs. *Genomics*, 89, 378–384.

Chrestensen, C.A., Shuman, J.K., Eschenroeder, A., Worthington, M., Gram, H. & Sturgill, T.W. (2007) MNK1 and MNK2 regulation in HER2-overexpressing breast cancer lines. *The Journal of biological chemistry*, 282, 4243–52.

Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., *et al.* (2007) Integration of biological networks and gene expression data using Cytoscape. *Nature protocols*, 2, 2366–82.

Corbit, K.C., Trakul, N., Eves, E.M., Diaz, B., Marshall, M. & Rosner, M.R. (2003) Activation of Raf-1 signaling by protein kinase C through a mechanism involving Raf kinase inhibitory protein. *The Journal of biological chemistry*, 278, 13061–8.

Courgeon, A.M., Maisonhaute, C. & Best-Belpomme, M. (1984) Heat shock proteins are induced by cadmium in Drosophila cells. *Experimental cell research*, 153, 515–21.

Cox, A.D., Hisaka, M.M., Buss, J.E. & Der, C.J. (1992) Specific isoprenoid modification is required for function of normal, but not oncogenic, Ras protein. *Molecular and cellular biology*, 12, 2606–15.

Creighton, C.J., Hilger, A.M., Murthy, S., Rae, J.M., Chinnaiyan, A.M. & El-Ashry, D. (2006) Activation of mitogen-activated protein kinase in estrogen receptor alpha-positive breast cancer cells in vitro induces an in vivo molecular phenotype of estrogen receptor alpha-negative human breast tumors. *Cancer research*, 66, 3903–11.

Csermely, P., Schnaider, T., Soti, C., Prohászka, Z. & Nardai, G. (1998) The 90-kDa molecular chaperone family: structure, function, and clinical applications. A comprehensive review. *Pharmacology & therapeutics*, 79, 129–68.

Czar, M.J., Owens-Grillo, J.K., Dittmar, K.D., Hutchison, K.A., Zacharek, A.M., Leach, K.L., *et al.* (1994) Characterization of the protein-protein interactions determining the heat shock protein (hsp90.hsp70.hsp56) heterocomplex. *The Journal of biological chemistry*, 269, 11155–61.

Das, A.K., Cohen, P.W. & Barford, D. (1998) The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *The EMBO journal*, 17, 1192–9.

Davenport, E.L., Zeisig, A., Aronson, L.I., Moore, H.E., Hockley, S., Gonzalez, D., *et al.* (2010) Targeting heat shock protein 72 enhances Hsp90 inhibitor-induced apoptosis in myeloma. *Leukemia*, 24, 1804–7.

Davis, R.J. (1999) Signal transduction by the c-Jun N-terminal kinase. *Biochemical Society symposium*, 64, 1–12.

Dhomen, N. & Marais, R. (2007) New insight into BRAF mutations in cancer. *Current opinion in genetics & development*, 17, 31–9.

Echeverría, P.C., Bernthaler, A., Dupuis, P., Mayer, B. & Picard, D. (2011) An interaction network predicted from public data as a discovery tool: application to the Hsp90 molecular chaperone machine. *PloS one*, 6, e26044.

Eckert, L.B., Repasky, G.A., Ulkü, A.S., McFall, A., Zhou, H., Sartor, C.I., *et al.* (2004) Involvement of Ras activation in human breast cancer cell signaling, invasion, and anoikis. *Cancer research*, 64, 4585–92.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32, 1792–7.

Erlich, R.B., Kahn, S.A., Lima, F.R.S., Muras, A.G., Martins, R.A.P., Linden, R., *et al.* (2007) STI1 promotes glioma proliferation through MAPK and PI3K pathways. *Glia*, 55, 1690–8.

Escusa-Toret, S., Vonk, W.I.M. & Frydman, J. (2013) Spatial sequestration of misfolded proteins by a dynamic chaperone pathway enhances cellular fitness during stress. *Nature cell biology*, 15, 1231–43.

Eustace, B.K., Sakurai, T., Stewart, J.K., Yimlamai, D., Unger, C., Zehetmeier, C., *et al.* (2004) Functional proteomic screens reveal an essential extracellular role for hsp90 alpha in cancer cell invasiveness. *Nature cell biology*, 6, 507–14.

Ewing, R.M., Chu, P., Elisma, F., Li, H., Taylor, P., Climie, S., *et al.* (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular systems biology*, 3, 89.

Farnsworth, C.L. & Feig, L.A. (1991) Dominant inhibitory mutations in the Mg(2+)-binding site of RasH prevent its activation by GTP. *Molecular and cellular biology*, 11, 4822–9.

Farré, D., Roset, R., Huerta, M., Adsuara, J.E., Roselló, L., Albà, M.M., *et al.* (2003) Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic acids research*, 31, 3651–3.

Faustino, T.A.. & Cooper, T.A.. (2003) Pre-mRNA splicing and human disease. *Genes & development*, 419–437.

Favata, M.F., Horiuchi, K.Y., Manos, E.J., Daulerio, A.J., Stradley, D.A., Feeser, W.S., *et al.* (1998) Identification of a novel inhibitor of mitogen-activated protein kinase kinase. *The Journal of biological chemistry*, 273, 18623–32.

Fernández-Medarde, A. & Santos, E. (2011) Ras in cancer and developmental diseases. *Genes & cancer*, 2, 344–58.

Fierro-Monti, I., Echeverria, P., Racle, J., Hernandez, C., Picard, D. & Quadroni, M. (2013) Dynamic impacts of the inhibition of the molecular chaperone Hsp90 on the T-cell proteome have implications for anti-cancer therapy. *PloS one*, 8, e80425.

Forafonov, F., Toogun, O.A., Grad, I., Suslova, E., Freeman, B.C. & Picard, D. (2008) p23/Sba1p protects against Hsp90 inhibitors independently of its intrinsic chaperone activity. *Molecular and cellular biology*, 28, 3446–56.

Frith, M.C., Hansen, U. & Weng, Z. (2001) Detection of cis-element clusters in higher eukaryotic DNA. *Bioinformatics (Oxford, England)*, 17, 878–89.

Frontini, M., Imbriano, C., diSilvio, A., Bell, B., Bogni, A., Romier, C., *et al.* (2002) NF-Y recruitment of TFIID, multiple interactions with histone fold TAF(II)s. *The Journal of Biological Chemistry*, 277, 5841–5848.

Fukuyo, Y., Hunt, C. & Horikoshi, N. (2010) Geldanamycin and its anti-cancer activities. *Cancer letters*, 290, 24–35.

Gaiser, A.M., Brandt, F. & Richter, K. (2009) The non-canonical Hop protein from Caenorhabditis elegans exerts essential functions and forms binary complexes with either Hsc70 or Hsp90. *Journal of molecular biology*, 391, 621–34.

Gibbs, J.B., Pompliano, D.L., Mosser, S.D., Rands, E., Lingham, R.B., Singh, S.B., *et al.* (1993) Selective inhibition of farnesyl-protein transferase blocks ras processing in vivo. *The Journal of biological chemistry*, 268, 7617–20.

Goodson, M.L., Park-Sarge, O.K. & Sarge, K.D. (1995) Tissue-dependent expression of heat shock factor 2 isoforms with distinct transcriptional activities. *Molecular and cellular biology*, 15, 5288–93.

Grad, I., Cederroth, C.R., Walicki, J., Grey, C., Barluenga, S., Winssinger, N., *et al.* (2010) The molecular chaperone Hsp90α is required for meiotic progression of spermatocytes beyond pachytene in the mouse. *PloS one*, 5, e15770.

Grbovic, O.M., Basso, A.D., Sawai, A., Ye, Q., Friedlander, P., Solit, D., *et al.* (2006) V600E B-Raf requires the Hsp90 chaperone for stability and is degraded in response to Hsp90 inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 57–62.

Guerrero, S., Casanova, I., Farré, L., Mazo, A., Capella, G. & Mangues, R. (2000) K- ras Codon 12 Mutation Induces Higher Level of Resistance to Apoptosis and Predisposition to Anchorage-independent Growth Than Codon 13 Mutation or Proto-Oncogene Overexpression. *Cancer research*, 60, 6750–6756.

Guettouche, T., Boellmann, F., Lane, W.S. & Voellmy, R. (2005) Analysis of phosphorylation of human heat shock factor 1 in cells experiencing a stress. *BMC biochemistry*, 6, 4.

Guo, Y., Guettouche, T., Fenna, M., Boellmann, F., Pratt, W.B., Toft, D.O., *et al.* (2001) Evidence for a mechanism of repression of heat shock factor 1 transcriptional activity by a multichaperone complex. *The Journal of biological chemistry*, 276, 45791–45799.

Haga, N., Saito, S., Tsukumo, Y., Sakurai, J., Furuno, A., Tsuruo, T., *et al.* (2010) Mitochondria regulate the unfolded protein response leading to cancer cell survival under glucose deprivation conditions. *Cancer science*, 101, 1125–32.

Hajj, G.N.M., Arantes, C.P., Dias, M.V.S., Roffé, M., Costa-Silva, B., Lopes, M.H., *et al.* (2013) The unconventional secretion of stress-inducible protein 1 by a heterogeneous population of extracellular vesicles. *Cellular and molecular life sciences : CMLS*, 70, 3211–27.

Hall, B.G. (2013) Building phylogenetic trees from molecular data with MEGA. *Molecular biology and evolution*, 30, 1229–35.

Hall, T. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*, 41, 95–98.

Hamilton, M. & Wolfman, a. (1998) Ha-ras and N-ras regulate MAPK activity by distinct mechanisms in vivo. *Oncogene*, 16, 1417–28.

Han, W., Li, X. & Fu, X. (2011) The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutation research*, 727, 86–103.

Hanahan, D., Jessee, J. & Bloom, F.R. (1991) Plasmid transformation of Escherichia coli and other bacteria. *Methods in Enzymology*, Methods in Enzymology, 204, 63–113.

Hanahan, D. & Weinberg, R.A. (2000) The hallmarks of cancer. *Cell*, 100, 57–70.

Hanahan, D. & Weinberg, R.A. (2011) Hallmarks of cancer: the next generation. *Cell*, 144, 646–74.

Hancock, C.N., Macias, A., Lee, E.K., Yu, S.Y., Mackerell, A.D. & Shapiro, P. (2005) Identification of novel extracellular signal-regulated kinase docking domain inhibitors. *Journal of medicinal chemistry*, 48, 4586–95.

Hanlon, M., Sturgill, T.W. & Sealy, L. (2001) ERK2- and p90(Rsk2)-dependent pathways regulate the CCAAT/enhancer-binding protein-beta interaction with serum response factor. *The Journal of biological chemistry*, 276, 38449–56.

Hartl, F.U. & Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science (New York, N.Y.)*, 295, 1852–8.

Heikkila, J.J., Schultz, G.A., Iatrou, K. & Gedamu, L. (1982) Expression of a set of fish genes following heat or metal ion exposure. *The Journal of biological chemistry*, 257, 12000–5.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O. V, *et al.* (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic acids research*, 26, 362–7.

Hernández, M.P., Sullivan, W.P. & Toft, D.O. (2002) The assembly and intermolecular properties of the hsp70-Hop-hsp90 molecular chaperone complex. *The Journal of biological chemistry*, 277, 38294–304.

Hiang, Y.Y., Sian, K.C., Ling, F.L., Too, S. & Cun, W. (2010) Cloning and Activity Analysis of Human Choline Kinase Beta Promoter. *Health and the Environment Journal*, 1, 12–16.

Hill, C.S. & Treisman, R. (1995) Transcriptional regulation by extracellular signals: mechanisms and specificity. *Cell*, 80, 199–211.

