UNIVERSITY OF CALGARY

A Role of ING1 Proteins in Cell Senescence

by

Mohamed A. Soliman

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

CALGARY, ALBERTA

SEPTEMBER 2008

UNIVERSITY OF
CALGARY

Abstract

The ING1 proteins affect cell growth and apoptosis by modulating chromatin structure. The major splicing isoforms of the ING1 locus are ING1a and INGlb. While INGlb mediates an apoptotic response, the function of ING1a is currently unknown. Here, we show that the INGla:INGlb ratio is altered by several-fold in senescent, compared to low passage primary fibroblasts. Furthermore, we show that INGla induces many features of senescence.

To expand our understanding of ING1 functions, a cross-species *in silico* approach was used to identify potential human ING1 interacting proteins. Novel interactions, with p38MAPK and MEKK4, were identified and biochemically verified. None of the validated interactions were predicted by conventional tools tested. The Bioinformatics approach described can be used to predict novel interactions for other human proteins with yeast homolog(s). These data expand our knowledge of the roles that ING proteins play and demonstrate a novel role for ING1 proteins in differentially regulating senescence.

Acknowledgements

I would like to deeply thank my mentor Dr. Karl Riabowol for consistent support, guidance and encouragement. Thanks for his patience while my experiments were not working and his wisdom in dealing with both my frustrations and triumphs. Thanks for his welcoming character and very easy-going personality that encouraged me to overcome the cultural and language barriers that I encountered, in a relatively short time. He also let me discover significant questions to address in my research, and he respected my opinion and research findings. Thanks for his confidence in my scientific knowledge and for giving me the opportunity to improve myself in writing and in criticizing research work. Above all, I would like to thank him for the intellectual freedom he gave me, his trust in my capacity to think independently and accept my ideas even when we have different scientific opinions. Thanks for the great friendly journey!

I would like to thank Drs. Randal Johnston and Jerome B. Rattner for the wonderful discussions we have had about science, politics, and global aspects of life, while serving on my supervisory committee. Their open-minded personalities helped me think out the box, celebrate diversity of ideas, and accept positive criticism. Thanks for their great help and advice during my Master's thesis and their great support to continue my academic career the way I would like. I would also like to thank Dr. Greg Moorhead for serving on my thesis examination committee...much appreciated!

Thanks also to all the Riabowol Lab members, past and current, for the collegial environment and their friendship. Special thanks to Dr. Michael Russell for helping with my first steps in the lab, Ms. Annie Fong for teaching me tissue culture techniques, and Drs. Xiaolan Feng and Keiko Suzuki for helping me appreciate science and teaching me how it is essential to be both constantly well-driven and a hard worker to succeed in academia.

Thanks to Dr. Philip Berardi for establishing the first solid observations that led to my thesis project, Pinaki Bose and Sitar Shaw for revising my thesis writing. Thanks also to all the members of SACRI; I learned new techniques, shared ideas, used lab facilities, or borrowed reagents from almost every single lab in the cancer biology group, which was

greatly appreciated. Thanks to Elizabeth Long and Alexander Klimowicz for the fantastic long discussions about the future of science in the country, and to Drs. Arcellana-Panlilio and Wang in the Sothern Alberta Microarray Facility for helping me with numerous microarray experiments. I also appreciate the help I received from the Bioinformatics group, especially Paul Gordon for the great collaboration we had with them.

I would like to acknowledge the Alberta Heritage Foundation for Medical Research, the Alberta Cancer Board, and the Canadian Institute of Health Research for generous awards and studentships.

Finally, thanks to my family and my friends here and back home for their constant support and love. We shared frustrations, happiness and success moments. Thanks for understanding the time- and effort-demanding nature of my work. Without you, my success would not have been possible.

Table of Contents

List of Tables

List of Figures and Illustrations

List of Symbols, Abbreviations and Nomenclature

bp	base pair
cDNA	complementary deoxyribonucleic acid
Da	Dalton
g	gram
hrs	hours
kb	kilobase
kDa	kiloDalton
M	molar
mg	milligram
mins	minutes
mL	millilitre
mM	millimolar
µM	micromolar
oligo(dT)	oligo-deoxythymidillic acid
rpm	revolutions per minute
°C	degrees Celsius
FITC	Fluorescein Isothiocyanate

Techniques and Reagents

DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco' s Modified Eagle's Medium
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
IP	immunoprecipitation
PBS	phosphate buffered saline
PI	propidium iodide
RT-PCR	reverse transcriptase polymerase chain reaction
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Tris	tri-hydroxymethyl aminomethane

Proteins and biological molecules

ACF1	ATP-utilising chromatin assembly and remodelling factor
Arf	alternative reading frame
CBP/p300	histone acetyltransferase-associated transcription factor
CDK	cyclin-dependent kinase
G_0	gap 0 phase (quiescence)
G1	gap 1 phase
G2	gap 2 phase
Gadd45	growth arrest and DNA-damage-inducible
GADPH	glyceraldehyde phosphate dehydrogenase
HAT	histone acetyltransferase
HDAC	histone deacetylases

HDM2	human double minute-2
INK4a	inhibitor of kinase 4a
ISWI	ATP-dependent chromatin-remodelling complex
JMJD2A	jumonji domain-containing 2A
MAPK	mitogen-activated protein kinase
MAP3K4/MEKK4	mitogen activated protein kinase kinase kinase 4
MCMT	DNA (cytosine-5)-methyltransferase
NURF	nucleosome remodelling factor
PAF	PCNA-associated factor
PCNA	proliferating cell nuclear antigen
PIKK	phosphatidylinositol 3-kinase (PIK)-related protein kinase
PIP	phosphatidylinositol phosphate
PIP4Kβ	phosphatidylinositol 4-kinase type 2 beta
PtdIns	phosphatidylinositol
Rb	retinoblastoma
Sin3A	transcriptional regulatory protein and is part of HDAC1/2 complexes.
SV40	simian virus 40
WDR5	WD repeat domain 5, a member of the WD repeat protein family
XPG	Xeroderma pigmentosum complementation group G

CHAPTER 1: INTRODUCTION

Section I. The ING tumor suppressors

ING tumor suppressors; interacting partners and biological functions

The INhibitor of Growth (ING) proteins are a well-conserved family of type II tumour suppressors [1] that are frequently deregulated in different cancer types [2,3]. Phylogenetic analyses showed that ING family members are present in diverse organisms, including rats, frogs, fruit flies, worms, fungi and plants, *inter alia* [1]. Consistent with a role as a tumor suppressor, loss of ING1 (the first member of the ING family of proteins to be discovered) increases sensitivity to ionizing radiation in mice and in mouse-derived cells, and increases B-cell lymphoma occurrence in mouse knockout models [4,5]. ING1 was discovered using PCR-mediated subtractive hybridization of normal mammary epithelial cell cDNA against breast cancer cell lines, followed by screening of a senescent cell cDNA library and a biological screen to isolate genes down-regulated in cancers [6]. Shortly after its role as a tumour suppressor was observed [6], roles for ING1 in regulating apoptosis [7] and cell replicative life span [8] were reported, facilitated in part via interactions between ING proteins and p53 [9-13]. Although neither ING1 nor p53 now seem to be required for the activity of the other [5], many studies have reported synergism between them in inducing apoptosis.

In addition to interacting with p53 and associated proteins such as ARF [14], ING1b, the most highly expressed isoform of ING1, also interacts with PCNA [15] and with PCNA-binding partners such as GADD45 [16] and PAF [17]. ING1 and other ING proteins also associate with CBP/p300 [9,18,19] and we speculate that ING1b probably forms complexes with other PCNA-binding proteins including Fen1, MCMT and XPG, among others [20]. Since PCNA is critical for DNA replication and repair and interacts with DNA polymerase δ, it is not surprising that ING1b has been implicated in sensing [15] and responding to [16] DNA damage. Recently, two mechanisms have been identified that also contribute to ING's ability to regulate diverse biological processes: 1) the binding of bioactive signalling phospholipids to ING [21-23] and 2) the interaction of the ING PHD region with methylated histone tails [24-27]. Activation of ING1 and ING2 proteins by binding phospholipids probably regulates the roles of ING in apoptosis and the DNA damage response. Reading of the histone code and the subsequent contribution of ING proteins in modulating chromatin

structure by increasing (ING1b) or decreasing (ING1a) histone H3 and H4 acetylation levels [18] probably accounts for regulation of gene expression by ING. Involvement of the ING proteins in these major biological processes - apoptosis, the DNA damage response and senescence - has temporal overlap in different phases of the cell cycle as diagrammed in Figure 1. Together, the recent reports suggest the exciting possibility that the ING proteins transduce stress signals by binding ephemeral lipid-signalling molecules [28], localizing to chromatin and reading of the local histone code, subsequently contributing to epigenetic regulation through the targeting of (HAT) and (HDAC) activities.

The association of the ING proteins with methylated histones, as well as with HAT and HDAC complexes, also links the regulation of two major pathways - histone methylation and acetylation. This action of the ING provides an example of how the histone code could be read in a combinatorial fashion whereby ING proteins first recognize and bind the tri-methylated, or to a lesser extent di-methylated histone “hot spots”, and then recruit different HAT or HDAC complexes to them [29,30]. Recruitment of these complexes either activates or represses gene transcription according to the HAT or HDAC activity associated with the particular ING protein. Recruitment of HAT and HDAC complexes may also play a role in facilitating DNA repair through regulating chromatin structure [31,32]. Since INGs 1-5 [33], NURF [34,35], heterochromatin protein 1 [36,37] and polycomb proteins [38,39] (among others) bind similar regions of methylated histone H3, important questions remain to be investigated. What determines specificity of the different proteins for particular modified lysine residues on H3? What are the relative affinities of these proteins to methylated histone H3? Is subcellular localization a factor in determining binding specificity and affinity? Does interaction with p53, which has also been linked to HAT [40] and HDAC [41] activities, modify ING specificity of binding to methylated histone H3? Recent advances regarding ING family structure and function provide us with several clues that may be the key to understanding these critical aspects of ING protein biology.

Structural features of the ING family

Most members of the ING family of genes in vertebrates encode multiple splicing isoforms [42,43]. Bioinformatics analyses of all known ING protein isoforms revealed several conserved regions highlighted in Figure 2, most of which have been linked to specific functions [1]. Some of these regions are found in a subset of the INGs such as the PCNA-interacting-protein (PIP) motif (region 1 of Figure 2) found only in ING1b which mediates interaction with PCNA in a DNA damage-inducible manner [15]. The ING1b protein also contains a region of sequence homology to the bromodomain (PBD, region 2 in Figure 2), the function of which remains unknown. In a region comparable to that occupied by the PIP and PBD of ING1b, human ING2-5 proteins contain a leucine zipper-like (LZL) motif consisting of 4-5 conserved leucine or isoleucine residues spaced seven amino acids apart, which can form a hydrophobic face near the N-terminus. Although little is known regarding the function of this motif, the LZL was reported to affect ING2 function in DNA repair and apoptosis [44].

In contrast to the domains and motifs described above, all INGs from all species share a plant homeodomain (PHD) finger [29] that contains a C4-H-C3 zinc finger motif (four cysteines, one histidine and three cysteines; region 6 of Figure 2) located near the C-terminus. The PHD has recently been shown to interact specifically with methylated forms of histone H3, particularly those trimethylated on lysine 4 [45]. This domain is the most highly conserved feature of the INGs and is found in ~150 members of the human proteome [46]. The next best-conserved structural feature is found only in the ING proteins and is called the novel conserved region (NCR, now known as lamin-interaction domain or LID) [1]. The sequence KIQI/KVQL of the NCR/LID is particularly well conserved and may participate in binding to HAT or HDAC complexes since the N-terminal 125 amino acids of ING1b, which includes this region, binds to the SAP30 protein of the Sin3/HDAC complex [47]. Another highly conserved region is the nuclear localization signal (NLS) located upstream of the PHD finger region that is present in most INGs. This region contains three potential nucleolar targeting signals (NTS) found only in ING1 and ING2, two of which target ING1 to the

nucleoli in response to different stresses [48], and which interact with the importin proteins Karyopherin-α and -β for nuclear import [49]. A 14-3-3-binding motif lies in a short region between the PHD and NLS domains, centered upon serine 199 (S199). Upon phosphorylation of S199, 14-3-3 binds ING1b, resulting in relocalization of ING1b from the nucleus to the cytoplasm. This inhibits the ability of ING1b to induce expression of the CDK inhibitor p21 [50], consistent with the idea that ING1 subcellular localization regulates its biological functions [51]. A short region closely juxtaposed to the PHD called the polybasic region (PBR) is found in ING1 and ING2, which are closely related evolutionarily [1] and functionally [30,52]. Initially believed to bind the PHD of ING2 [21], phosphatidylinositol monophosphates (PIP; for lipid nomenclature [28]), including PtdIns3*P*, PtdIns4*P*, PtdIns5*P* were later reported to bind to the PBR found adjacent to the PHD fingers of different proteins including ING2 [22], predicting a two-stage activation in which PHD fingers interactions with histones occur downstream of PIP binding to the PBR. This mechanism is discussed further below in the light of recently reported differences in the binding of different ING family members to both phosphoinositides (PIs) and modified forms of core histones.

ING1 splicing isoforms

Several splice isoforms have been reported for different ING family members. The two predominant ING1 isoforms are ING1a, which is a product of exon 2 and exon 1b, and ING1b which is a product of exon2 and exon 1a. ING1b appears to be the highest expressed form among the different ING1 isoforms in cultured fibroblasts and epithelial cells, thus, the majority of previous studies have focused on the function of ING1b. Both ING1a and ING1b appear to have opposing functions and different cellular partners. For example, ING1b binds to PCNA through the PCNA Interacting Protein domain (PIP) in response to UV irradiation and induces apoptosis while ING1a does not show any interaction with PCNA under stress conditions [15]. In addition, microinjection of ING1b construct induces acetylation of both H3 and H4 histones while microinjection of ING1a inhibits the acetylation of both histones [18].

ING and p53: Do they need each other?

Initial studies in which ING1b was shown to be able to block the growth of cells with functional p53, but not cells in which p53 was inactivated by SV40 T-antigen, suggested that ING1b might functionally interact with p53 [8]. This idea was significantly expanded upon in a report indicating that ING1b physically associated with p53, that ING1b was required for p53-dependent expression of the p53 target gene p21, and that ING1b was essential for p53 activity and vice versa [10]. Functional synergy between the ING1, -2, -4 and -5 (but not ING3) proteins and p53 in inducing apoptosis and p53 target gene expression has subsequently been reported by a number of groups [9,11,13,53,54]. Although how ING proteins activate p53 is unknown, evidence exists supporting two mechanisms. First, ING1 and ING2 are able to enhance levels of p53 acetylation on lysine (K) 382 [9,13] and ING1 binds and inhibits hSir2, which is known to deacetylate this residue, hence inactivating p53 [55]. Acetylation of K382 is also known to increase during cell senescence [56], consistent with the increased activity of p53 and increased expression of INGs 1 and 2 in this process. An alternative mechanism is that ING1 may stabilize p53 through binding p53 and blocking access of MDM2, contributing to increased levels, and therefore activity of p53 in response to ING1 [53]

Although the mechanism by which ING proteins activate p53 is an important and controversial point that remains to be fully elucidated, a recent report in which murine ING1 and p53 knockout lines were used [5] suggests that in contrast to the previous report of an obligatory link between ING1 and p53 [10], these two proteins have fully independent functions. In this study, ING1 deletion increased cell proliferation in both p53 wild-type and p53-deficient fibroblasts. Moreover, ING1 deletion did not affect p53 dependent functions such as oncogene-induced senescence and growth arrest following DNA damage nor did it rescue p53-dependent embryonic lethality in MDM2-deficient mice. Furthermore, in contrast to the previous observations that ING1b plays a proapoptotic role following DNA damage in primary [15,57] and established cells [7,11,54], in murine ING1b knockout cells a robust increase in apoptosis following gamma irradiation was seen both *in vitro* and *in vivo* that correlated with increased expression of the pro-apoptotic factor Bax. How this fits with the

observation that the same ING1b knockout animals develop follicular B-cell lymphomas [4,5] is paradoxical. However, previous studies indicating a pro-apoptotic function for ING1b have largely relied upon overexpression models and human cancer cell lines. So, while normal levels of ING1b may protect normal mouse cells from apoptosis as seen in the knockout model, perhaps through a growth inhibitory activity that allows appropriate initiation of DNA repair as previously suggested by PCNA interaction mutants [15], overexpression beyond physiological ranges may unbalance pathways and result in the induction of apoptosis for reasons that are not linked to normal cell physiology. Determining the range of normal and cancer cell types in which ablation of ING1 induces apoptosis and the cell backgrounds and levels at which ING1 overexpression can induce cell death should prove informative in understanding the complex relationship that exists between the p53 and ING1 tumor suppressors.

A role for ING stress signalling

PIs have an important role in mediating a variety of biological processes including response to stress and, because they regulate essential cellular functions, PI metabolism is tightly regulated at the subcellular level [58]. Historically, PIs were thought to be synthesized and metabolized in the cytoplasm, however, a plethora of more recent data supports the notion that PIs also reside in the nucleus where they have essential roles in normal cell function [58]. Since the discovery of nuclear PIs in the 1970s, several reports have suggested their involvement in the regulation of gene expression, DNA repair and telomere maintenance, but the effectors of their involvement in these processes remained obscure. To identify receptors for nuclear PIs, a large-scale analysis was undertaken, in which 100,000 peptides representing most of the human proteome were screened. This screen identified ING2 as the most avid binder of PtdIns5*P* [21]. ING2 bound, in decreasing order of avidity, to PtdIns5*P*, PtdIns3*P*, PtdIns4*P*, PtdIns(4,5)P_2. The ING1 PHD also bound these three monophosphorylated species with high affinity. Production of PtdIns5*P* in cells resulted in relocalization of ING2 from the cytoplasm to the nucleus suggesting that binding of PtdIns5*P* induces interaction of ING2 with chromatin. This finding was later confirmed by studies

showing that stress signals increase nuclear PtdIns5*P*-ING2 interaction by a mechanism involving p38MAP kinase (mitogen-activated protein kinase) activation. Specifically, stress-activated p38MAP kinase was found to inhibit PIP4Kβ lipid kinase which functions to reduce levels of PtdIns5*P* [23]. Based upon structural modeling it appeared that the ING2-PtdIns5*P* interaction was mediated via the PHD domain of ING2 [21]. Consistent with this idea, mutations of the positively charged residues in the PHD domain abrogated the ING2-PtdIns5*P* interaction and subsequent activation of p53-dependent apoptosis after genetic insult. However, the six lysine and/or arginine amino acid residues that were mutated to demonstrate specificity of the interaction appear to reside primarily in the short polybasic region (PBR) directly adjacent to the PHD finger. Indeed, a more recent report showed that the PBR of the Pf1 protein, which also harbours a PHD finger, and the PBR of ING2 are both necessary and sufficient for binding PtdIns5*P* [22]. In this study, mutations of the zinc-coordinating cysteine residues or chelation of zinc had no apparent effect on PIP binding. In contrast, binding to PIP was lost when the polybasic region was deleted. PIP-PBR binding is also consistent with reports showing that the polybasic region of other proteins, such as N-WASP (neural Wiskott-Aldrich Syndrome protein), bind to PtdIns(4,5)P_2 resulting in increased actin polymerization [59]. Also consistent with PIs binding the PBR, but not the PHD finger of ING proteins, ING4 (which lacks a PBR but contains a consensus PHD) did not bind PIP [27]. Finally, because these two studies [21,22] reported different relative binding affinities of this bioactive lipids to tagged ING2 fragments, the tags themselves may influence affinity suggesting that the use of full-length proteins may provide additional insights into the biological consequences of ING-PIP binding. Given the lack of other experimental reports and the different lipid-binding assay systems used, the possibility that the PHD finger domain may act as a “cofactor” in helping the PBR bind to PIs cannot be fully excluded. Regardless of these details, the discovery of ING proteins as nuclear PI receptors provides an important mechanism for how the ING1 and ING2 proteins, the only ING proteins to possess PBR, act to regulate gene expression in response to stress. Examining the effects of PtdIns5*P* and PtdIns4*P* phosphatases on ING localization to chromatin and whether deletion of the PBR in ING1 and 2 alters stress-induced ING activity

should help to better define the functional differences between members of the ING family.