Hirai, S. (1996) Protein Kinase C delta Activates the MEK-ERK Pathway in a Manner Independent of Ras and Dependent on Raf. *Journal of biological chemistry*, 271, 23512–23519.

Ho, C., Wang, C., Mattu, S., Destefanis, G., Ladu, S., Delogu, S., *et al.* (2012) AKT (v-akt murine thymoma viral oncogene homolog 1) and N-Ras (neuroblastoma ras viral oncogene homolog) coactivation in the mouse liver promotes rapid carcinogenesis by way of mTOR (mammalian target of rapamycin complex 1), FOXM1 (forkhead box M1)/SKP2,. *Autoimmune, Cholestatic and Biliary Disease*, 55, 833–45.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.-L., *et al.* (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature*, 415, 180–3.

Hoeflich, K.P., O'Brien, C., Boyd, Z., Cavet, G., Guerrero, S., Jung, K., *et al.* (2009) In vivo antitumor activity of MEK and phosphatidylinositol 3-kinase inhibitors in basal-like breast cancer models. *Clinical cancer research*, 15, 4649–64.

Holmberg, C. & Hietakangas, V. (2001) Phosphorylation of serine 230 promotes inducible transcriptional activity of heat shock factor 1. *The EMBO ...*, 20, 3800–3810.

Holmes, J.L., Sharp, S.Y., Hobbs, S. & Workman, P. (2008) Silencing of HSP90 cochaperone AHA1 expression decreases client protein activation and increases cellular sensitivity to the HSP90 inhibitor 17-allylamino-17-demethoxygeldanamycin. *Cancer research*, 68, 1188–97.

Hopper, A.K., Pai, D.A. & Engelke, D.R. (2010) Cellular dynamics of tRNAs and their genes. *FEBS letters*, 584, 310–7.

Horibe, T., Kohno, M., Haramoto, M., Ohara, K. & Kawakami, K. (2011) Designed hybrid TPR peptide targeting Hsp90 as a novel anticancer agent. *Journal of translational medicine*, 9, 8.

Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabé, R.R., *et al.* (2010) International network of cancer genome projects. *Nature*, 464, 993–8.

Hunger-Glaser, I., Fan, R.S., Perez-Salazar, E. & Rozengurt, E. (2004) PDGF and FGF induce focal adhesion kinase (FAK) phosphorylation at Ser-910: dissociation from Tyr-397 phosphorylation and requirement for ERK activation. *Journal of cellular physiology*, 200, 213–22.

Hunger-Glaser, I., Salazar, E.P., Sinnett-Smith, J. & Rozengurt, E. (2003) Bombesin, lysophosphatidic acid, and epidermal growth factor rapidly stimulate focal adhesion kinase phosphorylation at Ser-910: requirement for ERK activation. *The Journal of biological chemistry*, 278, 22631–43.

Hunter, M.C., O'Hagan, K.L., Kenyon, A., Dhanani, K.C.H., Prinsloo, E. & Edkins, A.L. (2014) Hsp90 binds directly to fibronectin (FN) and inhibition reduces the extracellular fibronectin matrix in breast cancer cells. *PloS one*, 9, e86842.

Ip, Y.T. & Davis, R.J. (1998) Signal transduction by the c-Jun N-terminal kinase (JNK) — from inflammation to development. *Current Opinion in Cell Biology*, 10, 205–219.

Ito, T., Nakayama, T., Ito, M., Naito, S., Kanematsu, T. & Sekine, I. (1998) Expression of the ets-1 proto-oncogene in human pancreatic carcinoma. *Modern pathology*, 11, 209–15.

Ito, Y., Takeda, T., Okada, M. & Matsuura, N. (2002) Expression of ets-1 and ets-2 in colonic neoplasms. *Anticancer research*, 22, 1581–4.

Iwashita, S. & Song, S.-Y. (2008) RasGAPs: a crucial regulator of extracellular stimuli for homeostasis of cellular functions. *Molecular bioSystems*, 4, 213–22.

Jiang, J., Maes, E.G., Taylor, A.B., Wang, L., Hinck, A.P., Lafer, E.M., *et al.* (2007) Structural basis of J cochaperone binding and regulation of Hsp70. *Molecular cell*, 28, 422–33.

Johannessen, C.M., Boehm, J.S., Kim, S.Y., Thomas, S.R., Wardwell, L., Johnson, L.A., *et al.* (2010) COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature*, 468, 968–72.

Johnson, B.D., Schumacher, R.J., Ross, E.D. & Toft, D.O. (1998) Hop modulates Hsp70/Hsp90 interactions in protein folding. *The Journal of biological chemistry*, 273, 3679–86.

Jolly, C. (2000) Role of the Heat Shock Response and Molecular Chaperones in Oncogenesis and Cell Death. *Journal of the National Cancer Institute*, 92, 1564–1572.

Jun, J., Mandoiu, I.I. & Nelson, C.E. (2009) Identification of mammalian orthologs using local synteny. *BMC genomics*, 10, 630–643.

Kamal, A., Thao, L., Sensintaffar, J., Zhang, L., Boehm, M.F., Fritz, L.C., *et al.* (2003) A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature*, 425, 407–10.

Kamb, A., Wee, S. & Lengauer, C. (2007) Why is cancer drug discovery so difficult? *Nature reviews. Drug discovery*, 6, 115–20.

Kampinga, H.H. & Craig, E.A. (2010) The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature reviews. Molecular cell biology*, 11, 579–92.

Kampinga, H.H., Hageman, J., Vos, M.J., Kubota, H., Tanguay, R.M., Bruford, E.A., *et al.* (2009) Guidelines for the nomenclature of the human heat shock proteins. *Cell stress & chaperones*, 14, 105–111.

Kapeli, K. & Hurlin, P.J. (2011) Differential regulation of N-Myc and c-Myc synthesis, degradation, and transcriptional activity by the Ras/mitogen-activated protein kinase pathway. *The Journal of biological chemistry*, 286, 38498–508.

Karnoub, A.E., Symons, M., Campbell, S.L. & Der, C.J. (2004) Molecular basis for Rho GTPase signaling specificity. *Breast cancer research and treatment*, 84, 61–71.

Kawakami, Y., Kitaura, J., Yao, L., McHenry, R.W., Kawakami, Y., Newton, A.C., *et al.* (2003) A Ras activation pathway dependent on Syk phosphorylation of protein kinase C. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 9470–9475.

Khokhlatchev, A. V, Canagarajah, B., Wilsbacher, J., Robinson, M., Atkinson, M., Goldsmith, E., *et al.* (1998) Phosphorylation of the MAP Kinase ERK2 Promotes Its Homodimerization and Nuclear Translocation. *Cell*, 93, 605–615.

Kiefhaber, T., Rudolph, R., Kohler, H.-H. & Buchner, J. (1991) Protein Aggregation in vitro and in vivo: A Quantitative Model of the Kinetic Competition between Folding and Aggregation. *Nature Biotechnology*, 9, 825–829.

Kim, M.J., Pelloux, V., Guyot, E., Tordjman, J., Bui, L.-C., Chevallier, A., *et al.* (2012) Inflammatory pathway genes belong to major targets of persistent organic pollutants in adipose cells. *Environmental health perspectives*, 120, 508–14.

Kim, W.-J., Kim, E.-J., Kim, S.-K., Kim, Y.-J., Ha, Y.-S., Jeong, P., *et al.* (2010) Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. *Molecular cancer*, 9, 1–9.

Kirschke, E., Goswami, D., Southworth, D., Griffin, P.R. & Agard, D.A. (2014) Glucocorticoid receptor function regulated by coordinated action of the Hsp90 and Hsp70 chaperone cycles. *Cell*, 157, 1685–97.

Kittler, R., Zhou, J., Hua, S., Ma, L., Liu, Y., Pendleton, E., *et al.* (2013) A comprehensive nuclear receptor network for breast cancer cells. *Cell reports*, 3, 538–51.

Kortenjann, M., Thomae, O. & Shaw, P.E. (1994) Inhibition of v-raf-dependent c-fos expression and transformation by a kinase-defective mutant of the mitogen-activated protein kinase Erk2. *Molecular and cellular biology*, 14, 4815–24.

Kowenz-Leutz, E., Twamley, G., Ansieau, S. & Leutz, A. (1994) Novel mechanism of C/EBP beta (NF-M) transcriptional control: activation through derepression. *Genes & development*, 8, 2781–91.

Krawczak, M., Reiss, J., and Cooper, D.N. (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. *Human Genetics*, 90, 41–54.

Kubota, H., Yamamoto, S., Itoh, E., Abe, Y., Nakamura, A., Izumi, Y., *et al.* (2010) Increased expression of co-chaperone HOP with HSP90 and HSC70 and complex formation in human colonic carcinoma. *Cell stress & chaperones*, 15, 1003–11.

Kurashina, R., Ohyashiki, J.H., Kobayashi, C., Hamamura, R., Zhang, Y., Hirano, T., *et al.* (2009) Anti-proliferative activity of heat shock protein (Hsp) 90 inhibitors via beta-catenin/TCF7L2 pathway in adult T cell leukemia cells. *Cancer letters*, 284, 62–70.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–5.

Laufen, T., Mayer, M.P., Beisel, C., Klostermeier, D., Mogk, A., Reinstein, J., *et al.* (1999) Mechanism of regulation of Hsp70 chaperones by DnaJ cochaperones. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 5452–5457.

Lee, D.H. & Goldberg, A.L. (1998) Proteasome inhibitors: valuable new tools for cell biologists. *Trends in cell biology*, 8, 397–403.

Lee, M.O., Kim, E.O., Kwon, H.J., Kim, Y.M., Kang, H.J., Kang, H., *et al.* (2002) Radicicol represses the transcriptional function of the estrogen receptor by suppressing the stabilization of the receptor by heat shock protein 90. *Molecular and cellular endocrinology*, 188, 47–54.

Lee, T. & Young, R. (2000) Transcription of eukaryotic protein-coding genes. *Annual review of genetics*, 34, 77–137.

Legrier, M.-E., Yang, C.-P.H., Yan, H.-G., Lopez-Barcons, L., Keller, S.M., Pérez-Soler, R., *et al.* (2007) Targeting protein translation in human non small cell lung cancer via combined MEK and mammalian target of rapamycin suppression. *Cancer research*, 67, 11300–8.

Leung, S.Y., Chen, X., Chu, K.M., Yuen, S.T., Mathy, J., Ji, J., *et al.* (2002) Phospholipase A2 group IIA expression in gastric adenocarcinoma is associated with prolonged survival and less frequent metastasis. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 16203–16208.

Li, J., Soroka, J. & Buchner, J. (2012) The Hsp90 chaperone machinery: conformational dynamics and regulation by co-chaperones. *Biochimica et biophysica acta*, 1823, 624–635.

Li, L.-C. & Dahiya, R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, 18, 1427–1431.

Li, R., Pei, H. & Watson, D.K. (2000) Regulation of Ets function by protein - protein interactions. *Oncogene*, 19, 6514–23.

Li, T. & Sparano, J.A. (2003) Inhibiting Ras signaling in the therapy of breast cancer. *Clinical breast cancer*, 3, 405–16.

Li, Y.-Z., Zhao, P. & Han, W.-D. (2009) Clinicopathological significance of LRP16 protein in 336 gastric carcinoma patients. *World journal of gastroenterology : WJG*, 15, 4833–4837.

Liao, D.-X., Han, W.-D., Zhao, Y.-L., Pu, Y.-D., Mu, Y.-M., Luo, C.-H., *et al.* (2006) Expression and clinical significance of LRP16 gene in human breast cancer. *Chinese journal of cancer*, 25, 866–70.