A role for ING in reading the histone code

Nucleosomes consist of 146 bp of DNA wrapped around an octamer of two basic H2A, two H2B, two H3 and two H4 histone proteins. The tails of the core histones have abundant lysine (K) residues that extend out of the core nucleosome. These K residues are the targets of numerous covalent modifications including acetylation, ubiquitination, methylation and sumoylation [60] and these modifications, as well as phosphorylation of other residues form what has been coined the “histone code”. The modified histone tails act as docking sites for effector protein complexes that, through associated enzymatic activity (ex. HAT, HDAC and others) are believed to ‘open’ or ‘close’ chromatin, largely based on the modification of histone charge and the ability to retain stable nucleosome structure. This alternation of chromatin structure regulates the transcription of specific genes according to the type of enzymatic activity bound and the nature of the local gene promoters. Certain histone modifications or ‘marks’ have been associated with transcriptional activation and some have been associated with transcriptional repression and the formation of heterochromatin, particularly on histones H3 and H4. Methylation of different K residues leads to distinct biological effects, and mono-, di- or tri-methylation patterns on the same K can lead to distinct functional outcomes. Together, these findings have led to the idea of combinatorial modifications on histone residues providing a histone code that can be interpreted in distinct ways in different cell types [61]. For example, binding by chromodomains (CHD) and regulation of transcriptional activation were recently reported [62,63] in which the CHD, through recognition and binding to methylated H3K4, recruited histone acetyltransferases (HAT) to chromatin where they acetylate specific lysine residues and open local chromatin structure, promoting transcriptional activation.

PHD fingers are a form of zinc finger found primarily in nuclear proteins that frequently contain an NLS. The canonical PHD finger of ~60 amino acids is found in all eukaryotes, has a conserved C4HC3 motif and binds two zinc atoms [46]. PHD fingers have long been suspected to interact with histones, for example the ACF1 PHD finger protein, a

subunit of the ISWI nucleosomal remodeling complex, interacts with nucleosomes via the central domains of core histones, promoting its binding to nucleosomal substrate [64]. PHD also act in a combinatorial fashion with the CHD domain to recruit HDAC complexes to the methylated H3K36 marker [65]. As detailed below, the ING PHD finger proteins have now joined the small but growing group of proteins currently implicated as interpreters of the histone code.

It is clear that stress-induced PIP binding activates ING1 and ING2, and that ING proteins are found in a number of HAT and HDAC complexes [29,30], but exactly how the activated ING proteins transduce stress signals has not been elucidated. One major transduction mechanism was clarified by several reports showing that the PHD finger domains of the ING2 and NURF proteins bound to the histone mark H3K4 in a methylation-sensitive manner [24,25,34,35]. Subsequently, a proteome-wide analysis in yeast identified additional proteins with PHD fingers which can selectively bind H3K4me3 and H3K36me3 (tri-methylated) [66]. The PHD regions of ING1-5 were all able to bind to H3K4Me3 peptides with dissociation constants (K_{ds}) ranging from 1.5 to 7.9 μM (K_{ds} for H3K9Me3 and H3K20Me3 were in the mM range, suggesting that they were not targets) [25] and the PHD finger from ING4 bound, but not in a methylation-sensitive manner [27]. All human and yeast ING proteins tested bound to trimethylated H3K4. In addition, the tri-methylated form was bound 10 times more avidly by ING2 than the di-methylated form, and 100 times more avidly than mono-methylated H3K4 [24,25]. Linking these two observations together (activated ING proteins binding to methylated H3K4 through their PHD fingers and ING proteins also binding specific HAT and HDAC complexes), suggests that ING can recruit HAT or HDAC activity to specific chromatin locales marked by H3K4. Indeed, yeast Yng1 can interact with H3K4me3 and recruit the NuA3-HAT complex to acetylate H3K14. This sequential tri-methylated histone binding and recruitment of HAT complexes to a specific genetic locale affected the transcription of certain genes [67]. Figure 3 shows a diagram linking the process of ING activation to interaction with methylated histones, altering adjacent histone acetylation status with associated HAT or HDAC complexes. This idea has been confirmed in yeast and mammalian cell systems where the Yng1 protein was shown to

recruit NuA4 to chromatin through binding to methylated H3K4 [26] and in which ING2-H3K4Me binding stimulated ING2-associated HDAC activity [24]. These studies make several points. First, they show for the first time that H3K4Me, previously considered to be a mark for gene activation, is not associated with a specific transcriptional outcome. Rather, it may be associated with either gene activation or repression according to the HAT or HDAC protein complex that can recognize the histone mark. Other histone methylation marks have also been reported to have different effects on transcription, for example, H3K9Me3, which has long been thought of as a heterochromatic marker was recently reported to be associated with transcriptional activation [68]. These different outcomes probably result from histone-binding proteins being associated with different HAT and HDAC complexes. The situation is considerably more complex given that several other proteins have been shown to recognize H3K4 methylation [69] such as CHD1 [63], NURF [34], JMJD2A demethylase [70] and WDR5 [71], and given that ING1 may also serve a role in the maintenance of pericentric heterochromatin through interaction with both DNA methyltrasferase complex and HDAC1/2 complex [72].

Although recent reports regarding the ING protein PHD fingers and binding of PIP by INGs 1 and 2 help to explain how some of the ING proteins exert their biological effects, a number of details remain to be resolved. For example, does p53, which interacts with acetylation complexes, affect ING1 function (and *vice versa*) through altering the histone code, being stabilized by ING1, or both? What are the relative affinities of intact ING proteins for binding to histone variants, as compared with those of the more commonly used peptides? Does lipid binding increase the affinity of ING1 and 2 proteins for histones and does histone binding lead to PIP displacement from the ING? Since the PIPs bind chromatin [73], do ING1 and 2 bind PIP and histones simultaneously or cooperatively? ING1 and 2 associate with both HAT [9,18] and HDAC [30,47] complexes and ING3 has been reported to be a subunit of the Tip60 HAT complex [30]. ING4 was detected in an HBO1 (histone acetyltransferase bound to ORC 1) HAT complex whereas ING5 appears to be part of two HAT complexes, HBO1 and MOZ-MORF (monocytic leukemia zinc finger protein-related factor) [30]. Despite the identification of these complexes, it is still unknown how

incorporation of ING proteins into different HAT-HDAC complexes is regulated. One fundamental question that remains to be addressed directly is whether ING proteins are targeted to histones to increase local HAT or HDAC activity by simultaneous binding to histones and HAT or HDAC complexes. One model of how this might occur based upon existing reports is shown in Figure 4. If binding to HAT or HDAC complexes occurs through the amino termini of the ING proteins that contain a conserved region [1] as previously reported for ING1-Sap30-HDAC1 complexes [47] then this model is plausible since the PHD finger region would be free to interact with histones. Variations on this "trimolecular" histone-ING-HAT or HDAC theory could include additional components for determining specificity. For example, ING could first be recruited by binding to specific transcription factors, lipid-signalling molecules or nuclear matrix components, and then this complex could be further stabilized by recognition of methylated H3K4 ensuring a reasonable 'residence time' for the complex to exert a biological effect [74]. Indeed, ING1 has been shown to repress AFP (alpha-fetoprotein) gene transcription through direct binding to A-T motifs, excluding HNF1 (hepatocyte nuclear factor 1) binding while targeting HAT activity to AFP promoter regions [13]. This binding may help recruit ING to chromatin, in the same vicinity as H3K4Me3 marks. Also consistent with this idea is that ING2 enrichment in chromatin and the accompanied biological outcome following DNA damage was dependent on binding to the lipid signalling molecule PtdIns5*P* [21], which should be able to occur independently of histone binding based upon binding sites on ING proteins (Figure 2). These reports [65,67] support the model shown in Figure 4, in which the stepwise recruitment of ING to specific gene loci via activation by PIP and binding of transcription factors and possibly other nuclear factors, occur upstream of the ING proteins recognizing and binding H3K4 and localizing HAT or HDAC activity to particular chromatin locales.

Figure 1: ING interactions through the cell cycle.

Once activated by the binding of phosphoinositides, and in particular phosphatidylinositol monophosphates (PIP), the ING1 protein has at least three major binding targets, which in the context of binding ING proteins, seem to primarily affect three biological functions. These are: (i) PCNA, involved in DNA replication and repair and apoptosis; (ii) p53, involved mainly with apoptosis and senescence; and (iii) histone H3 methylated on lysine (K) 4 (H3K4), which affects gene expression, and thus senescence and apoptosis. In the case of PCNA and histone H3, direct interactions have been demonstrated by point mutational studies and/or X-ray crystallography, whereas interaction with p53 might be direct or through associated proteins, such as ARF, MDM2 or others. Stress-inducible interactions with PCNA are thought to have their greatest consequences in S-phase when PCNA is predominantly nuclear and acts in DNA replication and repair. PCNA forms homotrimers and interacts with core replication and repair proteins, methylases, acetylases, GADD45 and p21, among others, perhaps explaining reports of ING interactions with these proteins. ING interactions with the p53 pathway have broader cell-cycle specificity during G1, S, and G2 phases and generally are reported in the context of apoptosis, whereas effects upon chromatin structure through histone H3 interaction are likely to be greatest during G1 and G2, and show a lesser effect during DNA synthesis. Therefore, effects on apoptosis and senescence are not limited to specific points in the cell cycle, and it should be noted that the three interacting proteins (p53, PCNA and H3K4) probably contribute to all of the three major biological functions to varying degrees. Although all ING proteins can interact with H3K4, only some are sensitive to the degree of histone methylation and only ING1 and ING2 seem to be bound and activated by PIP.

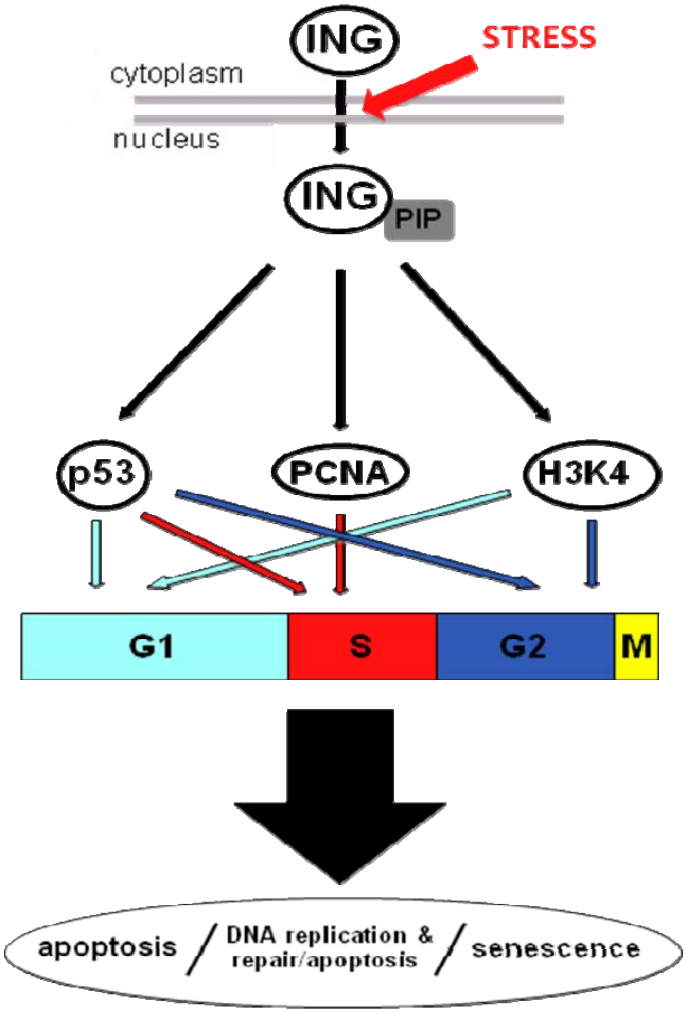


Figure 1: ING interactions through the cell cycle

Figure 2: ING1 protein domains and their interacting partners

A. ING protein domains. (i) The PCNA-interacting protein motif of ING1b binds specifically to PCNA, in competition with p21 (and potentially with GADD45 and other proteins in response to UV) and promotes ING1b-mediated apoptosis. This region is present only in ING1b. (ii) The region that shows sequence homology to bromodomains called a partial bromodomain (PBD) was identified by bioinformatics analysis, binds Sap30 of the mSin3–HDAC1 complex which might target HDAC, and possibly HAT activity for some ING proteins. (iii) The Lamin Interaction Domain (LID) identified by bioinformatics analyses is conserved in all ING proteins, is the second most highly conserved region in the ING protein family. (iv) The nuclear localisation signal (NLS) is conserved in most ING proteins and targets them to the nucleus through binding of the karyopherin-α and β transporter proteins. Nucleolar translocation sequences (NTS) are found within the NLS and target ING1 to the nucleolus under conditions of stress. (v) The ING1 proteins contain a 14-3-3 recognition motif to which 14-3-3 binds when ING1 is phosphorylated on serine 199. Binding promotes translocation of ING1 from the nucleus into the cytoplasm. (vi) The PHD finger found in the ING proteins is the most highly conserved region in the ING family in all species examined, from yeast to humans, and has been shown to bind to core histone H3K4 in a methylation-sensitive manner. Binding then promotes acetylation of nearby lysine residues through regulating HAT and/or HDAC activity, which can then alter transcription at specific genetic loci. (vii) The poly basic region (PBR) of ING1 and ING2 is both necessary and sufficient to mediate interaction with phosphoinositides and activate them. Activation might promote subcellular relocalisation and interaction with proteins and protein complexes.
B. Representation of the known ING proteins and their structural features. Numbers indicate the boundaries of the sequence motifs and domains in ING1b are drawn approximately to scale. In this diagram, PIP represents the PCNA-interacting protein motif, PBD is a sequence with partial homology to bromodomains, LZL is a leucine-zipper-like region, the LID is the lamin interaction domain, NLS represents the nuclear localisation

sequence, the NTS in p28ING4 represents a truncated NLS that retains a nucleolar translocation sequence but is not effective in localising proteins to the nucleus, PHD represents the highly conserved plant homeodomain finger and PBR is a polybasic region.

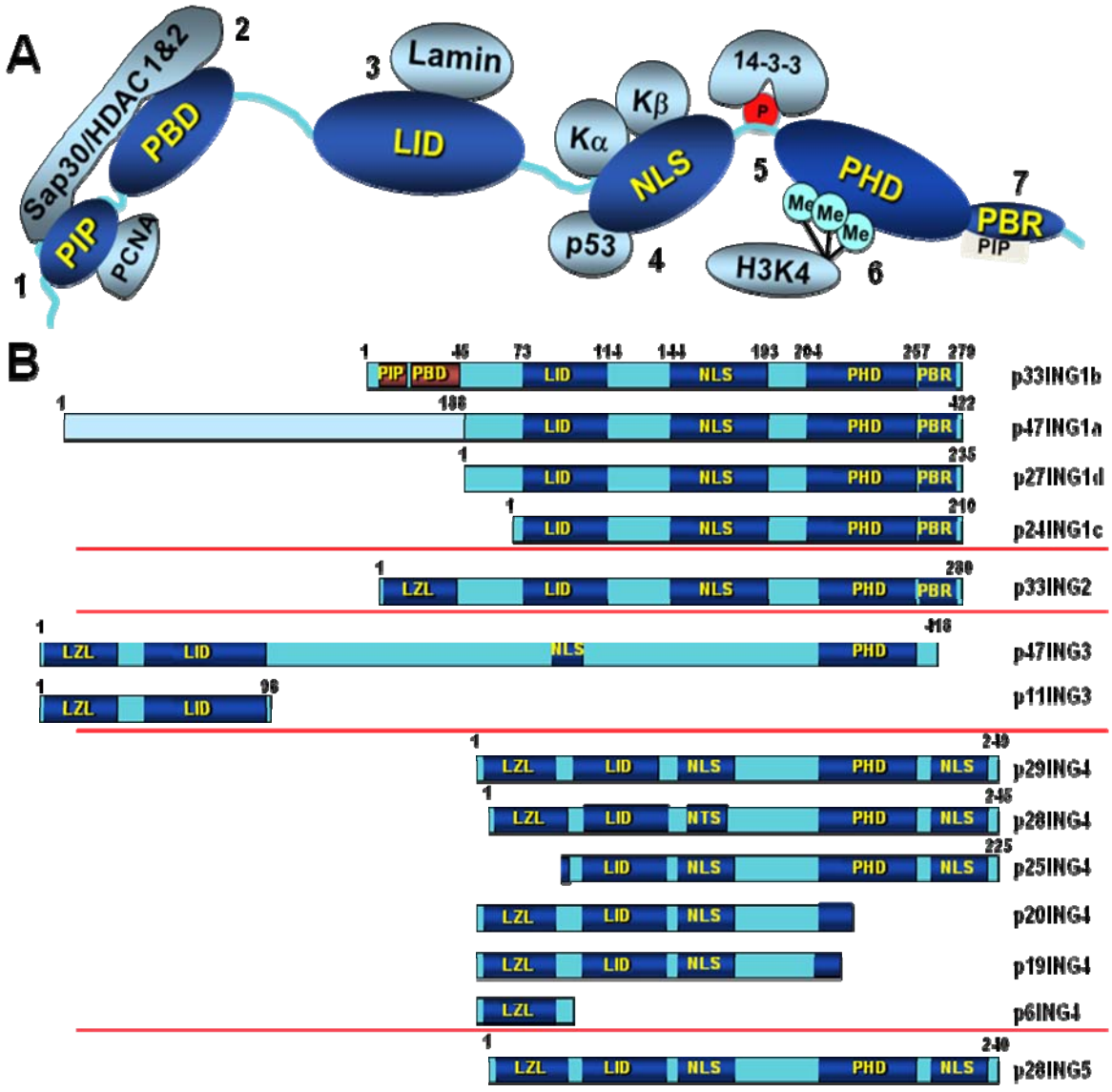


Figure 2: ING protein domains, interacting partners, and structural features

Figure 3: **Activating, interpreting and cycling of ING family proteins.**

(i) Activating: on exposure to cell stress (intra- or extracellular), p38MAPK is activated by upstream kinases. It then shuttles to the nucleus to phosphorylate PIP4Kβ, inhibiting its kinase activity, leading to increased nuclear PIP levels. PIP binds to ING1 and ING2, recruiting them to the chromatin, together with ING-associated HAT or HDAC complexes. Whether ING3–5 proteins are activated, and the mechanism for activation in response to some stresses are currently unclear. Whether PIP dissociate from ING proteins once they are recruited to the chromatin is also currently unknown but the regions of the ING protein that bind PIP and histones are closely juxtaposed as noted in Figure 2 (ii) Interpreting: ING binds specifically to H3K4me2/3 recruiting the ING-associated chromatin-remodelling complexes to specific chromatin regions and possibly specific genetic loci. This leads to distinct histone post-translational modifications, such as acetylation and deacetylation, leading to different biological outcomes such as apoptosis, cell cycle arrest or DNA repair. (iii) Cycling: once the ING-mediated changes in chromatin structure and subsequent transcriptional activation or repression are realized, ING proteins are thought to be removed and/or inactivated. One mechanism that might function in this role is the phosphorylation-dependent binding of ING proteins to 14-3-3 proteins, which leads to transport of ING from the nucleus to the cytoplasm. Dephosphorylation might lead to release from 14-3-3 and binding to karyopherin proteins α and β, which shuttle ING proteins back to the nucleus.