Liberek, K., Skowyra, D., Zylicz, M., Johnson, C. & Georgopoulos, C. (1991) The Escherichia coli DnaK chaperone, the 70-kDa heat shock protein eukaryotic equivalent, changes conformation upon ATP hydrolysis, thus triggering its dissociation from a bound target protein. *The Journal of biological chemistry*, 266, 14491–6.

Lima, F.R.S., Arantes, C.P., Muras, A.G., Nomizo, R., Brentani, R.R. & Martins, V.R. (2007) Cellular prion protein expression in astrocytes modulates neuronal survival and differentiation. *Journal of neurochemistry*, 103, 2164–76.

Lindahl, G.E., Stock, C.J., Shi-Wen, X., Leoni, P., Sestini, P., Howat, S.L., *et al.* (2013) Microarray profiling reveals suppressed interferon stimulated gene program in fibroblasts from scleroderma-associated interstitial lung disease. *Respiratory research*, 14, 1–14.

Lindquist, S. (1980) Varying patterns of protein synthesis in Drosophila during heat shock: implications for regulation. *Developmental biology*, 77, 463–79.

Liu, G.E., Weirauch, M.T., Tassell, C.P. Van, Li, R.W., Sonstegard, T.S., Matukumalli, L.K., *et al.* (2008) Identification of conserved regulatory elements in mammalian promoter regions: a case study using the PCK1 promoter. *Genomics, proteomics & bioinformatics*, 6, 129–43.

Liu, L., Cao, Y., Chen, C., Zhang, X., McNabola, A., Wilkie, D., *et al.* (2006) Sorafenib blocks the RAF/MEK/ERK pathway, inhibits tumor angiogenesis, and induces tumor cell apoptosis in hepatocellular carcinoma model PLC/PRF/5. *Cancer research*, 66, 11851–8.

Lo, F.-Y., Chang, J.-W., Chang, I.-S., Chen, Y.-J., Hsu, H.-S., Huang, S.-F.K., *et al.* (2012) The database of chromosome imbalance regions and genes resided in lung cancer from Asian and Caucasian identified by array-comparative genomic hybridization. *BMC cancer*, 12, 1–13.

Makhnevych, T. & Houry, W. a. (2012) The role of Hsp90 in protein complex assembly. *Biochimica et biophysica acta*, 1823, 674–82.

Mandal, A.K., Gibney, P.A., Nillegoda, N.B., Theodoraki, M.A., Caplan, A.J. & Morano, K.A. (2010) Hsp110 chaperones control client fate determination in the hsp70-Hsp90 chaperone system. *Molecular biology of the cell*, 21, 1439–48.

Mankame, T.P., Zhou, G. & Lingen, M.W. (2010) Identification and characterization of the human NOL7 gene promoter. *Gene*, 456, 36–44.

Marais, R., Wynne, J. & Treisman, R. (1993) The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell*, 73, 381–93.

Martins, V.R., Graner, E., Garcia-Abreu, J., Souza, S.J. De, Mercadante, A.F., Veiga, S.S., *et al.* (1997) Complementary hydropathy identifies a cellular prion protein receptor. *Nature Medicine*, 3, 1376–1382.

Matthews, K. (1992) DNA looping. *Microbiological reviews*, 56, 123–136.

Mayer, M.P. & Bukau, B. (2005) Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and molecular life sciences*, 62, 670–84.

McLaughlin, S.H., Sobott, F., Yao, Z., Zhang, W., Nielsen, P.R., Grossmann, J.G., *et al.* (2006) The co-chaperone p23 arrests the Hsp90 ATPase cycle to trap client proteins. *Journal of molecular biology*, 356, 746–58.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., *et al.* (2013) Analysis Tool Web Services from the EMBL-EBI. *Nucleic acids research*, 41, W597–600.

Meng, Y.G., Han, W.D., Zhao, Y.L., Huang, K., Si, Y.L., Wu, Z.Q., *et al.* (2007) Induction of the LRP16 gene by estrogen promotes the invasive growth of Ishikawa human endometrial cancer cells through the downregulation of E-cadherin. *Cell research*, 17, 869–80.

Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J. & Alba, M.M. (2002) PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics*, 18, 333–334.

Meyer, P., Prodromou, C., Liao, C., Hu, B., Roe, S.M., Vaughan, C.K., *et al.* (2004) Structural basis for recruitment of the ATPase activator Aha1 to the Hsp90 chaperone machinery. *The EMBO journal*, 23, 1402–10.

Michalak, P. (2008) Coexpression, coregulation, and cofunctionality of neighboring genes in eukaryotic genomes. *Genomics*, 91, 243–248.

Michel, G.P. & Starka, J. (1986) Effect of ethanol and heat stresses on the protein pattern of Zymomonas mobilis. *Journal of bacteriology*, 165, 1040–2.

Mirzoeva, O.K., Das, D., Heiser, L.M., Bhattacharya, S., Siwak, D., Gendelman, R., *et al.* (2009) Basal subtype and MAPK/ERK kinase (MEK)-phosphoinositide 3-kinase feedback signaling determine susceptibility of breast cancer cells to MEK inhibition. *Cancer research*, 69, 565–72.

Mizrak, S.C., Bogerd, J., Lopez-Casas, P.P., Párraga, M., Mazo, J. Del & Rooij, D.G. de. (2006) Expression of stress inducible protein 1 (Stip1) in the mouse testis. *Molecular reproduction and development*, 73, 1361–6.

Molhoek, K.R., Brautigan, D.L. & Slingluff, C.L. (2005) Synergistic inhibition of human melanoma proliferation by combination treatment with B-Raf inhibitor BAY43-9006 and mTOR inhibitor Rapamycin. *Journal of translational medicine*, 3, 39.

Morimoto, R.I. (1998) Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes & Development*, 12, 3788–3796.

Möröy, T. & Heyd, F. (2007) The impact of alternative splicing in vivo: mouse models show the way. *RNA*, 13, 1155–1171.

Mukhopadhyay, A., Deplancke, B., Walhout, A.J.M. & Tissenbaum, H.A. (2008) Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. *Nature protocols*, 3, 698–709.

Nabhan, A.R. & Sarkar, I.N. (2014) Structural network analysis of biological networks for assessment of potential disease model organisms. *Journal of biomedical informatics*, 47, 178–91.

Nadeau, J.H. (1989) Maps of linkage and synteny homologies between mouse and man. *Trends in Genetics*, 5, 82–86.

Nakai, A., Kawazoe, Y., Tanabe, M., Nagata, K. & Morimoto, R.I. (1995) The DNA-binding properties of two heat shock factors, HSF1 and HSF3, are induced in the avian erythroblast cell line HD6. *Molecular and cellular biology*, 15, 5268–78.

Nakai, A., Tanabe, M., Kawazoe, Y., Inazawa, J., Morimoto, R.I. & Nagata, K. (1997) HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. *Molecular and cellular biology*, 17, 469–81.

Nakajima, T., Kinoshita, S., Sasagawa, T., Sasaki, K., Naruto, M., Kishimoto, T., *et al.* (1993) Phosphorylation at threonine-235 by a ras-dependent mitogen-activated protein kinase cascade is essential for transcription factor NF-IL6. *Proceedings of the National Academy of Sciences of the United States of America*, 90, 2207–11.

Nakayama, T., Ito, M., Ohtsuru, A., Naito, S., Nakashima, M. & Sekine, I. (1999) Expression of the ets-1 proto-oncogene in human thyroid tumor. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc*, 12, 61–8.

Nassar, N., Singh, K. & Garcia-Diaz, M. (2010) Structure of the dominant negative S17N mutant of Ras. *Biochemistry*, 49, 1970–4.

Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R.C., Lee, H., *et al.* (2010) Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature*, 468, 973–7.

Nelson, G.M., Huffman, H. & Smith, D.F. (2003) Comparison of the carboxy-terminal DP-repeat region in the co-chaperones Hop and Hip. *Cell stress & chaperones*, 8, 125–33.

Nemoto, T., Ohara-Nemoto, Y. & Ota, M. (1992) Association of the 90-kDa heat shock protein does not affect the ligand-binding ability of androgen receptor. *The Journal of steroid biochemistry and molecular biology*, 42, 803–12.

Nicolet, C.M. & Craig, E.A. (1989) Isolation and characterization of STI1, a stress-inducible gene from Saccharomyces cerevisiae. *Molecular and cellular biology*, 9, 3638–46.

Nissim, O., Melis, M., Diaz, G., Kleiner, D.E., Tice, A., Fantola, G., *et al.* (2012) Liver regeneration signature in hepatitis B virus (HBV)-associated acute liver failure identified by gene expression profiling. *PloS one*, 7, e49611.

Notredame, C., Higgins, D.G. & Heringa, J. (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment11Edited by J. Thornton. *Journal of Molecular Biology*, 302, 205–217.

Nover, L., Scharf, K.D. & Neumann, D. (1989) Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Molecular and cellular biology*, 9, 1298–308.

Obayashi, T., Hayashi, S., Shibaoka, M., Saeki, M., Ohta, H. & Kinoshita, K. (2008) COXPRESdb: a database of coexpressed gene networks in mammals. *Nucleic acids research*, 36, D77–82.

Odunuga, O.O., Hornby, J.A., Bies, C., Zimmermann, R., Pugh, D.J. & Blatch, G.L. (2003) Tetratricopeptide repeat motif-mediated Hsc70-mSTI1 interaction. Molecular characterization of the critical contacts for successful binding and specificity. *The Journal of biological chemistry*, 278, 6896–904.

Odunuga, O.O., Longshaw, V.M. & Blatch, G.L. (2004) Hop: more than an Hsp70/Hsp90 adaptor protein. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 26, 1058–68.

Oikawa, T. & Yamada, T. (2003) Molecular biology of the Ets family of transcription factors. *Gene*, 303, 11–34.

Omerovic, J., Hammond, D.E., Clague, M.J. & Prior, I. a. (2008) Ras isoform abundance and signalling in human cancer cell lines. *Oncogene*, 27, 2754–62.

Onuoha, S.C., Coulstock, E.T., Grossmann, J.G. & Jackson, S.E. (2008) Structural Studies on the Co-chaperone Hop and Its Complexes with Hsp90. *Journal of molecular biology*, 379, 732–744.

Ostling, P., Björk, J.K., Roos-Mattjus, P., Mezger, V. & Sistonen, L. (2007) Heat shock factor 2 (HSF2) contributes to inducible expression of hsp genes through interplay with HSF1. *The Journal of biological chemistry*, 282, 7077–86.

Panaretou, B., Siligardi, G., Meyer, P., Maloney, A., Sullivan, J.K., Singh, S., *et al.* (2002) Activation of the ATPase activity of hsp90 by the stress-regulated cochaperone aha1. *Molecular cell*, 10, 1307–18.

Pandey, S.P. & Krishnamachari, a. (2006) Computational analysis of plant RNA Pol-II promoters. *Bio Systems*, 83, 38–50.

Patwardhan, C.A., Fauq, A., Peterson, L.B., Miller, C., Blagg, B.S.J. & Chadli, A. (2013) Gedunin inactivates the co-chaperone p23 protein causing cancer cell death by apoptosis. *The Journal of biological chemistry*, 288, 7313–25.

Paul, S., Barker, C.A., Turner, H.C., McLane, A., Wolden, S.L. & Amundson, S.A. (2011) Prediction of in vivo radiation dose status in radiotherapy patients using ex vivo and in vivo gene expression signatures. *Radiation research*, 175, 257–65.

Pedersen, a G., Baldi, P., Chauvin, Y. & Brunak, S. (1999) The biology of eukaryotic promoter prediction--a review. *Computers & chemistry*, 23, 191–207.