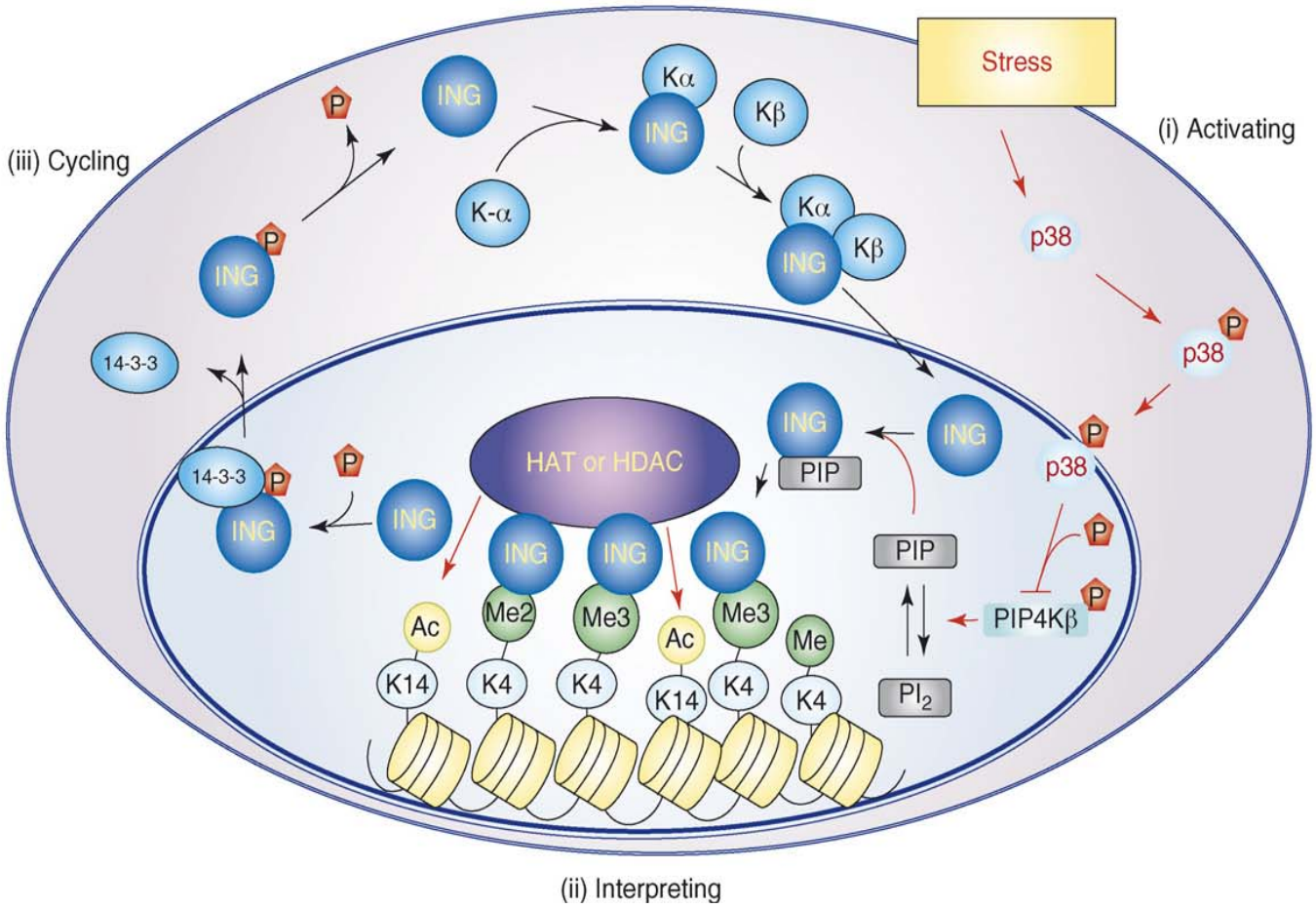


Figure 3: **Activating, interpreting and cycling of ING family proteins**

Figure 4: **A model for interpreting and translating the histone code.**

ING1 and ING2 are activated by binding of PIP, and all ING proteins might bind transcription factors including the ones shown, which have been reported to bind to subsets of ING. Interactions between ING proteins and other proteins might occur via PHD domains because one function of PHD is to promote protein interactions. These interactions between ING proteins and other transcriptional regulators might promote binding to methylated histone tails via the ING PHD finger regions through increased local concentration. The amino termini of ING are thought to bind to different members of major HAT or HDAC complexes. For example, Sap30 binding to ING1 would be expected to target HDAC1 and reduce local acetylation state. Other proteins that have been reported to bind to H3K4Me, such as CHD1, JMJD2A, and others, could compete for binding, antagonize, or collaborate with, the ING in modifying the state of chromatin acetylation. Only a subset of the known HAT or HDAC complexes with which ING proteins have been reported to interact are shown. Direct interaction with a member of a HAT or HDAC complex has only been shown for ING1 binding to Sap30. ING2 has not been shown to bind Sap30 although it associates with the mSin3a complex. Occupancy of complexes might occur with a single ING protein or more than one, and it is unclear if different ING proteins have the same or different roles when in these complexes, and how they are recruited to the complexes.

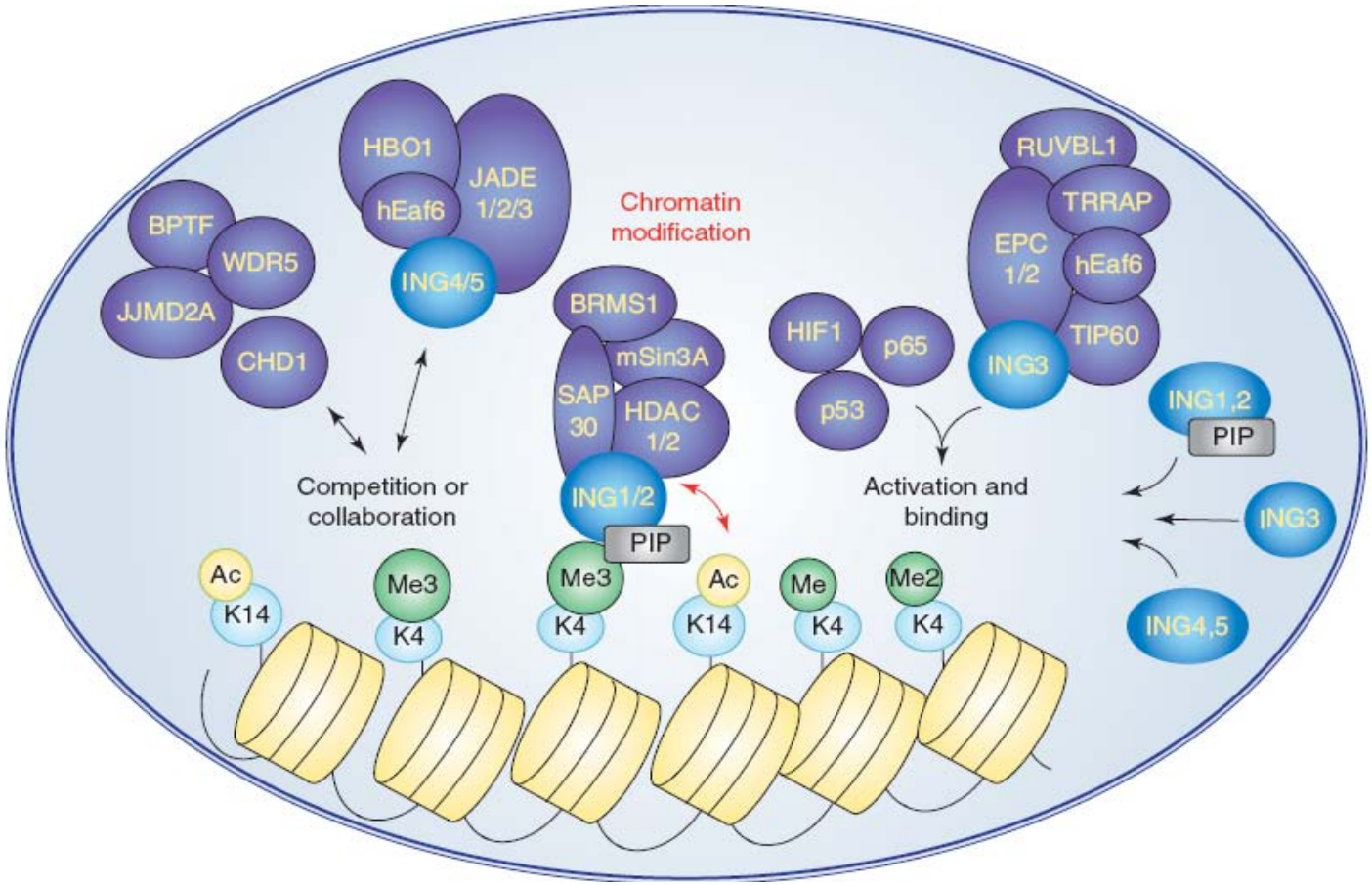


Figure 4: **A model for ING proteins interpreting and translating the histone code**

Section II. Cell senescence

Cell Senescence

Although the idea of cellular senescence had been discussed for years, Leonard Hayflick was the first to provide evidence that it prevents normal human fibroblasts from growing indefinitely in culture [75]. The cessation in growth seen as a function of population doublings was termed replicative senescence. At a certain proliferation limit, all normal diploid cells lose their capacity to divide. Non-dividing cells are viable for weeks, but are not able to grow despite the presence of nutrients and growth factors in the medium. Senescence, therefore, involves irreversible cell cycle arrest, although cells maintain metabolic activity and are characterized by changes in morphology and gene expression that distinguishes them from quiescent cells which are in a reversible form of cell cycle arrest.

Recently, it has become clearer that senescence is related to cancer in that cancer cells frequently escape the process of senescence to become immortal. Two main ideas have emerged that further link the fields of aging and cancer: 1) cell proliferation increases the risk of cancer development in multicellular organisms. As a defence mechanism, apoptosis (programmed cell death) and cell senescence may act to protect the organism from the effects of cell oncogenic transformation. Whereas apoptosis kills the cells, senescence irreversibly arrests cellular growth preventing malignancy, 2) senescence is most likely to be the *in vitro* form of *in vivo* aging, although the link between aging and cell senescence has not been fully established. This linkage implicates, however, that senescence may deplete tissues of proliferating or stem cells through senescence affecting the normal tissue renewal process and participating in age-associated degeneration.

Causes of aging and senescence

Although initial studies focused on investigating why normal human fibroblasts cannot grow continuously in tissue culture, subsequent studies showed that senescence can be induced by external and internal stimuli [76], with internal emanating from the chromosomal ends or telomeres.

Telomeres are stretches of hexanucleotide repeats (5'-TTAGGG-3' in vertebrates) and associated proteins that are located at the ends of chromosomes to protect them from being recognized as double strand breaks (DSBs) and, subsequent fusion with other chromosomes ends by DNA-repair processes [77]. DNA polymerases cannot completely replicate DNA ends, a phenomenon known as the end-replication problem. Accordingly, cells lose between 50 to 200 base pairs of telomeric DNA during each round of DNA replication [78]. Accumulated cell divisions render telomeres critically short and dysfunctional in normal human somatic cells. In contrast, in germ line and cancer cells, telomeres are maintained by the enzyme telomerase. Most somatic tissues totally lack telomerase, or it is expressed only transiently at levels that are insufficient to prevent telomere shortening. In cultured human fibroblasts, overexpression of telomerase was not only sufficient to prevent telomere shortening but also resulted in cell immortalization [79]. Some cells that escape senescence and continue to divide enter a state of **crisis**, a condition characterized by gross genomic instability and cell death. Cells that can avoid crisis and reactivate telomerase become immortalized. Thus, telomere erosion may play opposing roles in cancer development. On one hand, shorter telomeres induce cell senescence, blocking the growth of transformed cells. On the other hand, shortened telomeres may induce genomic instability which may then promote the development of cancer [80]. The first observation suggesting that telomere shortening is the major mechanism responsible for replicative senescence came in early 1990s [81] and it is currently the only known endogenous mechanism to induce senescence.

In addition to telomere attrition, cellular senescence can be induced by other multiple extrinsic factors such as DNA damaging agents [82], oxidizing agents [83], over expression of certain oncogenes [84] or tumour suppressors [85]. This ‘extrinsic form’ of induction of senescence occurs much more rapidly than that induced by telomere attrition and occurs while cells still have functional telomeres. This led to the distinction in the literature between ‘replicative senescence’ which refers to senescence due to population doublings and that is most likely induced by telomere erosion and ‘stress-induced premature senescence’ (SIPS), where senescence is induced more rapidly by exogenous factors other than loss of telomere segments [86].

Tumor suppressors and senescence: a coherent link

Tumor suppressors can be classified into two major categories: caretakers and gatekeepers based on the way they are believed act to prevent cancer [87]. Caretakers are envisioned to act by preventing replication and/or repairing DNA. In this mode, they serve as a defense line against oncogenic mutations and help maintain genome integrity. As a consequence, they can delay the onset of aging phenotypes caused by genome instability. In contrast, gatekeepers are thought to act by getting rid of potential cancer cells, i.e. cells that are at risk of neoplastic transformation, via induction of apoptosis or cell cycle arrest. The arrest may be transient to allow DNA damage repair to occur or stable arrest that might lead to cellular senescence [88]. It is presumed that gatekeepers function when internally or externally induced DNA damage cannot be repaired. Yet, they may do that at the expense of depleting tissues of proliferating or stem cells through induction of apoptosis or senescence affecting the normal tissue renewal process. In this context, they act antagonistically to the normal role of tumour suppressors in promoting longevity, demonstrating an antagonistic pleiotrophic mechanism of action. Antagonistically pleiotropic genes or processes are those that benefit organisms early in life (e.g. by suppressing cancer) but are detrimental later in life (e.g. by compromising tissue function).

Signalling pathways mediating senescence

Consistent with their roles in preventing cancer, tumour suppressors appear to act as crucial mediators of cellular senescence. Although there are several stimuli that can induce senescence, they all appear to ultimately rely on two gatekeeper tumour suppressor pathways, p53-p21 and p16-pRb [89]. Not surprisingly, the suppression of both the p53 and pRb pathways with oncogenes such as SV40 Large T-antigen, human papillomavirus E6 and E7 proteins or adenovirus E1A, together with telomerase activation, can inhibit cell cycle arrest and allow cells to bypass senescence [90]. Although these pathways can interact, they appear to respond to different stimuli and show cell-type-specific and species-specific differences in mediating senescence.

A. The ARF-p53-p21 pathway: Due to its importance as a genomic guardian and protectant against cancer progression, many studies have investigated the role of p53 in inducing senescence as a barrier to tumour progression. Stimuli inducing a DNA damage response, such as ionizing radiation, or shortened telomeres can induce senescence mainly through activation of the p53 pathway [91]. It is now well established that p53 activity increases in senescent cells [91,92]. However, loss of p53 delays, but does not prevent, the onset of replicative senescence in human fibroblasts [93]. The increased activity of p53 in senescent cells is attributed to post-translational modifications, mainly phosphorylation and acetylation. p53 is phosphorylated at serine 15 (S15) and acetylated at Lysine 382 (K382) during replicative senescence. Acetylation of p53 at K382 has been reported to be induced by ING2 [9]**,** among other factors. This pathway is regulated at different levels by different proteins. For example, the E3 ubiquitin-protein ligase MDM2 (HDM2 in humans) can mediate the proteosomal degradation of p53 and, hence, can inhibit p53-initiated senescence. One of the most interesting p53 upstream regulators is ARF (p19) in mice, the murine homologue of p14 in humans and a product of the INK4a locus. ARF is able to bind to and inhibit Mdm2, thus increasing p53 stability [94]. Murine fibroblats lacking p19ARF do not undergo senescence, indicating a crucial role of p19 in murine senescence [95]. p14 is not upregulated in senescent human keratinocytes [96] but its overexpression can induce a senescence phenotype in human fibroblasts [97], arguing that the p19-p53-p21 pathway may act in a species- and cell-type specific manner. Since the CDK inhibitor p21 is one of the most important downstream targets for p53-mediating cell cycle arrest, several studies have been conducted to determine the role of p21 in mediating senescence. As expected, p21 protein levels are increased in p53-induced senescence [98] and are also detected in screens for senescence-inducing genes [99].

B. The p16-pRb pathway: Another product of the INK4a locus is the p16 cyclin-dependent kinase inhibitor (CKI). It inhibits the CDK4-6/Cyclin D complex which mainly targets the retinoblastoma tumour suppressor protein and keeps it in the hypophosphorylated active form which can associate with E2F transcription factors. This association results in the inhibition

of transcription of cell cycle regulators such as Cyclin A and E. If released from its association with E2F, pRb prevents transcription initiation and cell cycle progression [100]. Like p21, p16 has been reported to be upregulated in cells undergoing replicative senescence [101,102]. Moreover, p16 inactivation results in life span extension of human fibroblasts and it is very frequently inactivated in immortal cell lines [103]. The mechanism behind upregulating INK4a protein expression may involve reduced expression of Polycomb proteins, such as BMI1 [104] and CBX7 [105]. Indeed, overexpression of either of these proteins extend lifespan of human or murine fibroblasts [106,107]. The mechanism through which p16-pRb mediates senescence may involve chromatin remodelling and formation of regions of heterochromatin known as senescence-associated heterochromatic foci (SAHF). These foci are believed to be involved in the permanent silencing of genes with crucial roles in cell proliferation such as E2F target genes [108]**.** Each SAHF contains portions of a single condensed chromosome, which is depleted for the linker histone H1 and enriched for HP1 proteins [109]. It has been reported that the p16-pRb pathway is crucial for generating SAHF and silencing proliferation-associated genes [108]. This finding established pRb as a crucial effecter of the senescence pathway and as a key component in the maintenance of irreversible cell cycle arrest.

C. Interplay between the two pathways in regulating senescence: Although not equivalent, both p16 and p21 functions as CDK inhibitors. Since pRb is known to be a target of CDKs, they can, hence, inhibit pRb phosphorylation and keep it in the active hypophosphorylated form. By doing this, p21 and p16 proteins can prevent E2F from transcribing genes that initiate proliferation since active pRb functions to inhibit E2F. For example, after DNA damage, the p53-p21 pathway is activated first followed by a secondary activation to the p16-pRb pathway [110]. Moreover, when p16-pRb activity is lost, p21 and p53 activities are upregulated [111]. Although the two pathways seem to be mutually regulated, cells respond differently when one of the two pathways is compromised. Cells that depend mainly on p53-p21 pathway to senesce can resume growth for some population doublings after p53 inactivation, albeit for fewer population doublings [93]. However, for

cells that senesce primarily through the p16-pRb pathway, inactivation of p53, p16 or pRb will not allow continued growth [111]. These observations have been interpreted to support two conclusions; first, the p16-pRb pathway is essential for induction of heterochromatinization by a poorly understood mechanism. Heterochromatinization is required for the perpetuation of the senescence state even when the senescence-inducing signals are removed. Second, the signalling pathways mediating senescence are cell-type specific to a degree, although pRb and p53 appear to be the main regulators of senescence in a majority of cell types.

Oncogene-induced senescence (OIS)

Oncogenes are mutant activated forms of normal genes that act in a dominant manner to transform cells. When an oncogene is activated in early tumourogenic stages, it produces abnormal signals which are interpreted as growth signals in some contexts, but normal cells frequently interpret these signals as a stress and respond by inducing apoptosis or senescence [112]. Mutations in senescence initiation pathways facilitate the establishment of unopposed-oncogene induced malignancy. This paradoxical role of oncogenes is known as oncogene-induced senescence (OIS). This phenomenon was first reported by Scott Lowe and colleagues when they overexpressed an oncogenic form of Ras in human fibroblasts and, to their surprise, the cells stopped growing and adopted a senescent phenotype [56]. Subsequently, activated forms of other members of the Ras pathway, such as Raf [113] and MEK [114], have been shown to induce similar phenotypes when overexpressed in normal diploid human cells. The idea of induction of senescence or tumour progression by the same agent sounds incompatible as one pathway halts the cell cycle while the other supports it progression. A model has been proposed to describe this apparent conundrum; most cells in the premalignant state undergo apoptosis or senescence in response to oncogenic stimulation, while a small fraction has the ability to continue growing. The balance between both groups of cells may determine the final fate of premalignant lesions [115]. Both actions of mutant Ras, for example, as oncogene and senescence-inducer can be reconciled if we consider that Ras transformation ability is only apparent in cells where some senescence signals are

malfunctioning. This is supported by the observation that cells defective in p19 alone or both p16 and p19 that are important upstream regulators of p53 and pRb, are transformed by Ras without evidence of other genetic changes that would be suspected to induce that transformation [56]. Moreover, inactivation of both the p53 and pRb senescence pathways, and overexpression of two oncogenes, (the simian virus 40 large-T oncoprotein and an oncogenic allele of H-Ras) together with the telomerase catalytic subunit (hTERT) result in direct tumourigenic conversion of normal human epithelial cells and fibroblasts to malignant cells [90]. This indicates that oncogene-induced transformation is only seen in an appropriate background of immortalization factors and it further expands upon a well-established concept in cancer research; multiple factors are involved in the generation of tumor cells. Although not initiated by telomere shortening, OIS has been proposed to use similar senescence signalling machinery since overexpression of Ras can induce p16 expression and drives the formation of SAHF [108]. These observations suggest the interesting idea that oncogenes, which usually thought to induce tumors, can induce senescence to limit tumorigenesis *in vivo.* Strong mitogenic stimulation by oncogene activation [116,117] or altered levels of tumor suppressors [118] have been shown to stop tumor growth at a benign stage predominantly by inducing senescence.

Senescence markers

Recent studies found that oncogene-induced senescence (OIS) acts at early stages of tumour formation in both mouse and human models to suppress tumour development [84]. This notion is strongly supported by the fact that senescence, at least by the current definition, occurs only in mitotic tissues (i.e. tissues that can proliferate and hence are more susceptible to malignancy) but not in postmitotic differentiated cells that have lost their proliferation capacity and are blocked from re-entering the cell cycle, including heart, skeletal muscles and brain neurons. Senescence markers (discussed below) have been observed mainly in epithelial, stromal (fibroblastic) and vascular (endothelial) cells that comprise the major renewable tissues and organs such as the skin, intestines, liver, and kidney, and the hematopoietic system. Senescence has also been observed in stem cells

[119]. When cells approach senescence, they undergo irreversible cell cycle arrest, develop resistance to apoptosis and show altered gene expression patterns [76]. However, to date, no markers have been identified that are exclusive to the state of senescence.

A. Growth arrest: The main feature of senescent cells is a lack of capacity to go through the cell cycle in response to mitogens, thus maintaining a DNA content similar to that in G1 phase, although they are still metabolically active. The inability to re-enter the cell cycle is thought to be primarily due to an increase in the expression of certain cell cycle inhibitors (see signaling pathways mediating senescence).

B. Resistance to apoptosis: Senescent cells of most strains appear resistant to the induction of apoptosis. For example, senescent human fibroblasts resist apoptosis caused by oxidative stress or growth factor deprivation [120]. Resistance to apoptosis may partly explain two important phenomena associated with senescence: first, stability of senescent cells in tissue cultures, and the increased number of senescent cells with age. The mechanism responsible for senescent cells being resistant to apoptosis is not well understood. One possibility is that p53 differentially activates cell cycle arrest genes rather than apoptosis-inducing genes [121]. An alternative explanation is that the altered gene expression profile in senescent cells may inhibit the expression of apoptotic genes, such as caspase 3 and others [122].

C. Altered gene expression: a global change in gene expression pattern happens in senescent cells [123]. As discussed before, p16 and p21 CDK inhibitors are expressed at higher levels as cells enter senescence. Moreover, many genes encoding proteins initiating or promoting re-entry into the cell cycle are repressed, such as cyclin A and PCNA [108]. At least in certain cell strains, gene repression is due to inactivation of the E2F transcription factor brought about by the senescence-associated increase in pRb activity. The observed gene repression may also be due to the Rb-dependent modulation of chromatin structure that leads to the formation of distinct senescence-associated heterochromatin foci (SAHF) prohibiting transcription factors from binding to the promoters of proliferation-associated genes [108].

D. Senescence associated β-galactosidase staining (SA-β-gal): Although its biochemical function is still unknown, the most widely accepted marker for senescence both *in vivo* and *in vitro* is the detection of the lysosomal β-galactosidase enzymatic activity at a suboptimum pH (pH 6) rather than pH 4 (optimum pH) using the chromogen X-gal. In the presence of β-galactosidase, X-Gal is converted to a blue colored reaction product [124]. In recent reports describing OIS, β-galactosidase staining was used to help define potentially senescent cells in premalignant tumours [115]. A recent study suggests that SA-β-gal is actually the lysosomal β-galactosidase and senescence-associated increases reflect an increase in lysosomal biogenesis [125].

E. Senescence associated heterochromatin foci (SAHF): The cause of permanent cell cycle arrest in senescence cells is unknown. Narita and colleagues provided a possible molecular explanation for the stability of the senescent phenotype as well as new insights into the molecular role of pRb in mediating senescence. They suggest that pRb contributes to initiating a state of irreversible chromatin remodelling that finally results in permanent cell cycle arrest [108]. This remodelling is associated with recruitment of proteins such as heterochromatin proteins (HP1) and association of pRb with E2F-responsive promoters, resulting in the stable repression of E2F target genes. These regions showed clusters of DAPI-stained heterochromatic regions which the authors designated as senescence-associated heterochromatic foci (SAHF). Additional proteins such as Bmi1 and others have also been suggested to contribute to SAHF formation and the other constituents of these foci are currently being defined.

F. Other markers: In a recent DNA microarray study to examine changes in gene expression during senescence in mouse models, increased levels of decoy death receptor (DCR), a p53-indicible gene and DEC1 transcription factor were detected in premaligant skin lesions, a state known to contain a high percentage of cells showing aspects of senescence, such as an increase in β-gal staining [112].

Epigenetic contributions to senescence

Epigenetic modifications consist of DNA methylation, histone post-translational modifications including, but not limited to, methylation, acetylation, ubiquitination, and phosphorylation, and ATP-dependent modifications of chromatin. Chromatin in young proliferating cells is largely in an ‘open’ or euchromatic state, at least, in the loci associated with transcription of proliferation regulatory genes. On the other hand, certain chromatin domains in senescent cells are packed into a ‘closed’ or heterochromatic structure and may be responsible for the inactivation of cell cycle-associated genes. The first link between aging and epigenetics came almost three decades ago linking genome remodelling that happens during the cell cycle, to cell aging [126]. Three years later, gradual loss of DNA methylation in senescent cells was reported [127]. Several reports came afterwards emphasizing the role of epigenetics in initiating and maintaining the senescence phenotype. Consistent with that, the activity of DNA methyltransferase also decreases with increasing population doubling in human diploid fibroblasts [128]. DNA demethylation plays a role in de-repressing a set of growth inhibitory genes (such as p21) with increased population doublings, perhaps contributing to withdrawal from the cell cycle [129].

Histone acetylation is also thought to play a very important role in mediating senescence. The histone acetyltransferase (HAT) p300 is an important integrator controlling transcription, differentiation, and DNA repair [130]. Fibroblasts deficient in p300 undergo senescence [131]. Moreover, p300 and CBP levels decrease with increased population doublings in melanocytes [132]. The decrease in levels of both proteins leads to a dramatic decrease in total histone H3 and H4 acetylation [132]. As a consequence of p300 downregulation, the cyclin E gene promoter has been shown to lose histone H4 acetylation which contributes to a decrease in cyclin E protein level in senescence. Moreover, the cyclin E promoter is also occupied by histone deacetylase (HDAC1) and Rb to repress transcription [132]. This is an interesting observation as it links decreased levels of histone acteylation and transcription repression to the p16-Rb pathway and the formation of SAHF. Paradoxically, HDAC inhibitors which can induce global histone acetylation can also induce a senescence-

like phenotype in human fibroblasts [133]. Both observations strongly suggest that the transition between growth competent and senescent cells requires a critical balance between acetylated and deacetylated chromatin. Probably, the loss of histone acetylation by loss of p300/CBP can shift the equilibrium towards formation of heterochromatin at specific loci and silencing of proliferation-associated genes. The loss of histone deacetylation induced by HDAC inhibitors, however, may also alter the expression of subset of genes while activating cell cycle inhibitory proteins in other regions. This idea is supported by the observation that HDAC1-deficient embryonic stem cells show reduced proliferation, which correlates with increased levels of the CDK inhibitors p21 and p27 [134]. Overall, it seems likely that balanced histone acetylation and deacetylation at specific genetic loci is required for proper control of cell growth and for the entry into cell senescence.

Section III. Protein interaction networks

Protein-protein interactions mediate a plethora of biological functions in the cell. The failure of proteins to interact with their corresponding partners may result in a failure to properly carry out their functions. The complexity of higher organisms may be due to increased complexity of protein interactions. For example, humans have a slightly larger number of genes than nematodes (around 25,000 versus 20,000, respectively), yet we are much more complex than nematodes, with humans containing $10^{12\text{-}14}$ cells and nematodes containing around 1000 cells. Perhaps this is because we have a more intricate protein interaction network. Elucidating the complete interactome network will reveal various unexpected nexi of proteins. Accordingly, unraveling protein interaction networks have garnered increasing interest in biological sciences research. Invention of new functional genomics and proteomics techniques allowed the investigation of all possible protein interactions in the entire genome. *S. cerevisiae* became the first model organism for which large-scale protein interaction data were presented [135,136]. Two independent studies used the yeast two-hybrid system and examined certain pools of protein interactions in yeast. Two years later, two more comprehensive, yet still incomplete, interactome networks were reported [137,138] using affinity purification/mass spectrometry. Now, several other methods have been used to generate global interactome data such as protein chips [139] and synthetic lethality [140], among others, and has been extended to include many other organisms, such as worm, fly and even humans.

Biochemical methods of determining large scale protein-protein interactions (PPI)

The yeast two-hybrid (Y2H) system was the first method used for the defining the biological protein interactome using a complete set of open reading frames (ORFs) in a eukaryotic organism [135,136]. Two domains are required for this assay: 1) a DNA binding domain (DBD) that can bind to DNA, and 2) an activation domain (AD) for transcriptional activation of DNA. When both domains come together, a reporter gene, which often encodes an essential gene, is transcribed [141]. The protein of interest is linked to the DBD (Bait protein) and the other protein is linked to the AC (Prey). If the bait and the prey interact with each other, the DBD and AD fusion become functional and the reporter gene is transcribed

[142]. The Y2H system has been utilized in two different ways to elucidate the yeast interactome: 1) protein array, where the PCR-generated ORFs were linked to GAL-AD and the transformed yeast was mated with yeast having the GAL-DBD and the diploid hybrids were selected by reporter, and 2) high throughput screening, where the PCR-generated ORFs were linked to either GAL-AD or GAL-DBD, transformed into yeast, and each of the DBD cells was mated to the AC library. Although the second approach gives much higher numbers of potential protein interactions and requires less work and less reagents, protein arrays give more confident results [143].

Y2H is a useful tool for detection of direct protein interactions. Moreover, it allows for the capture of transient interactions that may be missed by affinity purification methods. However, as noted above, the technique is notorious for detecting large number of false positive interactions. This may happen due to the activation of RNA polymerase by the protein of interest or by nonspecific binding of the bait and prey proteins to some endogenous proteins. Moreover, true interactions may be lost if false positive associations are rigorously screened out or if localization is an issue. For example, if the protein examined is localized to the membrane then it will not be able to activate the reporter gene as the interacting proteins must be localized to the nucleus. Also, if any of the proteins require posttranslational modifications to carry out its function, it may not be able to interact properly. Finally, protein folding may be affected when they are overexpressed in the yeast nuclear environment, and hence they may not interact.

Tandem affinity purification (TAP) tagging is an approach invented to examine PPI under near normal physiological conditions [144]. The method encompasses double tagging of the ORF followed by a double purification processe (that is why it is named tandem), most commonly used utilizing Staphylococcus protein A and calmodulin beads. The ORFs are overexpressed in yeast and form native complexes with other proteins. The complexes formed with the tagged protein are then affinity purified via a two-step purification procedure and analyzed by SDS-PAGE followed by MS [144]. Gavin *et al.* and Ho *et al* were the first to use this method in screening the yeast interactome [137,138]. In comparison to the Y2H approach (using BIND database as a reference, a standardized database for all biochemically

validated PPI [145]) the TAP tagging/MS approach yielded 3 fold more literature verified interactions per bait than each of the Y2H published interactomes [135,136]. This may be due to the fact that TAP tagging/MS allows for purification of protein complexes under more natural conditions than those generated by the Y2H system, allowing for detection of larger numbers of PPIs. It is also much more sensitive in that it can identify proteins present at only 15 copies per cell. Not only can membrane-associated proteins be purified through the TAP tagging process (which cannot be done using Y2H) but very large complexes can be isolated as well (around 1.5 Mega Dalton). Moreover, unstructured proteins that properly fold only when they are in complexes, but not when they are in a binary interaction, can be identified via tagging/MS. Despite its strength, TAP tagging may interfere with the formation of certain protein complexes due to steric hindrance by the tag [138]. Also, potentially low expression of the fusion protein may affect the protein's ability to interact with other proteins. New technologies are now being implemented to increase the reproducibility and the efficacy of the TAP/MS approach [146].

Due to the large amount of data generated by different biochemical approaches, several databases have been constructed to organize these data. For example, the biomolecular interaction network database (BIND) shows detailed description of the experimental approach used to identify a particular PPI [145]. The general repository of interaction databases (BioGRID) contains protein and genetic interactions from 13 species [147]. Another database of interest is the molecular interaction database (MINT), in which the PPI data are extracted from the literature and scores are given to each interaction according to the weight of evidence for each interaction [148]. The above mentioned databases, among others, have been successfully used to validate and carefully examine potential functionally relevant interactions from the vast amount of PPI generated by high-throughput screens.

Computational prediction of PPIs

Various computational algorithms have been generated to predict novel PPI. A list of some of the computational methods used is discussed below.

Genomic methods

Genomic methods make use of the complete genome sequence information as it provides data about how the genes are organized (gene order) [149]. The methods do not depend completely on sequence similarities between genomes, rather they examine functional links between co-localized genes. In general, if the genes of two or more proteins are physically close in different genomes (ex. both genes are in close positions on the same chromosome), it is highly probable that they interact. Moreover, genomic methods can compare protein pairs in one organism's genome to a fused single protein homologue in another higher organism's genome. If the fused gene contains different segments of two different genes in a different organism, the two proteins encoded by these two genes will be predicted to interact. Marcotte *et al.* developed a cross-species computational method to detect fusion events among genomes of different genomes [150]. The gene-fusion approach identified thousands of putative PPI in yeast and some of them were biochemically validated. For example, upon comparing yeast and *E. coli* genome sequences, it has been found that genes for the Gyr A and Gyr B subunits of the *E. coli* DNA gyrase are actually fused as a single gene in yeast topoisomerase II. This meant that Gyr A-Gyr B PPI is highly probable [150]. However, genomic approaches cannot predict PPI for proteins encoded by distantly-located genes. Also, in more complex organisms like humans, this approach may not be useful as the co-regulation of genes at the genome level is rare in higher organisms [150].

Protein three-dimensional structure-based PPI prediction

Three-dimensional structure of proteins can be used to predict PPIs [151]. The method involves fitting of the two potentially interacting proteins in a complex of known three-dimensional structure and examining the molecular basis of how the interaction may occur. Residues that make contact in a crystallographic complex are analyzed and an interaction is conserved as long as the contact residues are conserved [152]. InterPreTS (interaction prediction through tertiary structure) is a web based method that uses the 3D structure to validate PPIs that have been predicted by other methods [153]. Moreover, it allows visualization of the molecular details of predicted PPIs. Another web-friendly tool is

PRISM (protein interaction by structural matching), a database that compares the structure of protein pairs with a dataset of interfaces of ‘crystal’ interactions [154]. PRISM aligns the structures of the two proteins to calculate the interface similarity. It allows for searching protein interfaces and prediction of PPI. This protein three-dimensional structure based methods, however, can be used only for proteins or domains with known crystal structures.

Domain-based prediction of PPI

A number of PPI prediction techniques are based on domain structure of proteins. Using the conserved domains in the Pfam database (protein families), the probabilities of interaction between every pair of domains can be estimated [155] and these domain-domain interactions can be used to predict PPI. PreSPI (prediction system for protein interaction) is a computational tool that can predict PPI based on conserved domain-domain interactions [156].

Primary protein structure

Specific short amino acid sequences can mediate PPI. These sequences do not necessarily form or exist as part of protein domains. Indeed, it has been shown that amino acid sequence alone can be used effectively to predict PPIs [157]. Like other computational methods, primary sequence based tools depend on previously published PPI and use the literature to predict novel PPIs. For example, PIPE (protein interaction prediction engine) is a computational tool that can predict PPI for yeast proteins based only on the protein pair primary sequence [158]. PIPE predicts the interaction probability by measuring how often sequences pairs in two proteins co-exist in protein pairs known to interact. Surprisingly, the method showed an overall 75% accuracy, 89% specificity and 11% false positives for prediction of PPI in yeast [158], a success rate comparable to traditional biochemical techniques.

Weaknesses of computational prediction methods

Although computational methods provide an invaluable tool in predicting novel PPIs, its high-throughput nature allows false positive and negative predictions. False positives are not uncommon in most computational methods. Several reports estimated the number of PPIs in yeast to be between 10,000- 30,000. The most comprehensive yeast interactome dataset as of July 2008 showed around 7,000 interactions at most. This indicates that the number of false negative interactions may be a serious problem masking the real number of true PPI. The problem is mainly due to the lack of standards that can assess different computational approaches used, and limitations in validating the results of various methodologies. One possible approach to increase the accuracy of future tools will be to integrate more than one approach to predict interactions.

Hypothesis and specific aims:

Based upon the background provided, we hypothesized that ING1a and ING1b have differential effects on senescence and apoptosis. We also hypothesized that elucidating novel ING1 protein interaction will help us better understand the role of ING1 protein in diverse biological processes. To test our hypothesis, we examined the following specific aims: 1) analysis of the differential expression of ING1 isoforms in senescent human fibroblasts, 2) examination of a causal role of ING1a in promoting senescence-like features, 3) investigation of the molecular mechanism(s) mediating ING1a biological functions, 4) studying the effects of ING1a on gene expression patterns in human fibroblasts, 5) analysis of the degree of conservation of different ING protein domains across various species, 6) computational analysis of potential human ING interacting proteins, based on yeast and fly interactome datasets, 7) generation of functional biological networks for ING proteins, and 8) biochemical validation of a subset of predicted ING protein interactions.

CHAPTER TWO: METHODS AND MATERIALS

Cells and cell culture

Hs68 human diploid foreskin fibroblasts (ATCC# CRL-1635), WI38 human embryonic lung fibroblasts (ATCC# CCL-75) and immortalized human embryonic kidney (HEK) 293 cells (ATCC-CRL-1573) were purchased from American Type Cell Collection (ATCC). Hs68 fibroblasts were grown in low glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 4mM L-glutamine and 1mM sodium pyruvate. WI38 and HEK 293 cells were grown in high glucose Dulbecco's MEM supplemented with 10% FBS as suggested by ATCC. Cells were maintained in a humidified atmosphere of 5% CO_2 and 95% air at 37°C and media was changed every 2-3 days. Culture media and FBS were supplied by Gibco-BRL or Fisher Canada. Plastic tissue culture plates were supplied by Fisher, Sarstedt, or Corning Canada. The plates were non-pyrogenic and plasma coated intended for culture of adherent cell types. Plasma treatments ensured that hydrophobic polystyrene is permanently rendered hydrophilic to support cell attachment.

Freezing and thawing of cells

Cryopreservation was done by harvesting cells using trypsin-EDTA (Gibco-BRL) treatment followed by centrifugation at 1500rpm for 5 mins. Cells were then resuspended in a medium containing 10% FBS and 5-10% sterile dimethylsulfoxide (DMSO, Sigma) to yield approximately l-2 X $l0^6$ cells/mL. One mL aliquots of cell suspension were transferred to cryovials (Nalgene) and vials were placed at -80°C overnight. Frozen cells were then placed in liquid nitrogen for long-term storage.

To thaw cells, vials of frozen cells were removed from liquid nitrogen and placed in a 37°C water bath for 3 mins. The thawed cell suspension was quickly transferred to plates containing fresh culture media supplemented with 10% FBS then incubated at 37°C in an atmosphere of 5% CO_2. The next day, media was replaced to remove the DMSO and incubation was continued under the same conditions.

Splitting of cells

Cells reaching confluence were incubated in a trypsin-EDTA solution at 37°C in a 5% CO_2 incubator for 2-5mins until they lose anchorage from the tissue culture plates. The cells were harvested by triturating with an appropriate volume of DMEM containing 10% FBS. Cells were dislodged by titration and seeded onto fresh tissue culture dishes. Uniform spreading of cells was achieved by gently shaking plates containing medium and cells in cross-shape direction. Primary cells including Hs68 and WI38 fibroblasts were split at 1:2-1:4 ratios depending on the purpose of the experiment. Low-passage (young) cells used were between 28-37 MPDs (Hs68) and 22-28 MPDs (WI38) and high-passage (senescent) cells used were between 89-92 MPDs (Hs68) and 54-56 MPDs (WI38). At the indicated young *in vitro* age, cells remain competent for growth and traverse the cell cycle freely in response to mitogens. HEK 293 cells were passaged at 1:2-1:10 ratios depending on the growth rate of the cells and the purpose of the experiment.