Petanidis, S., Hadzopoulou-Cladaras, M. & Salifoglou, A. (2013) Cadmium modulates H-ras expression and caspase-3 apoptotic cell death in breast cancer epithelial MCF-7 cells. *Journal of inorganic biochemistry*, 121, 100–7.

Pimienta, G., Herbert, K.M. & Regan, L. (2011) A compound that inhibits the HOP-Hsp90 complex formation and has unique killing effects in breast cancer cell lines. *Molecular pharmaceutics*, 8, 2252–61.

Pirkkala, L., Nykanen, P. & Sistonen, L. (2001) Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *The FASEB Journal*, 15, 1118–1131.

Planell, N., Lozano, J.J., Mora-Buch, R., Masamunt, M.C., Jimeno, M., Ordás, I., *et al.* (2013) Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut*, 62, 967–76.

Pompliano, D.L., Rands, E., Schaber, M.D., Mosser, S.D., Anthony, N.J. & Gibbs, J.B. (1992) Steady-state kinetic mechanism of Ras farnesyl:protein transferase. *Biochemistry*, 31, 3800–7.

Ponzielli, R., Katz, S., Barsyte-Lovejoy, D. & Penn, L.Z. (2005) Cancer therapeutics: targeting the dark side of Myc. *European journal of cancer*, 41, 2485–501.

Powers, M. V, Clarke, P.A. & Workman, P. (2008) Dual targeting of HSC70 and HSP72 inhibits HSP90 function and induces tumor-specific apoptosis. *Cancer cell*, 14, 250–62.

Prapapanich, V., Chen, S. & Smith, D.F. (1998) Mutation of Hip's carboxy-terminal region inhibits a transitional stage of progesterone receptor assembly. *Molecular and cellular biology*, 18, 944–52.

Prodromou, C., Roe, S.M., O'Brien, R., Ladbury, J.E., Piper, P.W. & Pearl, L.H. (1997) Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell*, 90, 65–75.

Ptashne, M. & Gann, A. (1997) Transcriptional activation by recruitment. *Nature*, 386, 569–577.

Richardson, A.L., Wang, Z.C., Nicolo, A. De, Lu, X., Brown, M., Miron, A., *et al.* (2006) X chromosomal abnormalities in basal-like human breast cancer. *Cancer cell*, 9, 121–32.

Richter, K., Haslbeck, M. & Buchner, J. (2010) The heat shock response: life on the verge of death. *Molecular cell*, 40, 253–66.

Ridley, A.J. (2004) Rho proteins and cancer. *Breast cancer research and treatment*, 84, 13–9.

Riggs, D.L., Cox, M.B., Tardif, H.L., Hessling, M., Buchner, J. & Smith, D.F. (2007) Noncatalytic role of the FKBP52 peptidyl-prolyl isomerase domain in the regulation of steroid hormone signaling. *Molecular and cellular biology*, 27, 8658–69.

Rocha Dias, S. da, Friedlos, F., Light, Y., Springer, C., Workman, P. & Marais, R. (2005) Activated B-RAF is an Hsp90 client protein that is targeted by the anticancer drug 17-allylamino-17-demethoxygeldanamycin. *Cancer research*, 65, 10686–91.

Roe, S.M., Ali, M.M.U., Meyer, P., Vaughan, C.K., Panaretou, B., Piper, P.W., *et al.* (2004) The Mechanism of Hsp90 regulation by the protein kinase-specific cochaperone p50(cdc37). *Cell*, 116, 87–98.

Röhl, A., Tippel, F., Bender, E., Schmid, A.B., Richter, K., Madl, T., *et al.* (2014) Hop/Sti1 phosphorylation inhibits its co-chaperone function. *EMBO reports*, 16, 240–9.

Ronchi, A., Bellorini, M., Mongelli, N. & Mantovani, R. (1995) CCAAT-box binding protein NF-Y (CBF, CP1) recognizes the minor groove and distorts DNA Antonella. *Nucleic acids research*, 23, 4565–4572.

Roskoski, R. (2012) ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacological research : the official journal of the Italian Pharmacological Society*, 66, 105–43.

Ruckova, E., Muller, P., Nenutil, R. & Vojtesek, B. (2012) Alterations of the Hsp70/Hsp90 chaperone and the HOP/CHIP co-chaperone system in cancer. *Cellular & molecular biology letters*, 17, 446–58.

Ruohola, J.K., Valve, E.M., Karkkainen, M.J., Joukov, V., Alitalo, K. & Härkönen, P.L. (1999) Vascular endothelial growth factors are differentially regulated by steroid hormones and antiestrogens in breast cancer cells. *Molecular and cellular endocrinology*, 149, 29–40.

Russell, L.C., Whitt, S.R., Chen, M.S. & Chinkers, M. (1999) Identification of conserved residues required for the binding of a tetratricopeptide repeat domain to heat shock protein 90. *The Journal of biological chemistry*, 274, 20060–3.

Saini, K.S., Loi, S., Azambuja, E. de, Metzger-Filho, O., Saini, M.L., Ignatiadis, M., *et al.* (2013) Targeting the PI3K/AKT/mTOR and Raf/MEK/ERK pathways in the treatment of breast cancer. *Cancer treatment reviews*, 39, 935–46.

Salminen, A., Lehtonen, M., Paimela, T. & Kaarniranta, K. (2010) Celastrol: Molecular targets of Thunder God Vine. *Biochemical and biophysical research communications*, 394, 439–42.

Sanchez-Palencia, A., Gomez-Morales, M., Gomez-Capilla, J.A., Pedraza, V., Boyero, L., Rosell, R., *et al.* (2011) Gene expression profiling reveals novel biomarkers in nonsmall cell lung cancer. *International journal of cancer*, 129, 355–64.

Sandelin, A., Carninci, P., Lenhard, B., Ponjavic, J., Hayashizaki, Y. & Hume, D. a. (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nature reviews. Genetics*, 8, 424–36.

Sandqvist, A., Björk, J.K., Akerfelt, M., Chitikova, Z., Grichine, A., Vourc'h, C., *et al.* (2009) Heterotrimerization of heat-shock factors 1 and 2 provides a transcriptional switch in response to distinct stimuli. *Molecular biology of the cell*, 20, 1340–7.

Sangster, T.A., Lindquist, S. & Queitsch, C. (2004) Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 26, 348–62.

Santos, C.I. & Costa-Pereira, A.P. (2011) Signal transducers and activators of transcription-from cytokine signalling to cancer biology. *Biochimica et Biophysica Acta*, 1816, 38–49.

Sarge, K.D., Murphy, S.P. & Morimoto, R.I. (1993) Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localization and can occur in the absence of stress. *Molecular and cellular biology*, 13, 1392–407.

Sawhney, R.S., Cookson, M.M., Omar, Y., Hauser, J. & Brattain, M.G. (2006) Integrin alpha2-mediated ERK and calpain activation play a critical role in cell adhesion and motility via focal adhesion kinase signaling: identification of a novel signaling pathway. *The Journal of biological chemistry*, 281, 8497–510.

Scammell, J.G., Hubler, T.R., Denny, W.B. & Valentine, D.L. (2003) Organization of the human FK506-binding immunophilin FKBP52 protein gene (FKBP4). *Genomics*, 81, 640–643.

Scheufler, C., Brinker, a, Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., *et al.* (2000a) Structure of TPR domain-peptide complexes: critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. *Cell*, 101, 199–210.

Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., *et al.* (2000b) Structure of TPR Domain–Peptide Complexes. *Cell*, 101, 199–210.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671–675.

Schuetz, T.J., Gallo, G.J., Sheldon, L., Tempst, P. & Kingston, R.E. (1991) Isolation of a cDNA for HSF2: evidence for two heat shock factor genes in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 6911–6915.

Schulte, T.W., An, W.G. & Neckers, L.M. (1997) Geldanamycin-induced destabilization of Raf-1 involves the proteasome. *Biochemical and biophysical research communications*, 239, 655–9.

Schultz, J., Milpetz, F., Bork, P. & Ponting, C.P. (1998) SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences*, 95, 5857–5864.

Sepp-Lorenzino, L. & Rosen, N. (1998) A Farnesyl-Protein Transferase Inhibitor Induces p21 Expression and G1 Block in p53 Wild Type Tumor Cells. *Journal of Biological Chemistry*, 273, 20243–20251.

Seth, A. & Watson, D.K. (2005) ETS transcription factors and their emerging roles in human cancer. *European journal of cancer*, 41, 2462–78.

Shaul, Y.D. & Seger, R. (2007) The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochimica et biophysica acta*, 1773, 1213–26.

Shi, W. & Zhou, W. (2006) Frequency distribution of TATA Box and extension sequences on human promoters. *BMC bioinformatics*, 7, 1–12.

Shimokuni, T., Tanimoto, K., Hiyama, K., Otani, K., Ohtaki, M., Hihara, J., *et al.* (2006) Chemosensitivity prediction in esophageal squamous cell carcinoma: novel marker genes and efficacy-prediction formulae using their expression data. *International journal of oncology*, 28, 1153–62.

Shintani, S., Li, C., Ishikawa, T., Mihara, M., Nakashiro, K. & Hamakawa, H. (2004) Expression of vascular endothelial growth factor A, B, C, and D in oral squamous cell carcinoma. *Oral Oncology*, 40, 13–20.

Sidera, K. & Patsavoudi, E. (2008) Extracellular HSP90: conquering the cell surface. *Cell cycle (Georgetown, Tex.)*, 7, 1564–8.

Simpson, S., Woodworth, C.D. & DiPaolo, J.A. (1997) Altered expression of Erg and Ets-2 transcription factors is associated with genetic changes at 21q22.2-22.3 in immortal and cervical carcinoma cell lines. *Oncogene*, 14, 2149–57.

Skotheim, R.I., Lind, G.E., Monni, O., Nesland, J.M., Abeler, V.M., Fosså, S.D., *et al.* (2005) Differentiation of human embryonal carcinomas in vitro and in vivo reveals expression profiles relevant to normal development. *Cancer research*, 65, 5588–98.

Smith, D.F. (2004) Tetratricopeptide repeat cochaperones in steroid receptor complexes. *Cell stress & chaperones*, 9, 109–21.

Smith, J.R., Billy, E. de, Hobbs, S., Powers, M., Prodromou, C., Pearl, L., *et al.* (2013) Restricting direct interaction of CDC37 with HSP90 does not compromise chaperoning of client proteins. *Oncogene*.

Smith, J.R., Clarke, P.A., Billy, E. de & Workman, P. (2009) Silencing the cochaperone CDC37 destabilizes kinase clients and sensitizes cancer cells to HSP90 inhibitors. *Oncogene*, 28, 157–69.

Smith, J.R. & Workman, P. (2009) Targeting CDC37: an alternative, kinase-directed strategy for disruption of oncogenic chaperoning. *Cell cycle (Georgetown, Tex.)*, 8, 362–72.

Song, Y. & Masison, D.C. (2005) Independent regulation of Hsp70 and Hsp90 chaperones by Hsp70/Hsp90-organizing protein Sti1 (Hop1). *The Journal of biological chemistry*, 280, 34178–85.

Southworth, D.R. & Agard, D.A. (2011) Client-loading conformation of the Hsp90 molecular chaperone revealed in the cryo-EM structure of the human Hsp90:Hop complex. *Molecular cell*, 42, 771–81.

Sözeri, O., Vollmer, K., Liyanage, M., Frith, D., Kour, G., Mark, G.E., *et al.* (1992) Activation of the c-Raf protein kinase by protein kinase C phosphorylation. *Oncogene*, 7, 2259–62.

Spinelli, S.L., Malik, H.S., Consaul, S.A. & Phizicky, E.M. (1998) A functional homolog of a yeast tRNA splicing enzyme is conserved in higher eukaryotes and in Escherichia coli. *Proceedings of the National Academy of Sciences*, 95, 14136–14141.

Spuy, J. van der. (2000) *Murine Stress-Inducible Protein 1 (mSTI1): expression in response to constitutive activation of the RAS oncogene, biochemical characterisation and interaction with the 70 kilodaltan heat shock cognate protein 70 (HSC70). PhD Thesis.*

Srebrow, A. & Kornblihtt, A.R. (2006) The connection between splicing and cancer. *Journal of cell science*, 119, 2635–41.

Sreedhar, A.S. & Csermely, P. (2004) Heat shock proteins in the regulation of apoptosis: new strategies in tumor therapy: a comprehensive review. *Pharmacology & therapeutics*, 101, 227–57.

Stacey, D.W., Feig, L.A. & Gibbs, J.B. (1991) Dominant inhibitory Ras mutants selectively inhibit the activity of either cellular or oncogenic Ras. *Molecular and cellular biology*, 11, 4053–64.

Stargell, L. & Struhl, K. (1996) Mechanisms of transcriptional activation in vivo: two steps forward. *Trends in Genetics*, 12, 311–315.

Stephen, A.G., Esposito, D., Bagni, R.K. & McCormick, F. (2014) Dragging ras back in the ring. *Cancer cell*, 25, 272–81.

Stratton, M.R., Campbell, P.J. & Futreal, P.A. (2009) The cancer genome. *Nature*, 458, 719–24.

Studier, F.W. & Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology*, 189, 113–130.

Sugito, K., Yamane, M., Hattori, H., Hayashi, Y., Tohnai, I., Ueda, M., *et al.* (1995) Interaction between hsp70 and hsp40, eukaryotic homologues of DnaK and DnaJ, in human cells expressing mutant-type p53. *FEBS letters*, 358, 161–4.

Sun, L., Hui, A.-M., Su, Q., Vortmeyer, A., Kotliarov, Y., Pastorino, S., *et al.* (2006) Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. *Cancer cell*, 9, 287–300.

Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K.D., Karras, G.I., *et al.* (2012) Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell*, 150, 987–1001.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular biology and evolution*, 28, 2731–2739.

Tan, S.S., Ahmad, I., Bennett, H.L., Singh, L., Nixon, C., Seywright, M., *et al.* (2011) GRP78 up-regulation is associated with androgen receptor status, Hsp70-Hsp90 client proteins and castrate-resistant prostate cancer. *The Journal of pathology*, 223, 81–7.

Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., *et al.* (2008) An in vivo map of the yeast protein interactome. *Science (New York, N.Y.)*, 320, 1465–70.

Tavaria, M., Gabriele, T., Kola, I. & Anderson, R.L. (1996) A hitchhiker's guide to the human Hsp70 family. *Cell stress & chaperones*, 1, 23–8.

Thorne, J.L., Campbell, M.J. & Turner, B.M. (2009) Transcription factors, chromatin and cancer. *The international journal of biochemistry & cell biology*, 41, 164–75.

Tiroli-Cepeda, A.O. & Ramos, C.H.I. (2011) An overview of the role of molecular chaperones in protein homeostasis. *Protein and peptide letters*, 18, 101–9.

Tong, M., Chan, K.W., Bao, J.Y.J., Wong, K.Y., Chen, J.-N., Kwan, P.S., *et al.* (2012) Rab25 is a tumor suppressor gene with antiangiogenic and anti-invasive activities in esophageal squamous cell carcinoma. *Cancer research*, 72, 6024–35.

Towbin, H., Staehelin, T. & Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America*, 76, 4350–4.

Trepel, J., Mollapour, M., Giaccone, G. & Neckers, L. (2010) Targeting the dynamic HSP90 complex in cancer. *Nature reviews. Cancer*, 10, 537–49.

Trinklein, N.D., Murray, J.I., Hartman, S.J., Botstein, D. & Myers, R.M. (2004) The Role of Heat Shock Transcription Factor 1 in the Genome-wide Regulation of the Mammalian Heat Shock Response. *Molecular biology of the cell*, 15, 1254–1261.

Tsantoulis, P.K. & Gorgoulis, V.G. (2005) Involvement of E2F transcription factor family in cancer. *European journal of cancer*, 41, 2403–14.

U.S. National Institutes of Health. (2014) ClinicalTrials.gov [WWW Document]. URL http://www.clinicaltrials.gov/ct2/results?term="HSP90"+AND+"cancer"&Search=Search [accessed on 2014].

Venables, J.P. (2004) Aberrant and Alternative Splicing in Cancer Aberrant and Alternative Splicing in Cancer. *Cancer research*, 64, 7647–7654.

Vogel, J.L., Parsell, D.A. & Lindquist, S. (1995) Heat-shock proteins Hsp104 and Hsp70 reactivate mRNA splicing after heat inactivation. *Current biology : CB*, 5, 306–17.

Voss, A.K., Thomas, T. & Gruss, P. (2000) Mice lacking HSP90beta fail to develop a placental labyrinth. *Development (Cambridge, England)*, 127, 1–11.

Walsh, N., Larkin, A., Swan, N., Conlon, K., Dowling, P., McDermott, R., *et al.* (2011) RNAi knockdown of Hop (Hsp70/Hsp90 organising protein) decreases invasion via MMP-2 down regulation. *Cancer letters*, 306, 180–9.

Wang, J. & Hannenhalli, S. (2006) A mammalian promoter model links cis elements to genetic networks. *Biochemical and biophysical research communications*, 347, 166–77.

Wang, J., Park, J.-S., Wei, Y., Rajurkar, M., Cotton, J.L., Fan, Q., *et al.* (2013) TRIB2 acts downstream of Wnt/TCF in liver cancer cells to regulate YAP and C/EBPα function. *Molecular cell*, 51, 211–25.

Wang, T.-H., Chao, A., Tsai, C.-L., Chang, C.-L., Chen, S.-H., Lee, Y.-S., *et al.* (2010) Stress-induced phosphoprotein 1 as a secreted biomarker for human ovarian cancer promotes cancer cell proliferation. *Molecular & cellular proteomics : MCP*, 9, 1873–84.

Wang, Z., Zhou, J., Fan, J., Qiu, S.-J., Yu, Y., Huang, X.-W., *et al.* (2008) Effect of rapamycin alone and in combination with sorafenib in an orthotopic model of human hepatocellular carcinoma. *Clinical cancer research*, 14, 5124–30.

Wegele, H., Haslbeck, M., Reinstein, J. & Buchner, J. (2003) Sti1 is a novel activator of the Ssa proteins. *The Journal of biological chemistry*, 278, 25970–6.

Wegele, H., Müller, L. & Buchner, J. (2004) Hsp70 and Hsp90--a relay team for protein folding. *Reviews of physiology, biochemistry and pharmacology*, 151, 1–44.

Welch, W.J. & Suhan, J.P. (1985) Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat-shock treatment. *The Journal of cell biology*, 101, 1198–211.

Wheeler, A.P. & Ridley, A.J. (2004) Why three Rho proteins? RhoA, RhoB, RhoC, and cell motility. *Experimental cell research*, 301, 43–9.

Whelan, S. & Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular biology and evolution*, 18, 691–699.

Wilhelm, S.M., Carter, C., Tang, L., Wilkie, D., McNabola, A., Rong, H., *et al.* (2004) BAY 43-9006 exhibits broad spectrum oral antitumor activity and targets the RAF/MEK/ERK pathway and receptor tyrosine kinases involved in tumor progression and angiogenesis. *Cancer research*, 64, 7099–109.

Wilkerson, D.C., Skaggs, H.S. & Sarge, K.D. (2007) HSF2 binds to the Hsp90, Hsp27, and c-Fos promoters constitutively and modulates their expression. *Cell stress & chaperones*, 12, 283–90.

Willmer, T. (2011) *The role of Hsp90 / Hsp70 organising protein (Hop) in the Proliferation , Survival and Migration of Breast Cancer Cells*. *MSc Thesis*.

Willmer, T., Contu, L., Blatch, G.L. & Edkins, A.L. (2013) Knockdown of Hop downregulates RhoC expression , and decreases pseudopodia formation and migration in cancer cell lines. *Cancer Letters*, 328, 252–260.

Wingender, E., Chen, X., Hehl, R., Karas, H., Liebich, I., Matys, V., *et al.* (2000) TRANSFAC: an integrated system for gene expression regulation. *Nucleic acids research*, 28, 316–9.

Wray, G. a, Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M. V, *et al.* (2003) The evolution of transcriptional regulation in eukaryotes. *Molecular biology and evolution*, 20, 1377–419.

Wu, M., Wu, Z.-F., Kumar-Sinha, C., Chinnaiyan, A. & Merajver, S.D. (2004) RhoC induces differential expression of genes involved in invasion and metastasis in MCF10A breast cells. *Breast cancer research and treatment*, 84, 3–12.

Xi, H.Q., Zhao, P. & Han, W.D. (2010) Clinicopathological significance and prognostic value of LRP16 expression in colorectal carcinoma. *World journal of gastroenterology : WJG*, 16, 1644–8.

Xu, W., Marcu, M., Yuan, X., Mimnaugh, E., Patterson, C. & Neckers, L. (2002) Chaperone-dependent E3 ubiquitin ligase CHIP mediates a degradative pathway for c-ErbB2/Neu. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 12847–52.

Yamamoto, S., Subedi, G.P., Hanashima, S., Satoh, T., Otaka, M., Wakui, H., *et al.* (2014) ATPase activity and ATP-dependent conformational change in the co-chaperone HOP. *The Journal of biological chemistry*, 289, 9880–9886.

Yang, B.S., Hauser, C.A., Henkel, G., Colman, M.S., Beveren, C. Van, Stacey, J., *et al.* (1996) Ras-mediated phosphorylation of a conserved threonine residue enhances the transactivation activities of c-Ets1 and c-Ets2 . Ras-Mediated Phosphorylation of a Conserved Threonine Residue Enhances the Transactivation Activities of c-Ets1 and c-Ets2. *Molecular and cellular biology*, 16, 538–547.

Yasuda, R., Harvey, C.D., Zhong, H., Sobczyk, A., Aelst, L. van & Svoboda, K. (2006) Supersensitive Ras activation in dendrites and spines revealed by two-photon fluorescence lifetime imaging. *Nature neuroscience*, 9, 283–91.

Young, A., Lou, D. & McCormick, F. (2013) Oncogenic and wild-type Ras play divergent roles in the regulation of mitogen-activated protein kinase signaling. *Cancer discovery*, 3, 112–23.

Yura, T., Tobe, T., Ito, K. & Osawa, T. (1984) Heat shock regulatory gene (htpR) of Escherichia coli is required for growth at high temperature but is dispensable at low temperature. *Proceedings of the National Academy of Sciences of the United States of America*, 81, 6803–7.

Zanata, S.M., Lopes, M.H., Mercadante, A.F., Hajj, G.N.M., Chiarini, L.B., Nomizo, R., *et al.* (2002a) Stress-inducible protein 1 is a cell surface ligand for cellular prion that triggers neuroprotection. *The EMBO journal*, 21, 3307–16.

Zanata, S.M., Lopes, M.H., Mercadante, A.F., Hajj, G.N.M., Chiarini, L.B., Nomizo, R., *et al.* (2002b) Stress-inducible protein 1 is a cell surface ligand for cellular prion that triggers neuroprotection. *The EMBO journal*, 21, 3307–16.

Zhang, M., Botër, M., Li, K., Kadota, Y., Panaretou, B., Prodromou, C., *et al.* (2008a) Structural and functional coupling of Hsp90- and Sgt1-centred multi-protein complexes. *The EMBO journal*, 27, 2789–98.