Synchronizing of cells

In order to synchronize human fibroblasts, cells were grown to confluence and split in DMEM containing 10% FBS. Cells were then incubated at 37°C for 6-8 hrs to allow adherence. Following this short incubation, media containing FBS was aspirated and replaced with media lacking FBS (without washing with PBS) then cells were incubated at 37°C for 4 days. The sub-confluent serum-deprived cells were then stimulated by the addition of media containing 10% FBS enabling a large portion of the cell population which had been arrested in the G_0 phase of the cell cycle to synchronously enter the G_1 phase.

Cell fractionation

Cells were harvested in 1mL PBS and centrifuged at 10,000rpm for 2 mins. The supernatant was aspirated and the cell pellet was suspended in 300µL PBS and divided into 100 and 200µL portions. The 100 µl portion was centrifuged and the cell pellet was boiled with Lammeli sample buffer to be used as the whole cell lysate portion. The other portion

was pelleted, the supernatant was aspirated and the pellet was resuspended in 10 volumes PBS containing 0.1% NP-40 on ice. The cells were triturated 5 times through a 20-gauge needle using 1mL syringe, centrifuged at 10,000rpm for 2 mins and the supernatant was transferred to another tube and boiled with Lammeli sample buffer to be used as the cytoplasmic fraction. The pellet is triturated again on ice in PBS containing 0.1% NP-40 to ensure the removal of cytoplasmic membranes of ruptured cells. The nuclei are pelleted and resuspended in an appropriate volume of Lammeli sample buffer to be used as the nuclear fraction and the protein concentration of the three fractions (whole cell lysate, cytoplasmic and nuclear fractions) are checked with Coomassie blue staining. From the experiments done in the course of this study, the ratio of Lammeli sample buffer added to the different cell fractions was 2:2:1 (whole cell lystae: cytoplasmic: nuclear).

Cell transfection

Cells were seeded onto tissue culture plates at 40-50% confluence. 16-18 hrs later when cells reached 80-90% confluence, cells were transiently transfected using Lipofectamine-2000 transfection reagent (Invitrogen) or Fugene 6 (Roche). For Lipofectamine transfections, cells were washed with PBS and media was replaced with serum-free medium. Plasmid DNA and Lipofectamine reagent were then mixed separately in antibiotic and serum-free medium for 5 mins at room temperature then mixed together for 20 mins with agitation at room temperature. The ratio of µg of DNA to µL of Lipofectamine was 1:2. After 20 mins of incubation, the DNA:Lipofectamine complexes were added dropwise to cells. Growth medium was then replaced after incubation for 4-6 hrs at 37°C with fresh media containing 10% FBS. For Fugene 6 transfections, media was replaced with fresh antibiotic free FBS-containing media. Fugene 6 was added to antibiotic and serum-free media and incubated for 5 mins. DNA was added to the Fugene 6-media mixture and incubated for 20 mins at room temperature. The ratio of DNA:Fugene 6 used was 1:2. Amounts of DNA and Lipofectamine used were according to the manufacture's recommendation. The mixture was added dropwise to the plate of cells containing media without antibiotics and distributed evenly by gently shaking the plate forth and back several

times. The cells were incubated at 37°C in a 5% CO_2 incubator and harvested at the time points indicated in each experiment. Co-transfection with green fluorescence protein (GFP) was done to monitor for transfection efficiency. Transfected cells were then selected for further analysis.

Viral infection

Recombinant adenoviral vectors were used for gene delivery and expression in human primary fibroblasts as they can infect a broad variety of cell strains very efficiently and independent of cell division. The transfer of the reporter gene, GFP in this case, to almost 90-95% of the cells happens without sign of toxicity at a Multiplicity of Infection (MOI) of 100.

The AdEasy system

The AdEasy system is developed to simplify the process of generating adenoviral vectors. There is no ligation steps needed, as the process depends on efficient homologues recombination (HR) machinery present in *E. coli*, rather than inefficient human HR. The vectors used have a GFP gene incorporated into the adenoviral backbone, which allowed monitoring of infection efficiency. The AdEasy backbone lacks the E1 gene which renders the virus defective for replication and incapable of producing viral particles in target cells other than HEK 293 cells. HEK293 E1 cells have the E1 region integrated in their genome and therefore can provide the virus with the E1 gene to re-gain its reproduction capacity.

The generation of adenoviral vectors using the AdEasy system includes three steps: first, ING1a or ING1b was cloned into the transfer vector pAdTrackCMV, which has a GFP gene in a separate transcriptional unit (bicistronic). Second, the transfer vector now containing ING is linearized with Pac I and mixed with the adenoviral backbone, pAdEasy, which lacks the E1 gene. The mixture was then transformed into the BJ5183 bacterial strain, which contains highly efficient homologues recombination machinery. The potential recombinant adenoviral DNA was screened through analysis of the size of the DNA and restriction digestion. Finally, the recombinant DNA was transfected in HEK 293 cells using

Lipofectamine-2000. Plaque purification was performed and pure plaques are amplified on a large scale.

Viral production in HEK 293 cells

HEK 293 cells were grown on tissue culture plates to reach 100% confluence. The cells were then transfected with the recombinant DNA and the viral production process was monitored by GFP reporter under the fluorescence microscope (Zeiss). The HEK 293 cells lysis took place over 36-72 hrs. When 50% of the cells or more rounded up and detached from the plate, the cells and media was collected in 50mL Falcon tubes and frozen in a dry ice-ethanol mixture then thawed in a 37°C water bath followed by vigorous vortex to ensure complete release of viral particles from cells. The freeze-thaw process was repeated three times. The cell lysate was then centrifuged at 6000rpm at 4°C to remove cellular debris, and filtered twice through 0.45µm sterile filters to completely remove cell debris. The crude viral lysate (CVL) was then aliquoted in small volumes (10-100µL) and stored at -80°C. For amplification, 10µL of the CVL was used to infect one 15cm^2 plate of HEK 293 cells from which the CVL was made according to the above mentioned procedures. Fifteen 15cm^2 plate of HEK 293 cells were infected using the CVL made from the above mentioned step. When 50% of the cells were detached, media was discarded and the cells were collected in 15mL of media and the CVL was prepared as previously described.

Infection of cells

Hs68 and WI38 fibroblasts were infected at a MOI of 100 where 95% of the cells were infected and monitored by GFP expression with no toxicity observed. The amount of virus used for infection was calculated as follows: amount of CVL to be used (µL) = number of cells X MOI X 1000/pfu/mL. The CVL aliquots were thawed on ice before use. One day before infection, cells were split to reach 80% confluence next day. 16-18 hrs post-splitting, media was changed with an amount just enough to cover the plate to avoid viral dilution and obtain maximum infection efficiency. The calculated amount of virus (in my study, it was 1µL for a 12-well plate of human fibroblast) was then added to the media and distributed

evenly by gentle shaking several times. After incubation at 37°C incubator for 6 hrs, more media was added to the plate and cells were incubated under the same conditions until harvesting. Alternatively, media was removed and replaced with serum-free media and the calculated amount of virus was added. Media was replaced with DMEM/10% FBS 12 hrs post-infection and the cells were incubated under same condition until harvesting.

Plasmid preparation and DNA constructs

ING1a and ING1b were cloned into pCI vectors (Clontech). pGFP–N1 (Clontech) was used as a negative control and was co-transfected with ING1a or ING1b constructs to monitor for transfection efficiency and gate transfected cells for further analysis. Small scale and large scale plasmid preparations were done using the Qiagen plasmid preparation kits. These plasmid purification protocols are based on a modified alkaline lysis procedure, followed by binding of plasmid DNA to an anion-exchange resin under appropriate low-salt and pH conditions. RNA, proteins, dyes and low molecular weight impurities were removed by a medium salt wash. Plasmid DNA was then eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation.

Senescence-associated beta-galactosidase staining

Senescence-associated beta-galactosidase (SA β-gal) activity as a measure of senescence was detected as described [124]. Cells were washed with phosphate-buffered saline (PBS, pH 7.2), fixed with 3% formaldehyde in PBS for 5 mins at room temperature, washed with PBS (pH 6.0) or citric acid/phosphate buffer (pH 6.0), and stained for 16-24 hrs at 37°C. The staining was done under no CO_2 as it may affect the pH of the solution during the incubation period. The staining solution contained 1mg/mL of 5-bromo-4-chloro-3-indolyl-3-galactoside (X-gal) in PBS (pH 6.0), 5mM potassium ferrocyanide, 5mM potassium ferricyanide, 150mM NaCl and 2mM MgCl. The pH is adjusted to be suboptimal as buffers of pH 4.0 will stain for lysosomal β-gal activity which is present in both young and senescent cells.

Western blot and Co-immunoprecipitation-western (IP-western) assays

Preparation of total cell lysates

Cell culture plates were washed 3 times with the appropriate amount of cold (4°C) PBS and then scraped on ice. Cell suspensions were then transferred into 1.5mL eppendorf tubes, centrifuged at 13,000rpm for 1 min and the cell pellet was snap frozen in liquid nitrogen and stored in -80°C until used. For immunoprecipitation, cell pellets were lysed in 1mL RIPA lysis buffer (20mM Tris-HCl, pH 7.5, 100mM NaCl, 5mM KC1, 1mM EDTA, 0.25% deoxycholate, 0.25% Nonidet P-40, 0.25% Tween-20) containing EDTA-free protease inhibitor cocktail (Roche Diagnostics). For western blotting experiments, 200µL of 2X SDS sample buffer (2% SDS, 20% glyverol, 20mM TrisCl pH 8.0, 4% β-mercaptoethanol, 2mM EDTA, and 0.02% bromophenol blue) was added directly to the cell pellets (lysis buffer was prepared without bromophenol blue which was added after carrying out the Bradford assay to avoid interfering with the assay results). Cell pellets were sonicated on ice 3 times for 10 secs each, centrifuged for 15 mins at 4°C at 13000rpm and the supernatants were used immediately for immunoprecipitation or western blotting or aliqouted and frozen in -80°C.

Quantification of protein by the Bradford assay

The BSA concentration curve was done before measuring the concentrations of protein samples using the Bradford assay described in detail in the instruction manual. The Protein samples were diluted 100 times to avoid SDS interference with the Bradford reagent.

Western blotting

For western blots, denatured whole cell lysates in sample buffer were electrophoresed (SDS-PAGE) and transferred to polyvinylidene difluoride (PVDF) or nitrocellulose (NC) membranes for one hour at 70V. Detection of ING proteins was done with a panel of mouse monoclonal α-INGl (CAbs 1-4) in complete serum media diluted 1:1 in 5% milk (SACRI Antibody Services). Mouse α-INGl antibodies recognize both native and denatured forms of INGlb, INGla and the truncated ING1c isoform. Additional antibodies used for western blots in the course of this study include β-actin (Cell Signaling, # 4967, 1:1000), pRb (Santa Cruz, sc-50, 1:200), p16 (Cell Signaling, #4824, 1: 1000), PARP1 (Santa Cruz, sc-53643, 1: 200),

GFP (Abcam, 6556, 1: 2000), rabbit polyclonal and QM (Santa Cruz, SC-798, 1:200). Horseradish peroxidase (HRP)-conjugated secondary antibodies were used for western blotting (Amersham-Pharmacia Biotech, 1:2000 for goat α-rabbit and 1:5000 for goat α-mouse) and blots were developed using chemiluminescent substrate (Millipore). In UV irradiation experiments, media was removed and cells were exposed to 30 J/m^2 of UVB. After adding fresh media to the cells, plates were incubated at 37°C and harvested 2 hrs after UV.

Immunoprecipitation assays

ING1 mouse monoclonal hybridoma supernatants (CAbs 1-4) (15mL) were incubated with 40µL 1:1 slurry of sepharose G beads in PBS overnight on a rocker at 4°C. The next day, the hybridoma supernatant was centrifuged at 1,500rpm at 4°C for 5 mins to collect the antibody-saturated beads. The beads then were washed 3 times with ice-cold RIPA buffer and resuspended in 40µL ice-cold RIPA buffer. In parallel, cell lysates (see Preparation of total cell lysates) were pre-cleared by mixing with 20µL of 1:1 protein G sepharose beads slurry per 1mL cell lysate and incubated on a rocker at 4°C for 30 mins. Cell lysates were then centrifuged at 13,000rpm for 10 mins at 4°C to remove the beads. ING1 antibody saturated beads were added to the 1mL pre-cleared cell lysate and incubated on a rocker at 4°C for 4-6 hrs. The beads then were washed quickly 2 times with ice-cold RIPA buffer and resuspended in 30µL 2X SDS sample buffer and the supernatant was subjected to western blotting using the antibodies indicated.

Indirect immunoflourescence

48 hrs after infection, cells at 70-80% confluence were washed 3 times with PBS and then fixed with 4% paraformaldehyde in PBS for 10 mins at room temperature, washed with PBS, permeabilized by 0.5% Triton X-100 in PBS for 10 mins at room temperature, and washed with PBS. Fixed cells were incubated with phalloidin-conjugated-FITC to stain for actin. For Heterochromatin Protein 1 gamma (HP1γ) staining, cells were incubated with mouse anti-HP1γ (1:500) in 10% FBS/PBS for 30 mins, washed and then incubated with goat anti-mouse IgG-Alexa 488 (Cedarlane) (1:1000) in 10% FBS/PBS for 30 mins. After rinsing

three times with PBS, samples were mounted in 1µg/mL paraphenylenediamine in PBS/90% glycerol containing DAPI at 1µg/mL for 10 mins. Imaging was performed using a Zeiss Axiovert 200 microscope and AxioVision 4.5 software.

Flow cytometry and cell cycle analysis

To prepare adherent cells for Fluorescence Activated Cell Sorting (FACS) cell cycle analysis, cells were trypsinized until partially dislodged and an appropriate amount of complete media was used to further triturate cells. Samples were then centrifuged at 800rpm for 10 mins at 4°C in 15mL Falcon centrifuge tubes. The media was aspirated and the cell pellet was washed twice with ice-cold PBS/2mM EDTA/1% FBS to inhibit DNases and lubricate cells, respectively, and resuspended in 0.5mL of ice-cold PBS and 0.5ml of 3% formaldehyde for fixing cells and left on ice for 1 hr. Cells were centrifuged at 1,500rpm for 5 mins, washed with ice-cold PBS/2mM EDTA/1% FBS and resuspended in 0.5mL of the same mixture. 1mL of ice-cold 100% ethanol was then added dropwise with vortex to permeabilize cells and cells were either stored at -20°C or processed further. Cell suspensions were placed in 5mL polystyrene round bottom tubes (Becton Dickinson 35-2054) and centrifuged at 3,500rpm for 15 mins at 4°C and supernatant was aspirated. Cells were washed twice in PBS/2mM EDTA/1% FBS. The washed cells were resuspended in 800µL of PBS and incubated for 30 min at room temperature with 100µL of 1mg/mL RNase A. An additional 100µL of 100µg/mL propidium iodide was added to tubes (final volume 1mL), and samples were incubated in the dark for 30 mins at room temperature. FACS analysis was performed using a BD FACScan™ (BD Biosciences) at the University of Calgary Flow Cytometry Core Facility within an hour.

Microarray analysis

Total RNA was first extracted from experimental samples to be compared (cells infected with GFP as a reference sample and with GFP-ING1 as a test sample) and fluorescently labelled with two different dyes (Cy3 and Cy5) in a single round of reverse transcription. The fluorescently labelled cDNA probes were hybridized to a single array in a

competitive hybridization reaction. Detection of hybridized probes was achieved by laser excitation of the individual markers followed by scanning using a confocal scanning laser microscope. The Cy5/Cy3 or Cy3/Cy5 fluorescence ratio of each gene on the array reflects the relative abundance of the relevant transcript in the test mRNA pool versus that of the reference. Data are digitally color coded such that red represents genes transcriptionally upregulated in the test versus the reference, green represents genes downregulated and yellow represents those genes that exhibit no difference between test and reference samples.

RNA extraction from cultured cells

WI38 embryonic human lung fibroblasts (MPD 27) were infected with either GFP alone or with GFP-ING1 adenoviruses. 48 hrs after infection, cells were harvested in PBS and total RNA was extracted using Qiagen RNeasy kit according to the manufacturer's manual. This method combines the selective binding properties of a silica-based membrane with the speed of microspin technology. A high-salt buffer system allows up to 100μg of RNA longer than 200 bases to bind to the RNeasy silica membrane. Cells were first lysed and homogenized in the presence of a highly denaturing guanidine-thiocyanate–containing buffer, which immediately inactivated RNases to ensure purification of intact RNA. Ethanol was added to provide appropriate binding conditions, and the sample is then applied to an RNeasy mini-spin column, where the total RNA bound to the membrane and contaminants were efficiently washed away. RNA is then eluted in 30μL water. With the RNeasy procedure, all RNA molecules longer than 200 nucleotides were purified. The procedure provides enrichment for mRNA since most RNAs less than 200 nucleotides, such as 5.8S rRNA, 5S rRNA, and tRNAs, which together comprise 15–20% of total RNA, were selectively excluded.

Measurement of RNA quality and quantity

The RNA quantity and quality was determined using Bioanalyzer 2100 (Agilent). The ratio of S28 and S18 bands intensities was around 2:1.

Synthesis of cDNA from total RNA

cDNA was made from 10µg total RNA using a two-step reverse transcription method. First, DNAsed RNA was incubated with 1µL of 500ng/mL d(T) primer at 70°C for 10 mins and quickly chilled on ice. Then, 2µL of 10X StrataScript buffer, 1.5µL of 0.1M DTT, 1µL of 20X dNTP mixture, 0.5µL of 40U/µL RNase inhibitor were added and the samples were incubated at room temprature for 5 mins before 3µL of Affinity Script reverse transcriptase (FairPlay III, Stratagene) was added to each tube. After incubation at 42°C for 1 hr, 10µL 1M NaOH was added to the cDNA and the samples were heated at 70°C for 10 mins to inactivate enzymes and hydrolyse RNA. After cooling to room temperature, 10µL of HCl is added to neutralize the solution. To purify the cDNA, 4µL of 3M Na acetate (pH 4.5), 1µL 20mg/mL glycogen and 100µL ice-cold 100% ethanol were added to the reaction mixture and incubated over night at -20°C. The tubes were spin at 13,000rpm for 15 mins at 4°C and the supernatant was decanted. 0.5mL of ice-cold 70% ethanol was added and the tubes were centrifuged again using the same conditions mentioned above. After decanting the supernatant, the pellets were left to dry.

Flourescent labeling of cDNA probes

An indirect labelling was used to generate Cy3 and Cy5 fluorescent-labelled cDNA. Cy3 and Cy5 are preferentially used because they are readily incorporated by reverse transcription, they exhibit good photostability and most importantly, are widely separated in terms of their excitation and emission spectra. However, it is well known that Cy3 and Cy5 fluorescent dyes are incorporated at different rates in reverse transcription reactions and have difference quantum yields. This results in a difference in the Cy3 and Cy5 fluorescence intensities even when equal amounts of Cy3 and Cy5-labeled cDNA are present. To get around this dye-bias problem, reciprocal labelling experiments were done in which both test and reference samples are labelled with Cy3 and Cy5 dyes and *vice versa*. After resuspending the cDNA pellet in 5µL 2X coupling buffer, the mixture is incubated for 10

mins at 37°C. 5µL of the dye (Cy3 or Cy5) was added and the labelling took place by incubating the mixture for 1 hr at room temperature in the dark.

Hybridization of cDNA to microarray human gene chips

The labelled cDNA samples were then washed twice with 70% ethanol and then twice with 75% ethanol to remove any uncoupled fluorescent dye and subsequently combined together with 65µL Hyb solution (5µL yeast tRNA, 5µL fish sperm and 90µL DIG Easy solution) heated at 65°C for 2 mins, cooled to room temperature, and hybridized to 14K human oligo chips (Southern Alberta Microarray Facility) by incubating at 37°C under humidified condition for 18 hrs. After hybridization, arrays were washed 3-4 mins with 2X SSC/0.2% SDS, for 1 min with 0.2X SSC and for 1 min with 0.05X SSC. Slides were quickly dried and scanned using fluorescence laser microarray scanning device (Virtex).