Zhang, T., Hamza, A., Cao, X., Wang, B., Yu, S., Zhan, C.-G., *et al.* (2008b) A novel Hsp90 inhibitor to disrupt Hsp90/Cdc37 complex against pancreatic cancer cells. *Molecular cancer therapeutics*, 7, 162–70.

Zhao, P., Lu, Y. & Han, W. (2010) Clinicopathological significance and prognostic value of leukemia-related protein 16 expression in invasive ductal breast carcinoma. *Cancer science*, 101, 2262–8.

Zhao, R. & Houry, W.A. (2005) Hsp90 : a chaperone for protein folding and gene regulation 1. *Biochemical cellular biology*, 83, 703–710.

Zheng, Y., Xia, Y., Hawke, D., Halle, M., Tremblay, M.L., Gao, X., *et al.* (2009) FAK phosphorylation by ERK primes ras-induced tyrosine dephosphorylation of FAK mediated by PIN1 and PTP-PEST. *Molecular cell*, 35, 11–25.

Zhu, S., Yoon, K., Sterneck, E., Johnson, P.F. & Smart, R.C. (2002) CCAAT/enhancer binding protein-beta is a mediator of keratinocyte survival and skin tumorigenesis involving oncogenic Ras signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 207–12.

Zou, J., Guo, Y., Guettouche, T., Smith, D.F. & Voellmy, R. (1998) Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell*, 94, 471–80.

Zoubarev, A., Hamer, K.M., Keshav, K.D., McCarthy, E.L., Santos, J.R.C., Rossum, T. Van, *et al.* (2012) Gemma: a resource for the reuse, sharing and meta-analysis of expression profiling data. *Bioinformatics (Oxford, England)*, 28, 2272–3.

Chapter 7

Appendix

Electronic Appendixes contains the following:

Appendix A - Total *cis*-elements identified within the HOP promoter.

Table 6: Putative *cis*-elements identified within the HOP promoter

Appendix B - Fluorescence and Confocal Images.

Figure E 19B: Phosphorylation of FAK and MAPK is down-regulated by RAF Inhibition.

Figure E 20: The subcellular localisation of C/EBPβ and ETS-1 in Hs578T cells upon Sorafenib treatment.

Figure E 21A: The subcellular localisation of ETS-1 in Hs578T cells upon ERK2i treatment.

Appendix C - HOP Promoter Truncations

Figure E 23: Preliminary cloning of the truncated HOP promoter.

Chapter 7

Appendix

Appendix A: Total *cis*-elements identified within the HOP promoter.

Table 1: Putative cis-elements identified within the HOP promoter

Cis - element	Position along promoter in relation to TSS	Prediction Tool		
		TF Search	PROMO	Alibaba
GR-beta	-1497		X	
C/EBPbeta	-1494		X	
GR-alpha	-1490		X	
AP2-alpha	-1490		X	
Sp1	-1490			X
Pax-5	-1489		X	
P53	-1489		X	
RXR-alpha	-1482		X	
YY1	-1476		X	
Sp1	-1467			X
AP2-alpha	-1466		X	
C/EBPbeta	-1453		X	
Sp1	-1452			X
IRF-2	-1451		X	
Ik-2	-1448	X		
C/EBPbeta	-1432		X	
C/EBPbeta	-1419		X	
GATA-2	-1418	X		
NF-1	-1415			X
C/EBPbeta	-1414		X	
AR	-1408		X	
TFII-D	-1402		X	
C/EBPbeta	-1401		X	
GR-beta	-1397		X	
CPE	-1396			X
CACCC	-1396			X
YY1	-1396		X	
Sp1	-1395			X
MyoD	-1390			X
Pax-5	-1389		x	
P53	-1389		x	
E1	-1389			X
Olf-1	-1389			X
STAT4	-1381		X	
GR-beta	-1380			
GATA-1	-1373			X
AP2-alpha	-1370		X	
Sp1	-1369			X
Sp1	-1362			X
AP-1	-1361			X
C/EBPbeta	-1357		X	
NF-1	-1349			X
GR-alpha	-1342		X	
RAR-beta	-1331		X	
RXR-alpha	-1329		X	
MIG1	-1328			X
AP-2alpha	-1323			X
STAT4	-1320		X	
GATA-1	-1318		X	
OCT	-1315			X
GATA-1	-1314		X	
HNF-1	-1313			X
GR-beta	-1312		X	
GR-alpha	-1301		X	
PEA3	-1292		X	
T3R-alpha	-1290			X
T3R-beta	-1290			X
AR	-1285		X	
Sp1	-1279			X
C/EBPbeta	-1271		X	
Pax-5	-1276		X	
ENKTF-1	-1264		X	
RXR-alpha	-1258		X	
c-Jun	-1252		X	
c-Fos	-1251		X	
XBP-1	-1250		X	
TCF-1alpha	-1239			X
GR-alpha	-1238		X	

LEF-1	-1237		X	
SRY	-1237		X	
C/EBPbeta	-1235		X	
STAT4	-1234		X	
STATx	-1231	X		
c-Ets-1	-1229		X	
STAT4	-1227		X	
C/EBPbeta	-1223		X	
AP2-alpha	-1222		X	
RXR-alpha	-1217		X	
ER-alpha	-1216		X	
C/EBPbeta	-1205		X	
GR-beta	-1203		X	
Net	-1198			X
OCT	-1196			X
GR-beta	-1196		X	
Sp1	-1189			X
STAT4	-1172		X	
c-Ets-1	-1171		X	
C/EBPbeta	-1161		X	
AML-1a	-1160	X		
AP2-alpha	-1154		X	
Pax-5	-1153		x	
P53	-1153		x	
Sp1	-1149			X
AP2-alpha	-1145		X	
Sp1	-1133	X		X
Pax-5	-1131		X	
P53	-1131		X	
AP2-alpha	-1120		X	
Pax-5	-1119		X	
Sp1	-1118			X
CdxA	-1112	X		
P53	-1107		X	
ER	-1093			X
AP2-alpha	-1088		X	
GR-alpha	-1088		X	
ER-alpha	-1082		X	
FOXP3	-1081		X	
C/EBPbeta	-1080		X	
AP-1	-1074	X		X
GR	-1070			X
GR-alpha	-1046		X	
USF	-1033			X
Sp1	-1025			X
MZF1	-1024	X		
Sp1	-1019			X
P53	-1018		X	
GR-alpha	-1009		X	
Ftz	-1002			X
STAT4	-996		X	
E2F-1	-995		X	
N-Myc	-993	X		
Sp1	-992			X
USF	-990	X		
FOXP3	-984		X	
C/EBPbeta	-983		X	
GR-beta	-971		X	
STAT4	-970		X	
Sp1	-966			X
GR-alpha	-954		X	
C/EBPbeta	-948		X	
NFI/CTF	-948		X	
GCF	-942		X	
Sp1	-939			X
GCF	-935			
GATA-1	-931			X
GR-alpha	-927		X	
RXR-alpha	-924		X	
AML-1a	-918	X		
STAT4	-916		X	
NFI/CTF	-899		X	
C/EBPbeta	-899		X	
E2F-1	-895		X	
Sp1	-894			X
TFII-I	-892		X	
Sp1	-887			X
Pax-5	-885		X	
NF-kappa	-872			X
YY1	-867			X
AP2-alpha	-860		X	
GR-alpha	-859		X	
TFII-I	-858		X	
Sp1	-857			X
Pax-5	-854		X	

P53	-854		X	
YY1	-851			X
C/EBPbeta	-849		X	
Pax-5	-848		X	
P53	-848		X	
NFI/CTF	-843		X	
C/EBPbeta	-843		X	
YY1	-839			X
NF-muE1	-837			X
Arnt	-831			X
USF1	-830		X	
RXR-beta	-830			X
GBF2	-829			X
HBP-1b	-829			X
USF	-829	X		X
ER-alpha	-829		X	
c-Myc	-826		X	
USF1	-826		X	
c-Rel	-825	X		
NF-kap	-825	X		
N-Myc	-824	X		
SREBP	-824	X		
GR-beta	-820		X	
STATx	-820	X		
STAT4	-818		X	
HSF2	-815	X		
HSF1	-813	X		
TFII-I	-818		X	
c-Ets-1	-817		X	
STAT4	-812		X	
PR-B	-810		X	
PR-1	-810		X	
GR-beta	-809		X	
AP-2alpha	-788		X	
Sp1	-788			X
E1	-786			X
MyoD	-786			X
Sp1	-779			X
NF-1	-778			X
Pax-5	-778		X	
p53	-778		X	
Sp1	-772			X
AP-2alpha	-769		X	
Sp1	-766			X
GCF	-759		X	
AR	-754		X	
FOXP3	-753		X	
C/EBPbeta	-752		X	
Sp1	-728			X
AP-2alpha	-725		X	
GR-alpha	-725		X	
Sp1	-719			X
RXR-alpHA	-705		X	
Sp1	-701			X
YY1	-698		X	
GATA-1	-694	X		
Egr-1	-690	X		
Sp1	-689			X
WT1	-688			X
GCF	-688		X	
Sp1	-678			X
AP-2alpha	-677			X
Pax-5	-677		X	
p53	-677		X	
NF-1	-675			X
Pax-5	-671		X	
GR-alpha	-667		X	
E2F-1	-665		X	
GATA-1	-656	X		
Sp1	-656			X
AP-2alpha	-653			X
Sp1	-650			X
Sp1	-642			X
AP-2alpha	-640		X	
STAT4	-636		x	
Sp1	-627			X
Ik-2	-615	X		
IRF-1	-613		X	
GR-alpha	-613		X	
STAT4	-609		x	
C/EBPbeta	-604		X	
Sp1	-600			X
AP-2alpha	-599			X
GR-alpha	-597		X	
AP-2alpha	-597		X	
Sp1	-593			X
Adf-1	-588			X

Sp1	-585			X
AP-2alpha	-584			X
RXR-alpHA	-566		X	
Sp1	-562			X
D1	-548			X
SRF	-543			X
YY1	-541		X	
GATA-1	-536		x	
GATA-1	-536			X
Sp1	-532			X
Sp1	-532	X		
Sp1	-531		X	
Egr-1	-531			X
NF-1	-531			X
Pax-5	-530		X	
ETF	-530			X
p53	-530		X	
RXR-alpHA	-521		X	
Sp1	-520			X
AP-2alpha	-513		X	
GR-alpha	-512		X	
TFII-I	-509		X	
TFIID	-510		X	
C/EBPbeta	-508		X	
Sp1	-505			X
p53	-503		X	
GR-beta	-497		x	
CdxA	-495	X		
SRY	-491	X		
NF-Y	-491		X	
CP1	-490			X
CTF	-489			X
Pbx-1	-489	X		
C/EBPbeta	-488		X	
GR-beta	-486		x	
Ik-2	-485	X		
IRF-1	-485		X	
TFII-I	-481		X	
STAT4	-481		X	
C/EBPbeta	-479		X	
Sp1	-472			X
C/EBPbeta	-471		X	

AP-2	-471			X
GR-alpha	-469		X	
AP2	-469		X	
C/EBPbeta	-468		X	
Sp1	-463			X
Sp1	-463		X	
Adf-1	-460			X
NF-Y	-459		X	
RXR-alpha	-458		X	
Sp1	-457			X
AP-2	-456			X
C/EBPbeta	-456		X	
GR-beta	-455		x	
NF-Y	-453		X	
C/EBPbeta	-452		X	
STAT4	-450		X	
GR-beta	-448		x	
Sp1	-444			X
GR-beta	-440		x	
AP-2alpha	-433		X	
Sp1	-430			X
STAT4	-427		x	
D1	-420			X
YY1	-418		X	
AP-1	-415			X
SRF	-415			X
YY1	-413		X	
Sp1	-411			X
GATA-1	-408			X
NF-Y	-406		X	
Sp1	-404			X
Egr-1	-403			X
NF-1	-403			X
Sp1	-403	X	X	
EFI	-403			X
C/EBPbeta	-403		X	
Pax-5	-402		X	
p53	-402		X	
ETF	-402			X
GR-beta	-401		x	
YY1	-400		X	
TFII-I	-397		X	
RXR-	-393		X	

alpha				
YY1	-390		X	
GR-alpha	-384		X	
TFII-D	-382		X	
C/EBPbeta	-380		X	
Sp1	-377			X
p53	-375		X	
YY1	-372		X	
GR-beta	-369		X	
TFII-I	-369		X	
CdxA	-366	X		
C/EBPbeta	-363		X	
NF-Y	-363		X	
SRY	-362	X		
CP1	-362			X
CTF	-361			X
TFII-I	-361		X	
C/EBPalpha	-361		X	
C/EBPbeta	-360		X	
Pbx-1	-360	X		
GR-beta	-358		X	
YY1	-355		X	
C/EBPbeta	-351		X	
GR	-350		X	
C/EBPbeta	-340		X	
Sp1	-335		X	X
NF-Y	-331		X	
C/EBPalpha	-329		X	
C/EBPbeta	-328		X	
GR-beta	-327		X	
GR-beta	-326		X	
NF-Y	-325		X	
C/EBPbeta	-324		X	
STAT4	-322		X	
GR-beta	-320		X	
TEC1	-315			X
GR-beta	-312		X	
AP2	-305		X	
Sp1	-302			X
STAT4	-299		X	
AP-1	-287			X
Sp1	-283			X
NF-Y	-278		X	
EFI	-275			X
C/EBPbeta	-275		X	
C/EBPalpha	-273		X	
Sp1	-269			X
MZF1	-269	X		
GR-alpha	-267		X	
YY1	-264		X	
YY1	-262		X	
NF-AT1	-262		X	
C/EBPalpha	-262			X
NFKB	-261	X		X
NFKB	-261			X
NF-Y	-261		X	
STAT4	-261		X	
HSF2	-261	X		
c-Rel	-260	X		
NF-AT1	-260		X	
TFII-I	-257		X	
HSTF	-257			X
C/EBPalpha	-240			X
CTF	-239			X
EFI	-238			X
STAT4	-257		X	
HSF2	-251	X		
HSF1	-251	X		
C/EBPbeta	-238		X	
Pax-5	-233		X	
p53	-233		X	
GCF	-231		X	
Sp1	-230			X
GCF	-227		X	X
NF-1	-221			X
NFI/CTF	-214		X	X
GR-beta	-213		X	
C/EBPalpha	-212		X	
NF-Y	-211		X	
C/EBPbeta	-210		X	
NF-1	-210			X
C/EBPalpha	-206			X

TFII-D	-203		X	X
Sp1	-200	X		
NF-Y	-197		X	
C/EBPalpha	-195		X	
C/EBPbeta	-194		X	X
TFII-I	-193		X	
Sp1	-173			
SRF	-162			X
GR-beta	-162		X	X
Sp1	-158			
AP-2	-156			X
AP2	-151		X	X
GATA-1	-147	X		
GATA-2	-147	X		
Sp1	-143	X		
RAP1	-143			X
RXR-alpha	-138		X	X
TBP	-135		X	
CdxA	-135	X		
Sp1	-132	X		
GR-alpha	-130		X	X
Sp1	-123			X
NF-Y	-123		X	X
C/EBPalpha	-121		X	
C/EBPbeta	-120		X	
TFII-I	-119		X	
AML-1a	-112	X		
Egr-1	-110	X		
Sp1	-110			X
MZF1	-107	X		
WT1	-107			X
Sp1	-106			X
AP2	-102		X	X
RXR-alpha	-95		X	
FOX-P3	-93		X	
C/EBPbeta	-92		X	
STAT4	-87		X	
GR-beta	-86		X	
GR-beta	-85		X	
Sp1	-82			X
TFII-I	-73		X	X
NF-Y	-72		X	
C/EBPbeta	-69		X	
OCT	-67	X		
GR-beta	-67		X	X
YY1	-66		X	
Ets-2	-64		X	
Sp1	-64			X
Elk-1	-63		X	X
C-Ets-1	-62			X
MAF	-62			X
Ets-1	-61		X	X
STAT4	-59		X	
GATA-1	-52	X		
GATA-2	-52	X		
Sp1	-45		X	
Pax-5	-44		X	X
p53	-44		X	
T3R	-40	X		
TER-beat	-40			X
CP2	-36	X		
C/EBPbeta	-34		X	X
CT-1	-33	X		
NF-Y	-32		X	
C/EBPbeta	-29		X	X
GR-beta	-27		X	X
RelA	-10		X	
GR-beta	-8		X	
TFII-I	-8		X	
STAT4	-8		X	
AP2	21		X	
C/EBPbeta	24		X	
Ets-2	24		X	
Elk-1	25		X	
Ets-1	27		X	
Sp1	33		X	
Pax-5	34		X	X
p53	34		X	
AP-2	38	X		
Sp1	39			X
RXR-alpha	46		X	X
STAT4	67		X	
Sp1	70	X		
AP2	74		X	X
Sp1	76	X		

AML-1a	**86**	**X**		
NFI/CTF	**87**		**X**	**X**
Ik-2	**90**	**X**		
C/EBPbeta	**91**		**X**	
STAT4	**94**	**X**		

Appendix B: Fluorescence and Confocal Images.

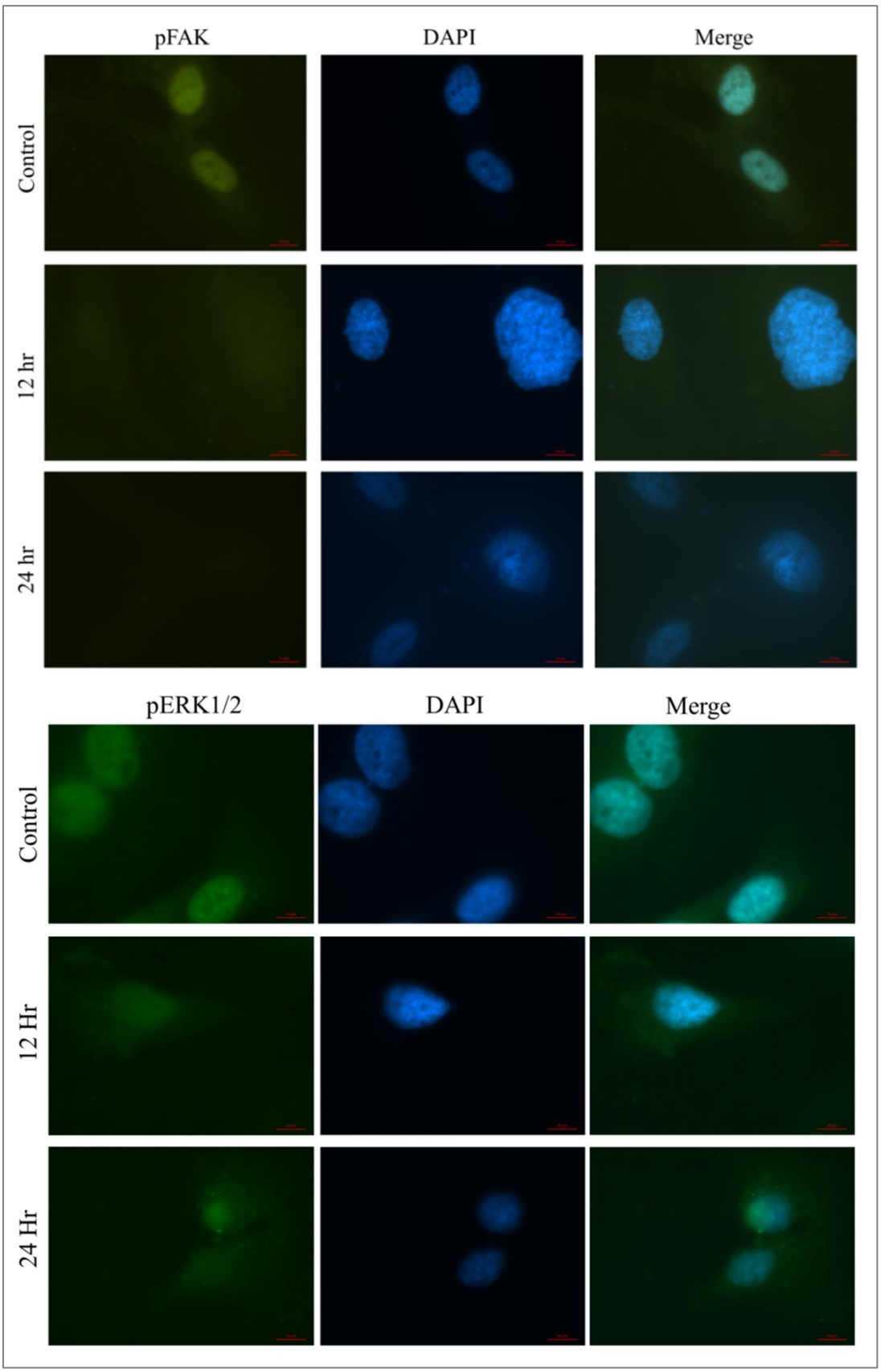


Figure E19B: Phosphorylation of FAK and ERK1/2 is down-regulated by RAF Inhibition.

The subcellular localisation of pFAK and pERK1/2 in Hs578T cells was determined using indirect immunofluorescence staining and fluorescence microscopy. Hs578T cells were seeded overnight onto glass coverslips and treated with 3.6 μM Sorafenib or dimethyl sulfoxide (DMSO) for 12 or 24 hours. Cells were incubated with pFAK or pERK1/2 primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss AxioVert AI FL-LED inverted fluorescence microscope with a 100x objective and were analysed by Zen Lite 2012. The scale bars indicate a length of 10 μm. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown.