Data analysis

To avoid non-specific experimental variation, which is a major caveat of this technology, dye-reversal experiments were performed for each ING gene assay. Data were first quantitated by Array-Pro software (Media Cybernetics) and then transferred to GeneTraffic bioinformation software (Iobion Informatic Company) for further data normalization, annotation and management.

Computational approaches

The process of creating a list of ING protein interaction consists of 8 broad steps. The first step was to determine the ING1 homologues in yeast. This was done using the Smith-Waterman sequence alignment tool [159] against the full set of *S. cerevisiae* genes downloaded from the yeast genome database (SGD) [160]. In step 2, we identified ING protein domains more specifically. In practice, we already know the domains ING contains, but the construction of domain models provides a quantification of their conservation among species. A multiple sequence alignment (MSA) was performed using CLUSTAL-W [161]. In step 3, the HMMER software (http://hmmer.janelia.org/) was used to build and calibrate

Hidden Markov Models (HMMs) from multiple, distinct conserved regions across species. In step 4, these HMMs were used to search for proteins in other species with the same domains. The database searched against was Uniprot, which provides a non-redundant set of all known eukaryotic genes. Each HMM search result was reviewed, and portions of database sequences deemed matching, primarily those with e-value of less than 10^{-5} were incorporated into the HMM. This searching and extra sequence incorporation was done iteratively (since new sequences in the HMM affect e-value results) until no new matches were found in Uniprot. The end result of step 4 was that for each domain we had an HMM representing the domain's very particular conservation across eukaryotic species. Step 5 was to compare the HMMs against the complete protein sets from model organisms with large amounts of interaction data, namely, worm [162], fly [163] and yeast (SGB). This was done using the HMM search, and identified model organism genes ING with the same domains as the human ING. Given HMM matches in yeast and at least one more model organism (the fly database in our case), we proceeded in step 6 to extract human ING interacting proteins from model organism sources ING-interactors lists. The source of the interaction data was either the yeast database [164] or the fly database [163]. Step 7 was to reduce the list of ING-interacting proteins in the model organisms to just those satisfying two conditions: showing strong pairwise homology among human, yeast and fly, and having interaction data in both yeast and fly. These filters reduced the potential human ING interacting proteins list dramatically and increased the confidence in our list of interacting partners. Finally in step 8, we viewed the human homologs of each ING-interacting protein in the STRING database [165] and assessed the biological relevance of the potential interactions in humans based on the premise that the ING-protein interaction is in fact conserved as a true interaction in humans. We have focused on the use of thorough methods in our approach to maximize the sensitivity of our results. It would be possible to substitute certain methods, such as BLAST [166] for Smith-Waterman, or InterPro models searches [167] for HMM building in these steps, with the caveat of reduced confidence values of the results.

CHAPTER 3: RESULTS

Section I. ING1a mediates cellular senescence

ING1 splicing isoforms are differentially expressed during cell senescence.

To examine whether the mRNA levels of the two major ING1 isoforms, ING1a and ING1b (Figure 5A), are expressed differently in young versus senescent human diploid fibroblasts, RT-PCR using isoform specific primers was done with GAPDH primers as an internal control for mRNA integrity and amplification efficiency. Interestingly, the level of INGlb mRNA decreased in senescent cells as previously reported [57]. In contrast, the level of INGla transcript increased significantly in senescent fibroblasts (Figure 5B). INGlb protein levels also decreased to a greater degree than mRNA levels while INGla levels increased in senescent cells (Figure 5C). The greater decrease in ING1b protein as opposed to mRNA levels may be due to altered phosphorylation of a residue recently shown to regulate ING1b proteosomal degradation [168].

INGla induces senescent cell morphology.

Although ING1a levels correlated with senescence, we wished to see if this isoform could play a causal role. To test whether INGla could induce a senescent morphology in normal fibroblasts, human fibroblasts were infected with GFP and either INGla or INGlb adenoviral constructs. The mouse monoclonal anti-ING1 (CAbs 1-4) mixture detected both ING1 isoforms with similar efficiency (Figure 6). The amounts of viral constructs used were adjusted to allow for similar expression levels of both isoforms (Figure 6). Figure 7A confirms the occurrence of SAHF in senescent cell nuclei (panel B) and highlights the similarity of nuclear phenotype and DNA staining patterns between senescent cells and cells expressing high levels of INGla (panel C). In contrast, ING1b induced nuclear changes reminiscent of apoptosis (panel D) as previously reported by many groups [reviewed in [3]]. We next examined cells for the presence of HP1γ foci, a known SAHF maker [108]. While nuclear HP1γ staining was homogenous in GFP and ING1b infected cells, ING1a infected cells showed distinct HP1γ foci (Figure 6C). These findings, coupled with previous reports showing that inhibiting ING1 gene expression by anti-sense RNA allows primary fibroblasts to undergo several additional population doublings when approaching senescence [8], support the idea that the ING1 family of tumor suppressors contribute to regulating replicative

lifespan and the senescent phenotype. We next examined whether ING1a can induce senescence-associated β-galactosidase activity, a marker shown repeatedly to be upregulated in senescent cells both *in vivo* and *in vitro* (see senescence markers in the introduction section). High enzymatic activity can be observed in ING1a-infected cells, compared to GFP or ING1b-infected cells (Figure 8A). The staining observed in GFP and ING1b-infected cells may be due to the heterogeneous cell population infected. Moreover, compared to fusiform non-infected cells, ING1a induced senescence large flattened cell morphology (Figure 8B). Overall, the data suggest a causal role of ING1a in mediating senescence.

ING1 isoforms have different effects upon the cell cycle.

Overexpression of ING1b has been shown in several reports to induce apoptosis as evidenced by a substantial sub-G1 peak in flow cytometry analysis [15,57]. Examining the effect of ING1a on cell cycle distribution showed that overexpression of INGla induced cell cycle arrest and the percentage of primary Hs68 cells in the G1 phase increased progressively with time despite culturing in complete medium (Figure 9A). To examine whether ING1a also induced apoptosis, we used an Annexin V assay. While the percentage of apoptotic cells increased dramatically over a 48 hr time course in cells transfected with ING1b construct, the percentage of apoptotic cells did not show a significant increase in response to transfected ING1a construct (Figure 9B). This is consistent with INGlb inducing the hyperacetylation of histones H3 and H4 [18] since increased histone acetylation can induce apoptosis [169]. These biochemical data corroborate the effects of the two ING1 isoforms upon cell morphology and, since senescent cells are known to be resistant to the induction of apoptosis, it suggested that INGla and INGlb might antagonize each other functionally as previously suggested for activation of p53 by ING proteins [170].

Interaction between INGl isoforms, p53 and Rb senescence pathways.

The activity of Rb [171] and p53 [92] increases during senescence and is thought to help maintain the state of senescence through effects upon chromatin, including the formation of SAHF [108]. To test whether the INGl isoforms interact directly with these senescence

pathways, we asked whether they differentially affected the expression of the pl6 and p21 cyclin-dependent kinase inhibitors (CDIs). p21 is a transcriptional target of p53 and is induced early in senescence, while pl6 helps maintain Rb in an active, hypophosphorylated state, and increases in amount throughout the course of cell senescence [56,101,102]. We did not detect reproducible increases in p21 levels upon ING1a overexpression in agreement with a previous report [13], however, ING1a expression substantially increased both pl6 and pRb levels (Figure 10). This is of particular interest since the Rb pathway has been directly implicated in the formation of SAHF [108,116,117] supporting the idea that ING1a promotes induction of the senescent phenotype through its impact on chromatin remodeling, in part via increasing Rb levels.

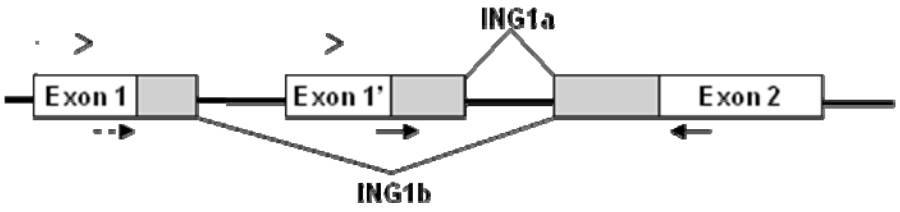
ING1a alters the gene expression profile of human fibroblasts

To assess whether ING1a can affect gene regulation, high-throughput gene expression analyses were performed with normal WI38 fibroblasts infected with replication-deficient adenoviral vectors encoding either GFP alone of GFP and ING1a. The cDNA was fluorescently labelled with Cy3 and Cy5 and *vice versa* (dye swap) to minimize the bias in incorporating different dyes in the reverse transcription reaction. Fluorescent-labeled cDNA probes were combined and hybridized to a single array containing 14,000 genes printed in duplicate. Detection of hybridized probes was achieved by laser excitation of the individual fluorescent markers followed by scanning using a confocal laser scanning microscope. In this study, 2 and -2 were set as the cutoff values for mean fold change (MFC). Tables 2 and 3 list the genes that appear to be regulated by the ING1a protein. In total, 13 (0.10%) and 176 (1.25%) genes displayed relatively higher and lower expression levels, respectively, This observation is in agreement with the preferential association of ING1a with the HDAC1 complex which is correlated with gene repression. Induction of Rb and p16 was not detected in the microarray data. This may be due to the fact that gene array experiments are frequently less sensitive that individual, direct assays of expression since they rely upon hybridization of a complex probe. The Panther database was used to classify ING1a-regulated genes into pathways. Tables 4 and 5 show a list of pathways affected by ING1a overexpression.

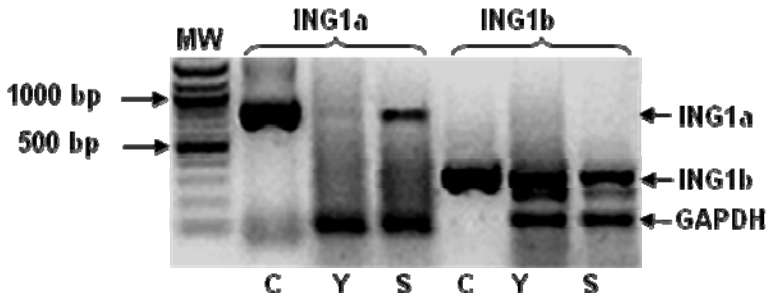
Figure 5: ING1 splicing isoforms are differentially expressed during cellular senescence

A. The two predominant isoforms of ING1 (INGla and INGlb) arise from alternative splicing. Regions that primers were designed to specifically amplify ING1 isoforms are shown on Exon 1 (dashed arrow) and Exon 1' (solid arrow). **B.** RNA from young (Y) or senescent (S) human fibroblasts was reverse-transcribed and the cDNA was analyzed by PCR to determine relative amounts of mRNA in harvested samples. GAPDH was used as an internal control to quantitate RNA isolation and PCR amplification efficiency. Dilutions of ING1 expression plasmids were used as positive controls to verify amplification products (C lanes). This experiment has been contributed by Svitlana Pastyryeva. **C.** Western blots of human fibroblast lysates from young log phase (Y), three-day serum-deprived quiescent (Q) and senescent (S) cells.

5 A


ING1a
Exon 1
Exon 1'
Exon 2
ING1b

5 B


MW
ING1a
ING1b
1000 bp
500 bp
ING1a
ING1b
GAPDH
C Y S C Y S

5 C

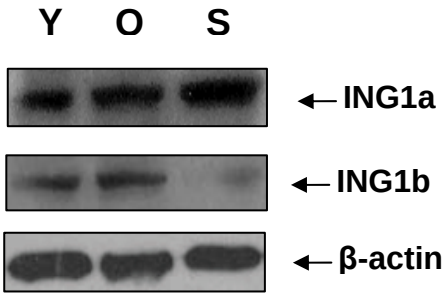

Y O S
ING1a
ING1b
β-actin

Figure 6: Similar levels of overexpression of ING1 splice isoforms

Blotting of lysates of human fibroblasts overexpressing GFP alone or together with ING1b or ING1a showed that the pan-specific anti-ING1 antibody (CAbs 1-4) used recognizes both forms of ING1 with similar efficiency. This is expected since the antibody is raised against a common region to ING1a and ING1b. The lower panel is a Commassie blue staining indicating equal protein loading. The second band in the ING1a lane is commonly observed when this isoform is overexpressed, which may be a degradation product or a post-translationally modified form of ING1a.

6

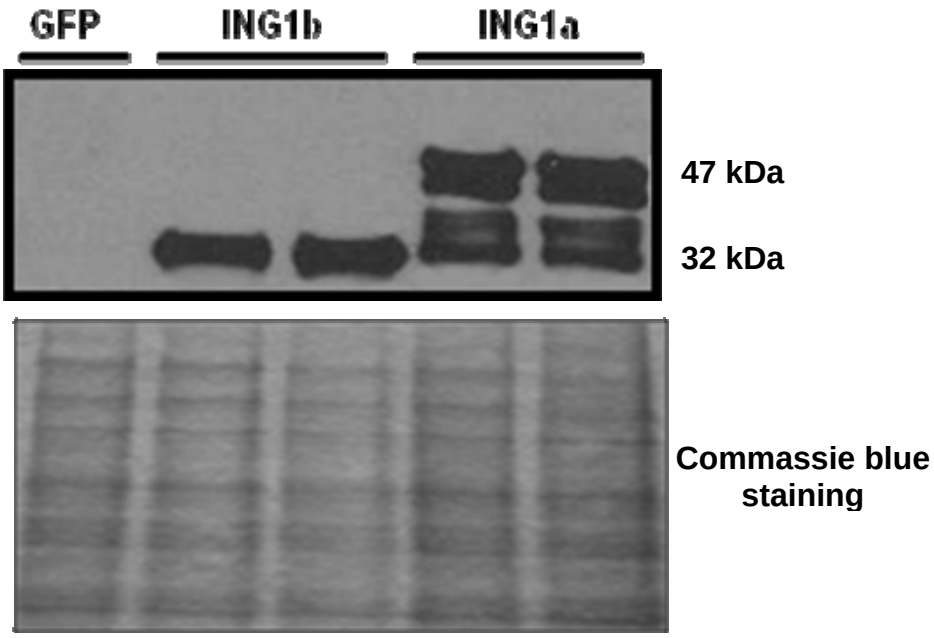
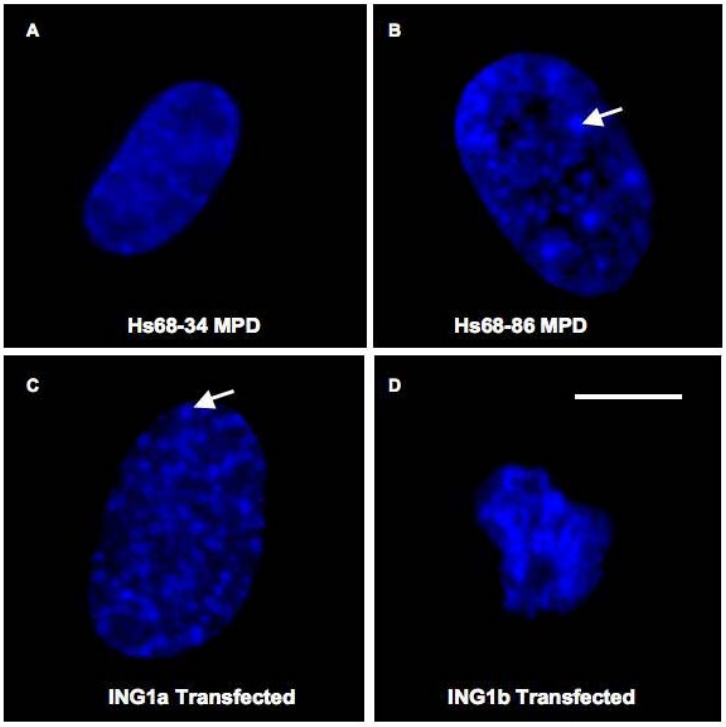


Figure 7: Morphology of cells overexpressing INGla versus INGlb

A. Representative nuclei from young (A), senescent (B), ING1a infected (C) and ING1b infected (D) cells, respectively. White arrows highlight SAHF in panel B and INGla induced heterochromatic foci in panel C. The bars represent 20 μM. **B.** Low passage human fibroblasts infected with GFP, GFP plus ING1b or GFP plus ING1a were fixed and stained with HP1γ antibodies. White arrows in the lower panels highlight one of the HP1γ foci seen in response to ING1a, but not to GFP or GFP plus ING1b.

7A


A
Hs68-34 MPD
B
Hs68-86 MPD
C
ING1a Transfected
D
ING1b Transfected

7B

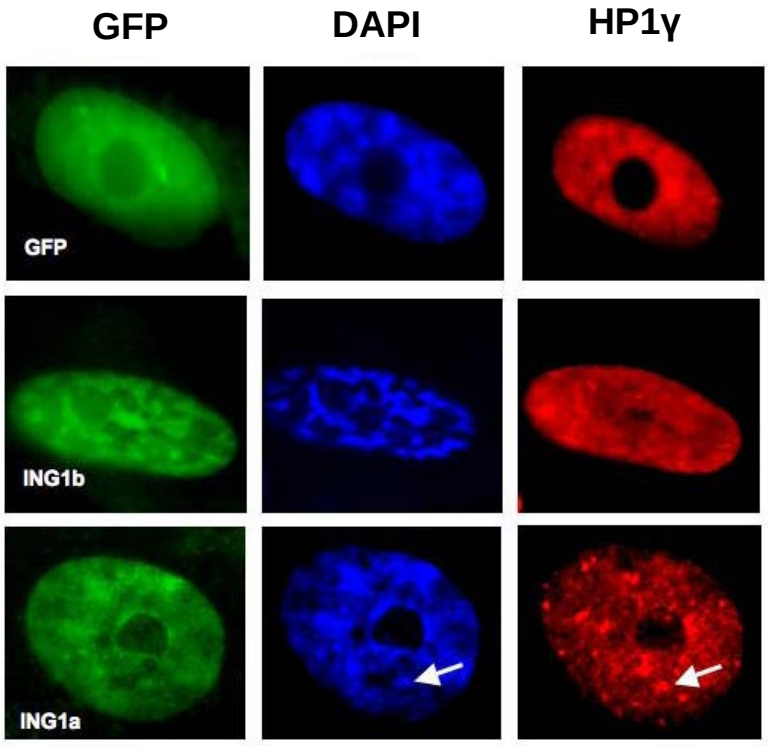
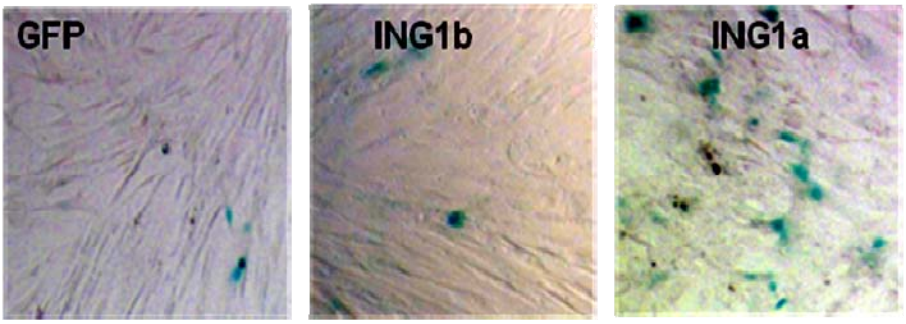


Figure 8: Induction of senescence-associated β-galactosidase activity and senescence cell morphology by ING1a

A. Low passage human fibroblasts were infected with a control GFP, GFP plus ING1a or GFP plus ING1b constructs. 48 hrs later, cells were fixed and stained for SA-β-gal activity. Infection efficiency was monitored by GFP flourescence and did not vary by more than 10% in different plates. **B.** Normarski interference contrast micrographs of human fibroblasts 48 hrs after infection with the indicated constructs.