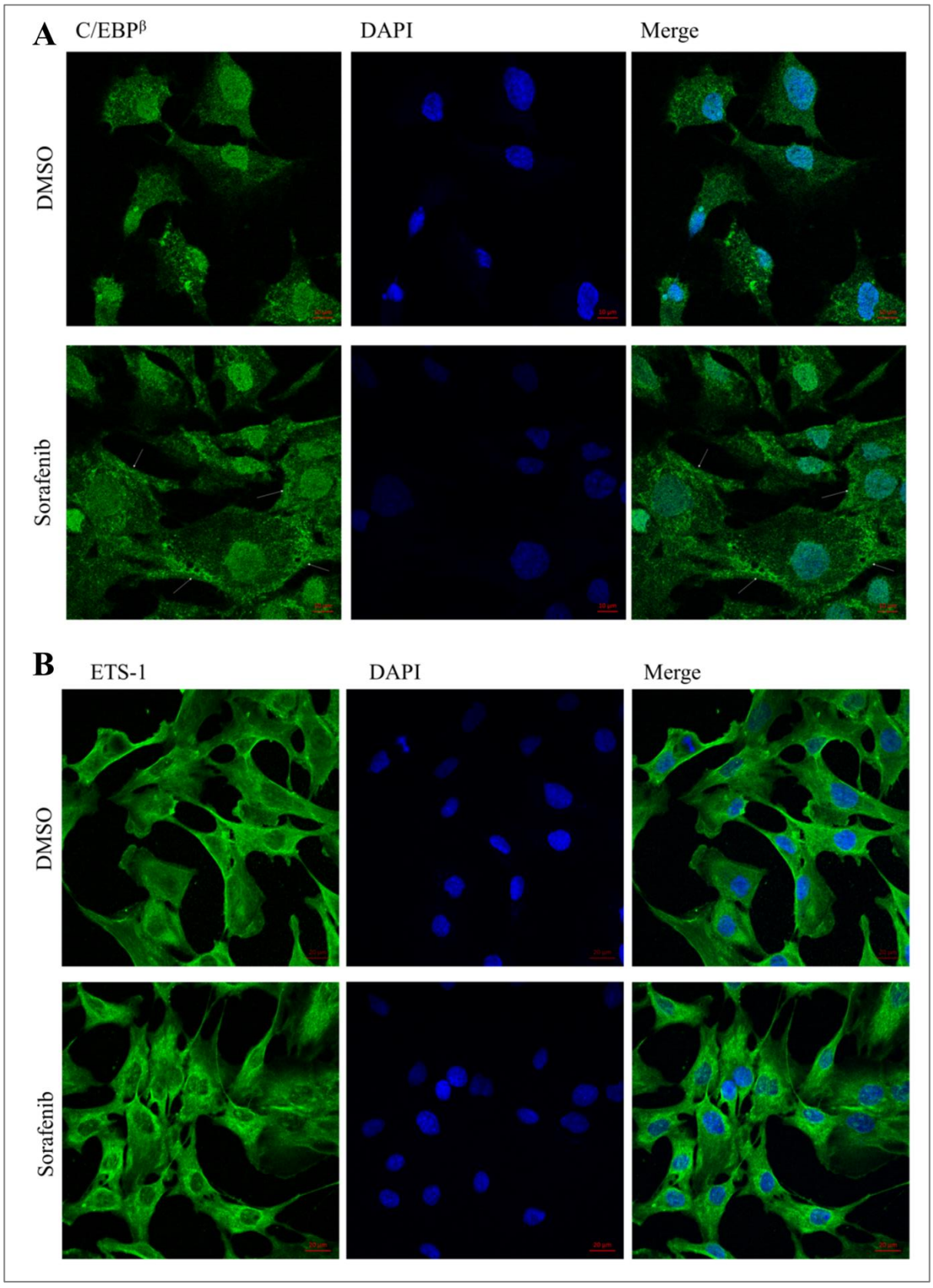


Figure E20: The subcellular localisation of C/EBP$^{\beta}$ and ETS-1 in Hs578T cells upon Sorafenib treatment.

The subcellular localisation of (A) C/EBP$^{\beta}$ and (B) ETS-1 in Hs578T cells was determined using indirect immunofluorescence staining and confocal microscopy. Hs578T cells were seeded overnight onto glass coverslips and treated with 3.6 µM Sorafenib or dimethyl sulfoxide (DMSO) for 24 hours. Cells were

incubated with C/EBP$^{\beta}$ or ETS-1 primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss LSM 780 Confocal Microscope with 60x and 100x objectives and were analysed by Zen Lite 2012. The scale bars represent a length of 10 µm or 20 µm as indicated. White arrows indicate C/EBP$^{\beta}$ staining in peripheral clusters within the Hs578T cells. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown.

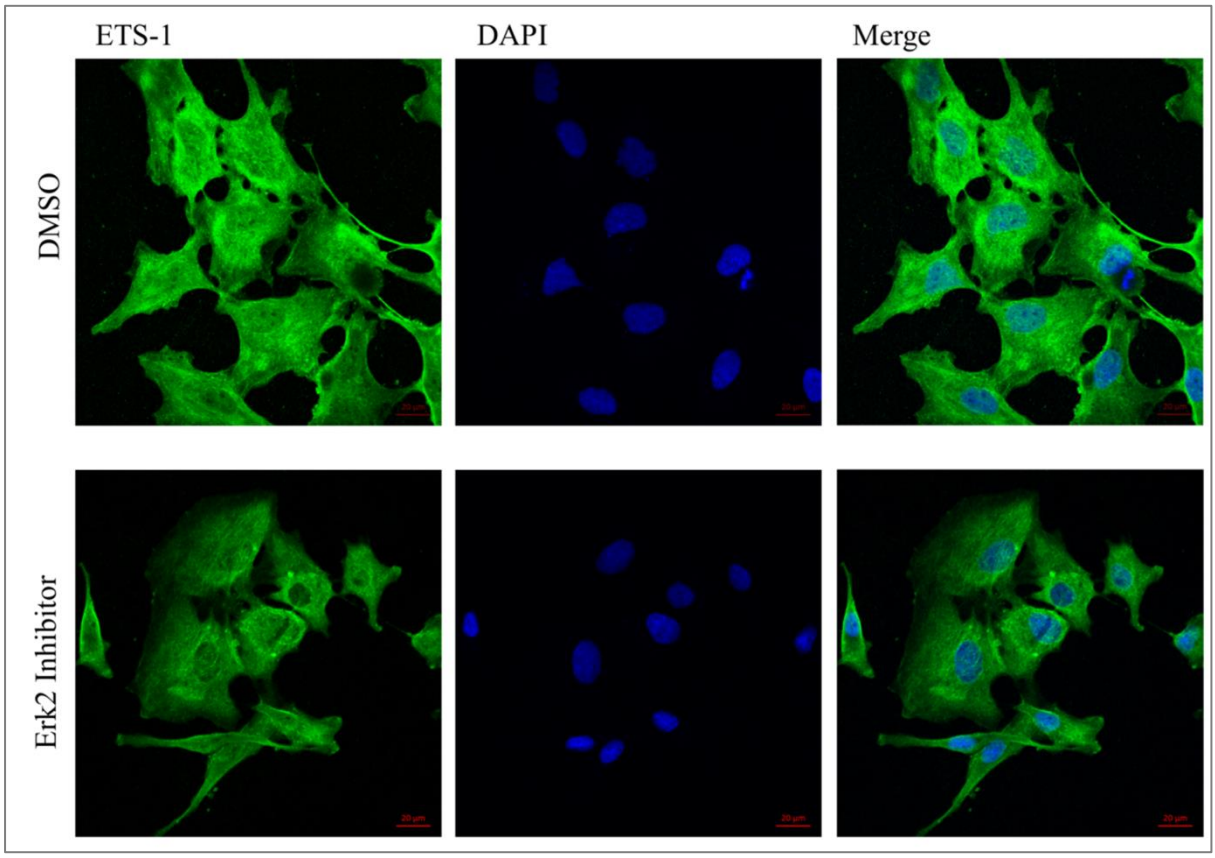


Figure E21A: The subcellular localisation of ETS-1 in Hs578T cells upon ERK2i treatment.

The subcellular localisation of ETS-1 in Hs578T cells was determined using indirect immunofluorescence staining and confocal microscopy (A). Hs578T cells were seeded overnight onto glass coverslips and treated with 60 µM ERK2i or dimethyl sulfoxide (DMSO) for 24 hours. Cells were incubated with ETS-1 primary antibodies (1:50 dilution) followed by donkey anti-rabbit-488 secondary antibody (green). The nuclei were stained with Hoechst-33342 (blue). Images captured using the Zeiss LSM 780 Confocal Microscope with a 100x objective and were analysed by Zen Lite 2012. The scale bars indicate a length of 20 µm. Images shown are representative of triplicate images captured from randomly selected fields, with the dominant morphology shown.

Appendix C: HOP Promoter Truncations

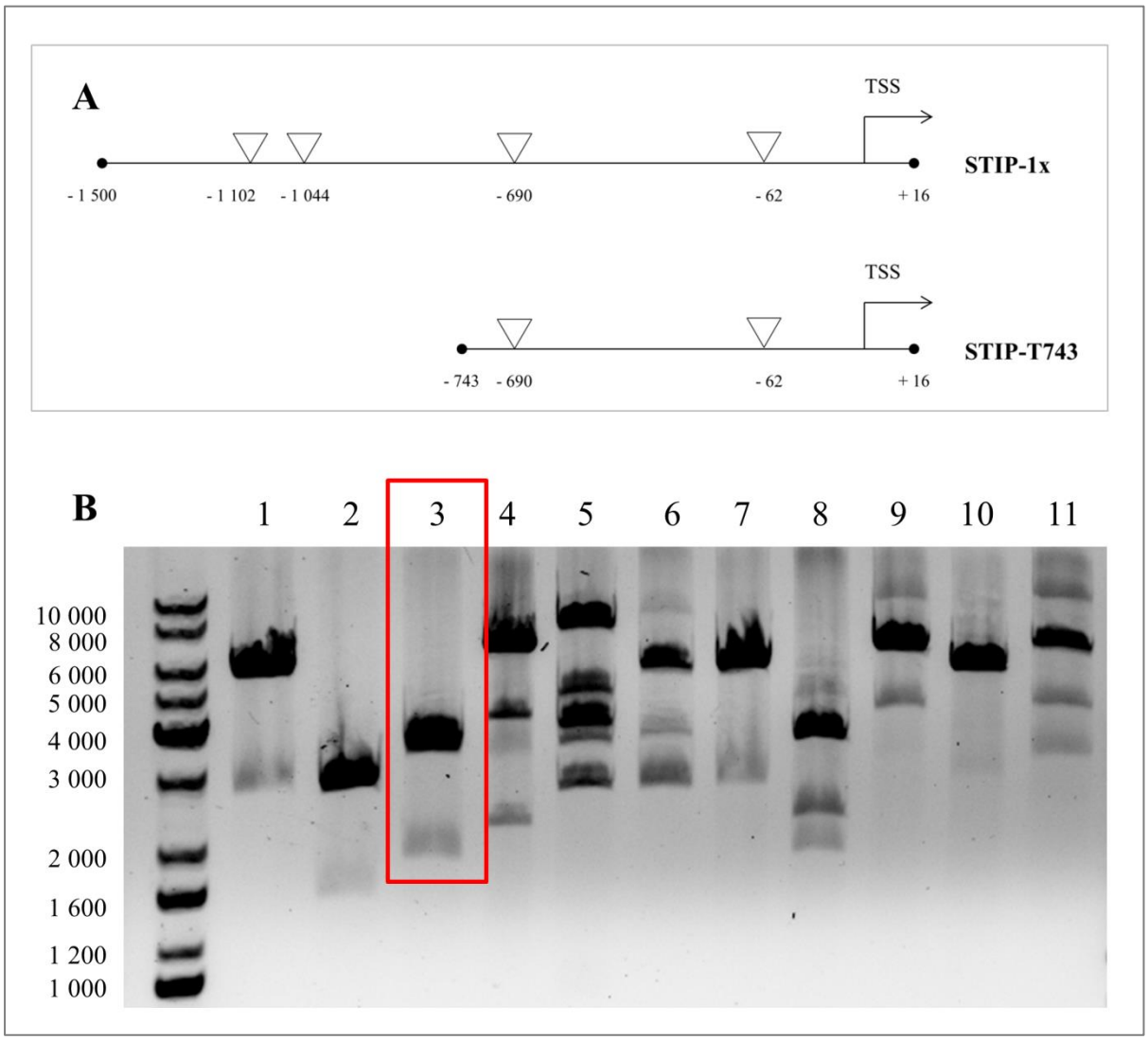


Figure E23: Cloning of the truncated HOP promoter.

A) Illustration of the design of the HOP promoter (STIP1-x) and the HOP promoter truncation (STIP1-T743). The triangles represent ETS-1 *cis*-elements along the promoter. Numbering of the promoter sequence is in relation to the transcription start site (TSS) of the HOP gene. B) Blunt end cloning of the STIP1-T743 into a pGL4 vector was unsuccessful. The only clone to contain the STIP1-T743 insert is highlighted in red, lane 3. However the direction of the insert is incorrect.