8A



8B

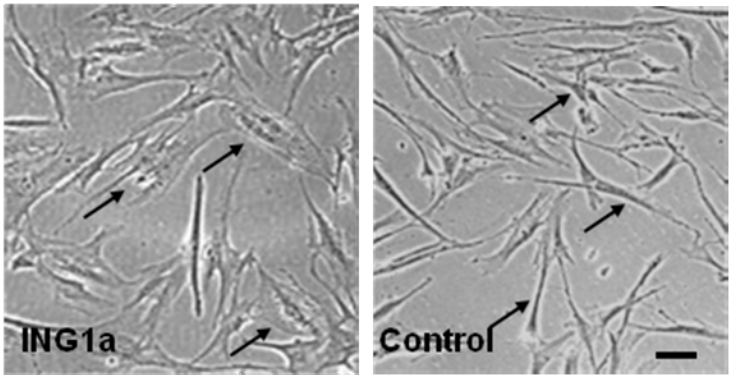
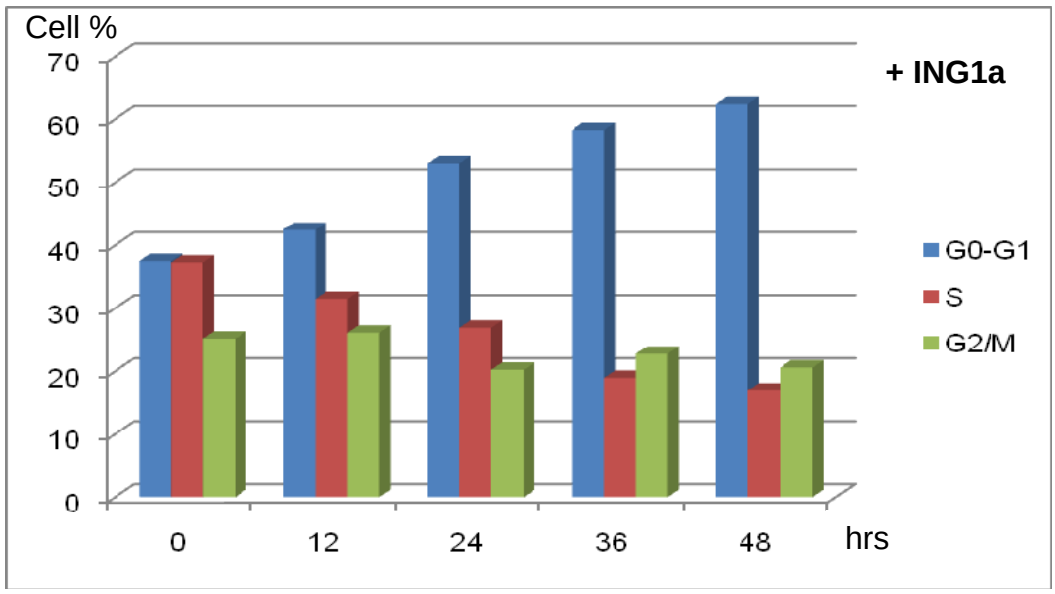


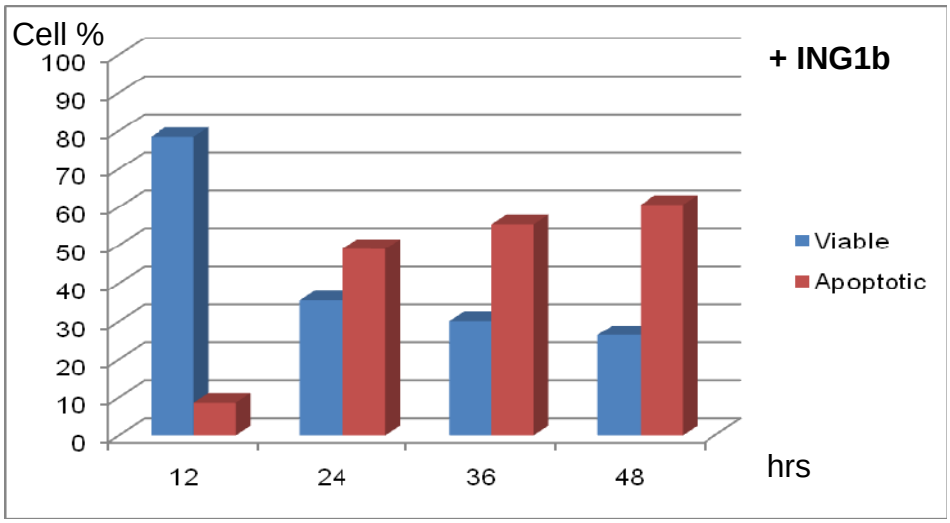
Figure 9: Differential effects of INGl isoforms on cell growth and apoptosis

A. Propidium iodide staining of human fibroblasts overexpressing INGla and analyzed by flow cytometry for DNA content over 48 hrs. The graph shows one of three independent experiments done which gave similar trends. **B.** Annexin V-FITC (BD Biosciences) analysis of cells overexpressing ING1a plus GFP or ING1b plus GFP that were harvested at 12, 24, 36 and 48 hrs. After harvesting, cells were incubated with Annexin V-FITC and propidium iodide for 30 minutes in the dark, followed by flow cytometry. GFP overexpression did not induce apoptosis to levels above background (data not shown). Graphs show the one of the three experiments done which showed similar trends (Annexin V staining was contributed by Philip Berardi).

9A



9B


Cell %
100
90
80
70
60
50
40
30
20
10
0
12
24
36
48
hrs
+ ING1b
Viable
Apoptotic

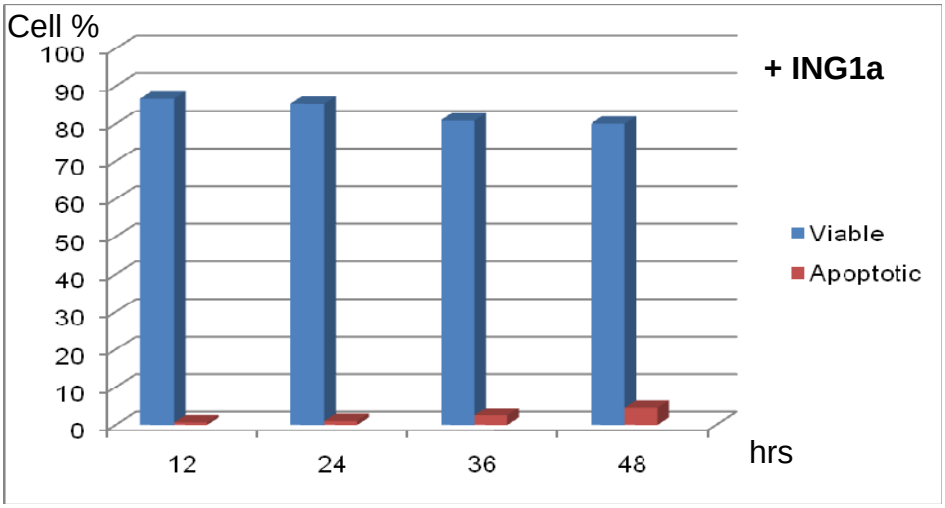

Cell %
100
90
80
70
60
50
40
30
20
10
0
12
24
36
48
hrs
+ ING1a
Viable
Apoptotic

Figure 10: Induction of the p16-pRb pathway by ING1a

A. Low passage human fibroblasts overexpressing GFP, GFP plus ING1a or left untreated (labeled Non) were harvested 48 hrs after infection and lysates were electrophoresed, and blotted with the indicated antibodies. GFP levels were used as a control to monitor infection efficiency, while actin serves as a loading control.

10

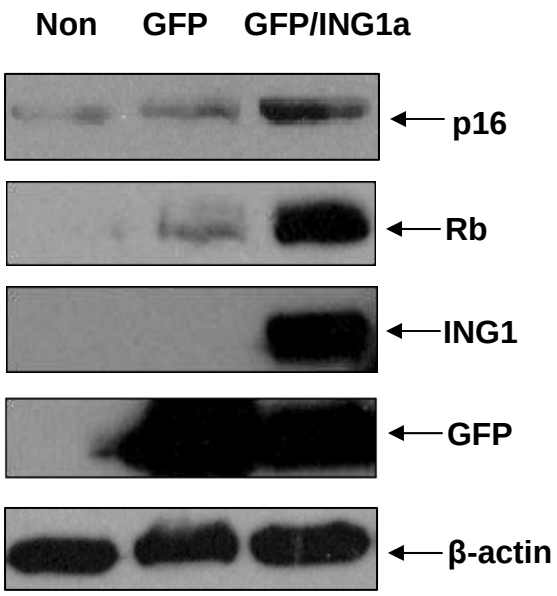


Figure 11: The quality of RNA used in microarray analysis

RNA was extracted from GFP or GFP plus ING1a-infected human diploid fibroblasts using Qiagen kit and was analyzed by Agilent® Bioanalyzer for integrity. The 28S and 18S bands are shown. The ratio of the intensities of the two bands is 1.9 and little degradation is seen, suggesting that the RNA was relatively intact.

11

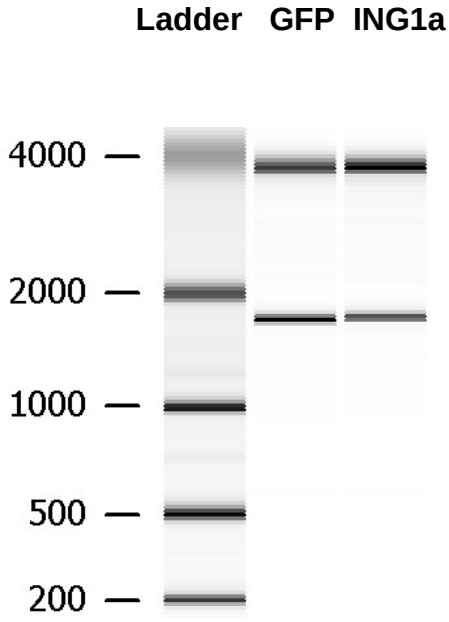


Figure 12: Scheme for the microarray protocol used in this study

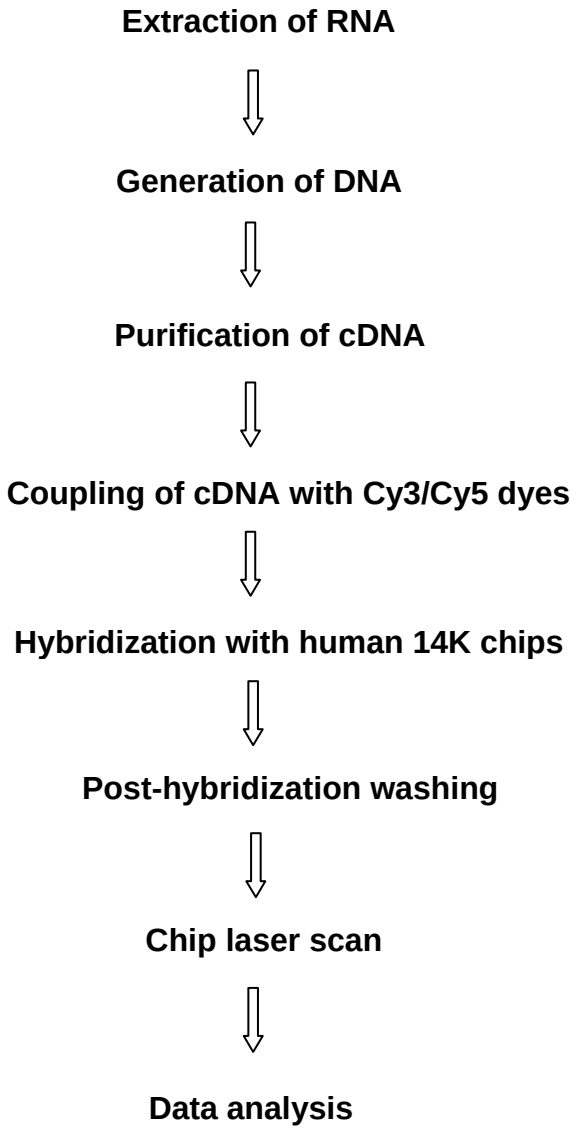


Table 1: List of genes with increased expression in response to ING1a

Gene Name	Gene Symbol	Num. Valid Spots	STDEV	Fold Change
AXL receptor tyrosine kinase	AXL	2	0.03	2.69
Lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	LSS	4	1.24	2.41
Eukaryotic translation termination factor 1	ETF1	4	1.23	2.37
Geranylgeranyl diphosphate synthase 1	GGPS1	4	1.33	2.33
Eukaryotic translation initiation factor 4E family member 2	EIF4E2	3	2.17	2.3
Cadherin 15, M-cadherin (myotubule)	CDH15	2	0.14	2.2
Uroporphyrinogen III synthase (congenital erythropoietic porphyria)	UROS	4	1.3	2.14
LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae)	LSM4	3	2.1	2.1
Matrix metallopeptidase 3 (stromelysin 1, progelatinase)	MMP3	4	1.01	2.08
MRNA; cDNA DKFZp434G1972 (from clone DKFZp434G1972)	NA	2	0.23	2.07
Pleckstrin homology domain containing, family G (with RhoGef domain) member 5	PLEKHG5	2	0.26	2.04
CAMP responsive element binding protein 5	CREB5	2	0.42	2.03
Transmembrane protein with EGF-like and two follistatin-like domains 2	TMEFF2	4	1.86	2.01

Table 2: List of genes with decreased expression in response to ING1a

UniGene Name	Gene Symbol	Num Valid Spots	STDEV	Fold Change
RAS, dexamethasone-induced 1	RASD1	4	0.25	8
Granzyme K (granzyme 3; tryptase II)	GZMK	4	3.08	5.93
Spleen tyrosine kinase	SYK	4	2.72	5.81
Chromosome 10 open reading frame 26	C10orf26	4	1.42	5.66
Regulatory factor X, 2 (influences HLA class II expression)	RFX2	4	2.35	5.54
Phosphoenolpyruvate carboxykinase 2 (mitochondrial)	PCK2	4	3.09	5.46
Fc fragment of IgG, low affinity IIa, receptor (CD32)	FCGR2A	4	2.57	5.31
AarF domain containing kinase 4	ADCK4	4	0.18	4.82
Cyclin-dependent kinase 8	CDK8	4	2.48	4.79
Wolf-Hirschhorn syndrome candidate 2	WHSC2	4	0.48	4.72
LIM domain only 6	LMO6	4	0.21	4.63
Leukocyte specific transcript 1	LST1	4	0.56	4.63
Axin 2 (conductin, axil)	AXIN2	4	0.48	4.47
Neurocan	NCAN	4	2.34	4.41
Apolipoprotein C-II	APOC2	4	2.71	4.41
Fasciculation and elongation protein zeta 1 (zygin I)	FEZ1	4	0.24	4.1
Mevalonate kinase (mevalonic aciduria)	MVK	4	2.08	4.1
NK2 transcription factor related, locus 5 (Drosophila)	NKX2-5	4	0.54	3.97
Glycine cleavage system protein H (aminomethyl carrier)	GCSH	4	0.48	3.92
Similar to cytoplasmic beta-actin	LOC648740	4	2.61	3.78
KH-type splicing regulatory protein (FUSE binding protein 2)	KHSRP	4	0.72	3.76
Insulinoma-associated 1	INSM1	4	2.74	3.73
Cytochrome P450, family 2, subfamily F, polypeptide 1	CYP2F1	4	2.42	3.68
RAB3A interacting protein (rabin3)-like 1	RAB3IL1	4	0.44	3.66
T-box, brain, 1	TBR1	4	1.52	3.66
SRY (sex determining region Y)-box 30	SOX30	4	0.35	3.61
Mitogen-activated protein kinase kinase kinase kinase 1	MAP4K1	4	1.98	3.61
Retinoblastoma-like 2 (p130)	RBL2	4	2.61	3.61
MRNA full length insert cDNA clone EUROIMAGE 1509279	NA	4	2.21	3.56
TAO kinase 2	TAOK2	4	0.74	3.51
Heparan sulfate proteoglycan 2	HSPG2	4	1.94	3.51
Dynamin 3	DNM3	4	2.56	3.48

UniGene Name	Gene Symbol	Num Valid Spots	STDEV	Fold Change
Clone 24627 mRNA sequence	NA	4	2.68	3.46
TRNA methyltransferase 6 homolog (S. cerevisiae)	CGI-09	4	0.41	3.36
Transcription elongation factor A (SII), 2	TCEA2	4	2.11	3.34
Eukaryotic translation initiation factor 4E binding protein 1	EIF4EBP1	4	0.13	3.32
Intestinal cell (MAK-like) kinase	ICK	4	1.94	3.3
Calcium channel, voltage-dependent, beta 3 subunit	CACNB3	4	1.83	3.27
Mitogen-activated protein kinase kinase kinase kinase 5	MAP4K5	4	2.21	3.27
Polymerase (DNA directed), beta	POLB	2	1.17	3.23
Lysyl-tRNA synthetase	KARS	4	1.91	3.14
Chaperonin containing TCP1, subunit 7 (eta)	CCT7	4	1.88	3.1
Nipsnap homolog 3B (C. elegans)	NIPSNAP3B	4	0.44	3.1
NK2 homeobox 1	TITF1	4	0.17	3.1
Component of oligomeric golgi complex 3	COG3	4	2.2	3.1
RAB3 GTPase activating protein subunit 2 (non-catalytic)	RAB3GAP2	4	2.04	2.99
Major histocompatibility complex, class II, DR beta 5	HLA-DRB5	4	2.74	2.99
Aldehyde dehydrogenase 9 family, member A1	ALDH9A1	4	1.64	2.99
Hypothetical protein BC004337	LOC90826	4	0.25	2.83
Tumor protein D52-like 2	TPD52L2	4	0.57	2.79
Male-specific lethal 3-like 1 (Drosophila)	MSL3L1	4	1.99	2.79
Protein tyrosine phosphatase-like A domain containing 1	PTPLAD1	4	1.07	2.77
Msh homeobox 1	MSX1	4	2.17	2.75
Nuclear receptor subfamily 1, group D, member 1	NR1D1	2	0.21	2.75
Spectrin repeat containing, nuclear envelope 2	SYNE2	4	0.12	2.73
Coiled-coil domain containing 52	CCDC52	4	0.25	2.73
Serine/threonine kinase 25 (STE20 homolog, yeast)	STK25	4	1.67	2.71
Glycine receptor, alpha 1 (startle disease/hyperekplexia)	GLRA1	4	1.78	2.71
Cyclin E1	CCNE1	4	1.97	2.7
Phospholipase C, delta 1	PLCD1	4	1.64	2.7
PMS2 postmeiotic segregation increased 2 (S. cerevisiae)	PMS2	4	3.18	2.68
Polybromo 1	PBRM1	4	0.98	2.64
Replication protein A3, 14kDa	RPA3	4	1.56	2.64
Ephrin-A5	EFNA5	4	2.05	2.64
Intraflagellar transport 140 homolog (Chlamydomonas)	IFT140	4	0.23	2.62
Tripartite motif-containing 13	TRIM13	4	1.69	2.6
Acetyl-Coenzyme A acetyltransferase 2 (acetoacetyl Coenzyme A thiolase)	ACAT2	4	1.33	2.6
Gonadotropin-releasing hormone receptor	GNRHR	4	1.76	2.58

UniGene Name	Gene Symbol	Num Valid Spots	STDEV	Fold Change
Gem (nuclear organelle) associated protein 4	GEMIN4	4	0.56	2.57
Glutamate receptor, ionotropic, N-methyl D-aspartate 1	GRIN1	4	0.35	2.57
Complement component 1, q subcomponent-like 1	C1QL1	4	0.28	2.55
TruB pseudouridine (psi) synthase homolog 2 (E. coli)	TRUB2	4	0.65	2.53
CD1d molecule	CD1D	4	1.4	2.51
Signal transducer and activator of transcription 4	STAT4	4	2.25	2.51
Calcium channel, voltage-dependent, alpha 2/delta subunit 1	CACNA2D1	4	1.51	2.51
8-oxoguanine DNA glycosylase	OGG1	4	1.46	2.51
KIAA0776	KIAA0776	4	0.52	2.48
SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	SMARCC1	4	1.43	2.48
5-methyltetrahydrofolate-homocysteine methyltransferase	MTR	4	3.05	2.48
Glioma-associated oncogene homolog 1 (zinc finger protein)	GLI1	4	2.4	2.48
Hypermethylated in cancer 2	HIC2	4	2.24	2.46
GIPC PDZ domain containing family, member 2	GIPC2	4	1.37	2.46
Ciliary neurotrophic factor receptor	CNTFR	4	2.84	2.46
Basal cell adhesion molecule (Lutheran blood group)	BCAM	4	0.35	2.46
Coagulation factor XI (plasma thromboplastin antecedent)	F11	4	1.69	2.43
Bone morphogenetic protein 7 (osteogenic protein 1)	BMP7	4	2.14	2.43
Mediator complex subunit 23	CRSP3	4	0.5	2.39
Coenzyme Q3 homolog, methyltransferase (S. cerevisiae)	COQ3	4	0.15	2.39
Prostaglandin F receptor (FP)	PTGFR	4	1.4	2.38
Retinoic acid induced 1	RAI1	4	0.28	2.38
Patched homolog 1 (Drosophila)	PTCH1	4	1.85	2.38
Phosphatidylinositol-4-phosphate 5-kinase, type I, alpha	PIP5K1A	4	0.73	2.38
VAMP (vesicle-associated membrane protein)-associated protein B and C	VAPB	4	0.84	2.36
Ribosomal protein L39	RPL39	4	0.22	2.36
Adenosine A2b receptor	ADORA2B	4	1.53	2.36
Transforming growth factor beta regulator 4	TBRG4	4	1.54	2.35
Myosin IIIA	MYO3A	4	0.57	2.35
Phosphoinositide-3-kinase, class 2, beta polypeptide	PIK3C2B	4	0.23	2.35
Developmental pluripotency associated 4	DPPA4	4	0.21	2.33
EGF-like repeats and discoidin I-like domains 3	EDIL3	4	0.98	2.33
Wilms tumor 1	WT1	4	1.28	2.33
Dipeptidase 1 (renal)	DPEP1	4	1.2	2.31
REX1, RNA exonuclease 1 homolog (S. cerevisiae)	REXO1	4	1.96	2.3
Zinc finger protein 638	ZNF638	4	1.39	2.3

UniGene Name	Gene Symbol	Num Valid Spots	STDEV	Fold Change
Galanin receptor 1	GALR1	4	1.24	2.28
Dystroglycan 1 (dystrophin-associated glycoprotein 1)	DAG1	4	1.34	2.27
CD244 molecule, natural killer cell receptor 2B4	CD244	4	0.25	2.25
Chromosome 11 open reading frame 30	C11orf30	4	0.25	2.22
CDC42 small effector 2	CDC42SE2	4	1.23	2.2
MRS2-like, magnesium homeostasis factor (S. cerevisiae)	MRS2L	4	2.46	2.19
Olfactomedin 2	OLFM2	4	0.24	2.19
Squalene epoxidase	SQLE	4	1.19	2.19
Calcium channel, voltage-dependent, T type, alpha 1I subunit	CACNA1I	4	0.34	2.17
Adenylosuccinate lyase	ADSL	2	0.14	2.14
Ubiquitin associated protein 2-like	UBAP2L	4	0.87	2.13
Apolipoprotein E	APOE	4	0.11	2.13
Calmodulin binding transcription activator 2	CAMTA2	4	0.1	2.1
Natriuretic peptide receptor A/guanylate cyclase A (atrionatriuretic peptide receptor A)	NPR1	4	1.67	2.1
Chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant)	CXCL13	4	0.38	2.08
ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2	ST8SIA2	4	0.82	2.08
Pancreatic and duodenal homeobox 1	PDX1	4	1.58	2.08
Sorting nexin 1	SNX1	4	0.94	2.06
Uncoupling protein 2 (mitochondrial, proton carrier)	UCP2	4	2.36	2.06
Neurturin	NRTN	4	0.4	2.04
Basonuclin 2	BNC2	2	0.19	2.03
F-box and WD repeat domain containing 2	FBXW2	4	0.26	2.03
Neuregulin 2	NRG2	4	0.53	2.03
RNA binding motif protein, X-linked 2	RBMX2	4	0.28	2.03
Cytochrome b5 reductase 1	CYB5R1	4	0.3	2.03
Opioid growth factor receptor	OGFR	4	0.42	2.01
Peptidase (mitochondrial processing) alpha	PMPCA	4	1.13	2.01
Clone 23629 mRNA sequence	NA	4	0.65	2.01
Poly(rC) binding protein 2	PCBP2	2	0.13	2.01
Tumor necrosis factor receptor superfamily, member 13B	TNFRSF13B	4	0.17	2
CDNA FLJ20812 fis, clone ADSE01316	NA	4	2.14	2
L antigen family, member 3	LAGE3	4	1.78	2

Table 3: List of pathways encompassing genes upregulated in response to ING1a (classification was done using the Panther database)

Pathway	Num. of genes upregulated by ING1a
Unclassified	4
Heme biosynthesis	1
Plasminogen activating cascade	1
Axon guidance mediated by semaphorins	1
p38 MAPK pathway	1
Ubiquitin proteasome pathway	1
Cytoskeletal regulation by Rho GTPase	1
Cadherin signaling pathway	1
G-protein signaling pathway	1
Wnt signaling pathway	1

Table 4: List of pathways encompassing genes downregulated in response to ING1a (classification was done using the Panther database)

Pathway	Num. of genes downregulated by ING1a
Unclassified	111
p53 pathway feedback loops 2	3
p53 pathway	3
EGF receptor signaling pathway	3
Inflammation mediated by chemokine and cytokine signaling pathway	3
Cell cycle	2
Hedgehog signaling pathway	2
Thyrotropin-releasing hormone receptor signaling pathway	2
Oxytocin receptor mediated signaling pathway	2
Metabotropic glutamate receptor group III pathway	2
5HT2 type receptor mediated signaling pathway	2
Insulin/IGF pathway-protein kinase B signaling cascade	2
Endothelin signaling pathway	2
Apoptosis signaling pathway	2
FGF signaling pathway	2
Integrin signalling pathway	2
Heterotrimeric G-protein signaling pathway-Gi alpha and Gs alpha mediated pathway	2
Methionine biosynthesis	1
S adenosyl methionine biosynthesis	1
Formyltetrahydroformate biosynthesis	1
Cholesterol biosynthesis	1
Pyruvate metabolism	1
5HT3 type receptor mediated signaling pathway	1
Wnt signaling pathway	1
Circadian clock system	1
JAK/STAT signaling pathway	1
Angiogenesis	1
p53 pathway by glucose deprivation	1
De novo purine biosynthesis	1
Cortocotropin releasing factor receptor signaling pathway	1
Synaptic_vesicle_trafficking	1
Hypoxia response via HIF activation	1
PDGF signaling pathway	1
Axon guidance mediated by netrin	1
Metabotropic glutamate receptor group I pathway	1

Pathway	Num. of genes downregulated by ING1a
Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	1
Interleukin signaling pathway	1
Ionotropic glutamate receptor pathway	1
Beta2 adrenergic receptor signaling pathway	1
Beta1 adrenergic receptor signaling pathway	1
p38 MAPK pathway	1
Metabotropic glutamate receptor group II pathway	1
Muscarinic acetylcholine receptor 2 and 4 signaling pathway	1
Huntington disease	1
Blood coagulation	1
G-protein signaling pathway	1
Muscarinic acetylcholine receptor 1 and 3 signaling pathway	1
TGF-beta signaling pathway	1
Ubiquitin proteasome pathway	1
T cell activation	1
Alzheimer disease-amyloid secretase pathway	1
Nicotinic acetylcholine receptor signaling pathway	1
VEGF signaling pathway	1
B cell activation	1
Parkinson disease	1

Section II. Prediction and validation of novel ING interacting partners

Pair-wise alignment of YNGs (yeast INGs) and human INGs

The first member of the ING family (ING1) was discovered in humans and subsequently four more ING genes have been identified (ING2-5) [1]. Homologs of ING proteins also exist throughout the animal and plant kingdoms. Our group first reported the existence of functional homologs of ING proteins in yeast [172]. Three yeast proteins, YNG1, YNG2 and YNG3 (also known as Pho23), have been shown to bear considerable homology to the human ING proteins in their C-terminal region. Pairwise alignments between individual yeast and human ING proteins were generated [173]. Sequences of ING1-5 (including all known ING1 isoforms) and YNG1-3 were obtained from the NCBI Genbank database (http://www.ncbi.nlm.nih.gov/). Additionally, CLUSTAL-W [161] and T-COFFEE [174] multiple sequence alignment programs were used to generate multiple sequence alignments, from which we derived the additional pairwise alignments. Although the alignment scores were very close to each other, given the consistency of the results obtained from the various sequence alignment tools used, the following observations can be made: (i) YNG1 shows the highest degree of sequence homology to ING1, (ii) YNG2 shows considerable homology to ING4 and ING5, and (iii) Pho23 and ING3 are similar to each other. These results agree with previous reports of phylogenetic relationships among ING proteins [1] and also with a recent paper which attempts to classify ING proteins with respect to their association with either HAT or HDAC complexes [30].

Conservation of ING domains across species

Since members of the ING family of tumor suppressors show significant sequence conservation from yeast to humans, we proposed that functional interactions might also be conserved. From the available yeast interactome data, it is evident that the yeast counterparts of the ING proteins, also referred to as YNGs, interact with 1,075 other yeast proteins under normal physiological conditions [164]. Although the majority of these interactions have very low probability scores, and hence are likely artefacts of the detection method, several of them may be transient, but nonetheless real interactions. Because of the availability of a large amount of marginal, unanalyzed yeast interaction data [164], there was potentially valuable untapped data to guide selection of human ING-interacting protein candidates. The yeast

dataset has the advantage of being near saturation with regards to interactome coverage, so that almost all real interactions should be detected. The hypothesis was bolstered by the fact that many of the previously validated ING interactions in humans were also present in the yeast interactome data. We attempted to reconcile interactomes from multiple model organisms based on two different approaches: orthology [175] and interaction network topology techniques [176]. Neither provided new insights for novel ING interacting proteins. Given the richness of available yeast data, we designed a new approach to better predict ING interactions. The bioinformatics workflow devised to filter down the massive lists of yeast interactions to a few salient candidates for biochemical validation is illustrated in Figure 13, and can be generalized to be useful for many other proteins. Since conservation of interacting partners is often a function of conservation of domain structure within [177] or across species [178], the next step was to identify the domain structure similarities between human and yeast ING family proteins. In order to characterize possible interaction domains of ING-like proteins quantitatively, we used an iterative consensus building processes. This process consisted of building initial Hidden Markov Models (HMMs), probabilistic process to build a position-specific amino acid substitution model of previously identified domains, such as the leucine zipper-like (LZL) motif, plant homeodomain (PHD), lamin-interaction domain (LID), nuclear localization signal (NLS), and poly basic region (PBR) using the multiple sequence alignments reported in [1]. These domain models were searched against the Uniprot database (http://www.pir.uniprot.org/), which consists of non-redundant protein datasets for all species to identify proteins with domains closely related to the human INGs. All Uniprot proteins matching the human domain models were then added to the original model sequence to make them less species-specific and new HMMs were built based on the expanded list of sequences. This process was repeated until no new Uniprot matches were found. Because Uniprot contains data from many species, the iterative approach is a method to create domain models capturing sequence conservation amongst multiple species. The phylogenetic distribution and consensus sequences for the domains are illustrated in Figure 14, and significance thresholds are discussed.

Since PPIs are primarily based on specific domains, domain-specific models of amino acid substitutions for the various annotated domains of the ING family proteins were

generated. This allowed us to statistically examine the validation of the domain models across species. Results obtained from this method of analysis should be more sensitive and verifiable than the generalized substitution rates used by the current ortholog detection methods based on pairwise alignments. This improved sensitivity may be due to the accounting for the specific evolution of individual protein domains and/or the greater flexibility of HMMs over simple pairwise alignments.

Prediction of human ING interactors

The domain structure models generated using this approach were then used to identify possible ING-like proteins in model species for which interaction data is readily available. We investigated *D. melanogaster*, *C. elegans*, and *S. cerevisiae* interactome data as these species have the most extensively consolidated lists of interactions. We ran the generalized ING domain models against the proteomes of the above mentioned three species to verify if counterparts of human ING protein domains exist in them. The interaction databases to use depend on the conservation breadth of the gene to be investigated. Only one of the domains was found to be conserved in a single protein in *C. elegans* (PHD in Y51H1A.4, human ING1b homolog). Much better conservation of multiple ING domains was observed in *S. cerevisiae* and *D. melanogaster*. We therefore focused on identifying potential ING-interacting proteins in these two species. The domain conservation in these species is illustrated in Figure 15. As expected, the PHD domain is highly conserved across all three, with the highest overall homology with human ING2. Also noteworthy is that the fly ING1 homolog contains an LZL domain, whereas the human version does not. The LID domain is strongly conserved in all humans and fly INGs, but is not present in yeast. Interestingly, the PBR domain is weakly conserved in yeast's ING2 homolog, but not in fly. The inclusion thresholds for each domain model were: PBR 10^{-2}, PHD 10^{-3}, NLS 10^{-6}, LID 10^{-3}, and LZL 10^{-5}. The thresholds represent the weakest e-value for any sequence used to create the domain model, i.e. the lowest score for a known positive example. The different thresholds are a consequence of the varying natures of the domain models (length, amino acid composition, phylogenetic distribution, etc.). With the exception of PBR (10^{-2}), all domain

matches are well above these thresholds. PBR weak score may reflect the fact that the domain model was built using only human sequences (see Figure 14).

Using the taxonomic search tools of MAGPIE, the 1075 yeast genes were filtered to just those with human homologs. This left 381 proteins that both interact with YNGs in yeast, and have human orthologs is shown in Table 5. We reasoned that the probability of the fungal interaction being conserved in humans would be higher for those proteins that show homology in another higher eukaryote because this would be evidence for the maintenance of the interaction in the Metazoan lineage. We filtered the ING-interacting proteins found in both yeast and human against Drosophila database as the Drosophila ING showed high degree of domain conservation with the yeast and human ING proteins. Of the 36 ING-interactors identified in fly by FlyBase [163], only five had strong yeast homologs (e-value < 10^{-35}), and only three of these showed a high degree of sequence conservation in humans. These three fly genes (having putative conserved interacting partners in yeast) have 5 potential homologs in human, namely: hRPC155, PAK1B, MAP3K4 (MEKK4), p38aMAPK, and GSPT1. The Venn diagram in Figure 16 shows the overlapping sets of potential ING-interacting proteins in fly, yeast and human.

Comparison to existing datasets and methods

To evaluate the combined contribution of Krogan *et al* marginal data [164] and our prediction technique to the study of protein-protein interactions, we compared our results to those obtained from biochemical surveys, and other prediction algorithms [143]. Through an evaluation of the completeness of current yeast and human protein-protein interaction networks, making raw unfiltered results available to all researchers could help distinguish between real and spurious interactions. Table 6 summarizes the few ING interactions we could extract from the publicly available datasets surveyed by Hart *et al.* [143], covering most of the commonly used techniques, from yeast two-hybrid to tandem affinity purification. Truly unfiltered experimental data is only available for Krogan *et al.* [164], therefore we searched for how well observed/predicted ING interactions matched our technique's criteria and to what extent the unfiltered Krogan dataset and the other datasets overlapped with respect to data involving ING. It is clear from the table that YNG/ING is

poorly represented in several datasets, and that different methods produced different biases in which ING is detected. This supports the proposition even the well-studied yeast interactome is only about 50% elucidated by existing, filtered datasets [143]. Krogan *et al.* dataset included all of the ING data presented in the other databases. For this reason, we did not need to screen against any other datasets. For other protein families, one can perform the same set comparison, in order to decide if complete interaction coverage requires including more than one of the datasets from Table 6. Based on our success in identifying true ING interaction from Krogan *et al.* unfiltered dataset, we strongly agree with that the research community would be much better served by the release of raw interaction datasets in general for comparison and consolidation.

Other bioinformatics-based approaches have been used to predict interactions between proteins (for a review, see [179]). Sequence, domain, and motif structure based approaches form the basis of Bayesian network models [180]. Examining co-evolution of interacting proteins by comparing phylogenetic trees [181], correlating mutations [182], or gene fusion [183] also rely on sequence based approaches. Protein domain interface-based approaches also exist [184]. Other approaches such as gene expression, gene ontology annotations, and transcriptional regulation, can also be used to predict whether or not a group of proteins are members of the same complex. Our attempts to use conventional protein-protein interaction prediction tools on ING and YNG proteins did not yield results beyond those described in the various public interaction repositories as listed in [185] or predicted by literature text mining. An exhaustive comparison of our technique to others is beyond the scope of this study, but Table 7 summarizes the results of searching for ING1/2/3 interactions with a variety of tools employing various techniques.

It must be noted that our core predictions (hRPC155, PAK1B, MAP4K3, p38MAPK and GSPT1) do not overlap with other ortholog-based techniques [181,186], which would be the most reasonable comparison to make. Interestingly, our two marginal predictions, PDI1 and CDC37, concur with some methods in Table 7. The fact that none of the core predictions overlaps, but marginal ones do highlights the fact that different techniques were used to define orthologs. All of the methods in Table 7 used either InParanoid [187] or Homologene [188] to define interspecies gene mappings. The former maps only YNG1 and Pho23 to the

human counterparts we have identified, while the latter maps all three yeast ING equivalents to ING3. The uniqueness of our core predictions suggests that the technique we have developed provides added value over a straightforward multi-species prediction tools. Given an unfiltered dataset, it is possible that some of the techniques used in Table 7 that employ existing biochemical data would also predict some or all of our five candidate interactions. However, we are unaware of any follow-up studies by the authors of those tools using a raw dataset. It is not unreasonable to assume that the level of false positive predictions from these tools would increase substantially without some changes to their algorithms, which were built for "clean" input datasets. We do not suggest that our technique will find all true positives, because interactions are not always shared between multiple species, and not all interactions have been elucidated. Rather, our technique provides guidance for researchers working on proteins whose interactions are not successfully predicted using existing techniques.

Biochemical validation of potential human ING protein interactions

In order to select candidates for biochemical validation of human ING interactors, we compared our data with experimentally validated ING interactions in human, as listed in the STRING database [189]. Nine of the ten experimentally validated human ING interactors with yeast homologs in the 381 gene list had extremely weak interactions ($p < 0.017$). It therefore seems reasonable to biochemically validate any of the five potential human homologs, even though they had similarly low probability scores according to available yeast data. The fact that none of the five candidate human homologs were found in the validated list of ING interaction from the STRING database is not surprising, since the human interactome dataset is at present not nearly as saturated as that for yeast.

To restrict the list of five candidates further, we considered the biological relevance of the potential interactions to the known functions of INGs. Accordingly, the choice was amongst PAK1b, MAP3K4 (MEKK4) and p38MAPK in descending order of homology among the 3 species (e-values 10^{-98}, 10^{-45}, and 10^{-26} respectively). We wanted to test if in such cases the yeast data could be used alone in successfully predicting human ING interactors. Based on reagent availability, and scientific relevance of the interactions (see the

discussion section), we chose to biochemically validate three putative interactions (MEKK4 and p38MAPK) using co-immunoprecipitation (Co-IP) followed by western blot analysis. Since the two proteins that we chose to investigate, MEKK4 and p38MAPK are closely linked in a stress pathway [190], we chose to confirm these interactions under both normal and stress conditions approaches, namely UV. Figures 17 and 18 show that overexpressed p38MAPK and overexpressed MEKK4 showed strong signals in ING1 immunoprecipitates, but not in the negative control glutathione-S-transferase (GST) immunoprecipitates. Figure 19 shows that ING1 immunocomplexes from untransfected cells, but not GST immunocomplexes contain both p38MAPK and MEKK4, confirming that this interaction occurs between endogenous proteins. Unlike the case for ING1-PCNA interactions that are increased by UV-induced stress [15], treatment of cells with a UV dose sufficient to induce a stress response did not markedly alter the degree of kinase-ING1 interaction. Input lanes also show that robust signals were obtained for both the p38MAPK and MEKK4 proteins in control western blots of lysates used, under conditions where little, if any, signal was seen for ING1. This indicates that high levels of the kinases are expressed in our experimental cell system, compared to ING1 protein. Relatively high expression levels of the kinases compared to ING1 likely explains why reciprocal IP-western assays did not clearly demonstrate a detectable interaction (data not shown), since only a small portion of the kinases would be expected to interact with ING1 protein based upon their apparent relative stoichiometry.

Figure 13: General workflow for identifying possible protein interaction candidates for a given gene of interest

The particular databases and genes used in our study are given in italics. Data in figures 13-20 and Tables 5-7 has been generated in collaboration with Dr. Christoph Sensen lab and includes contributions from Paul Gordon.

13

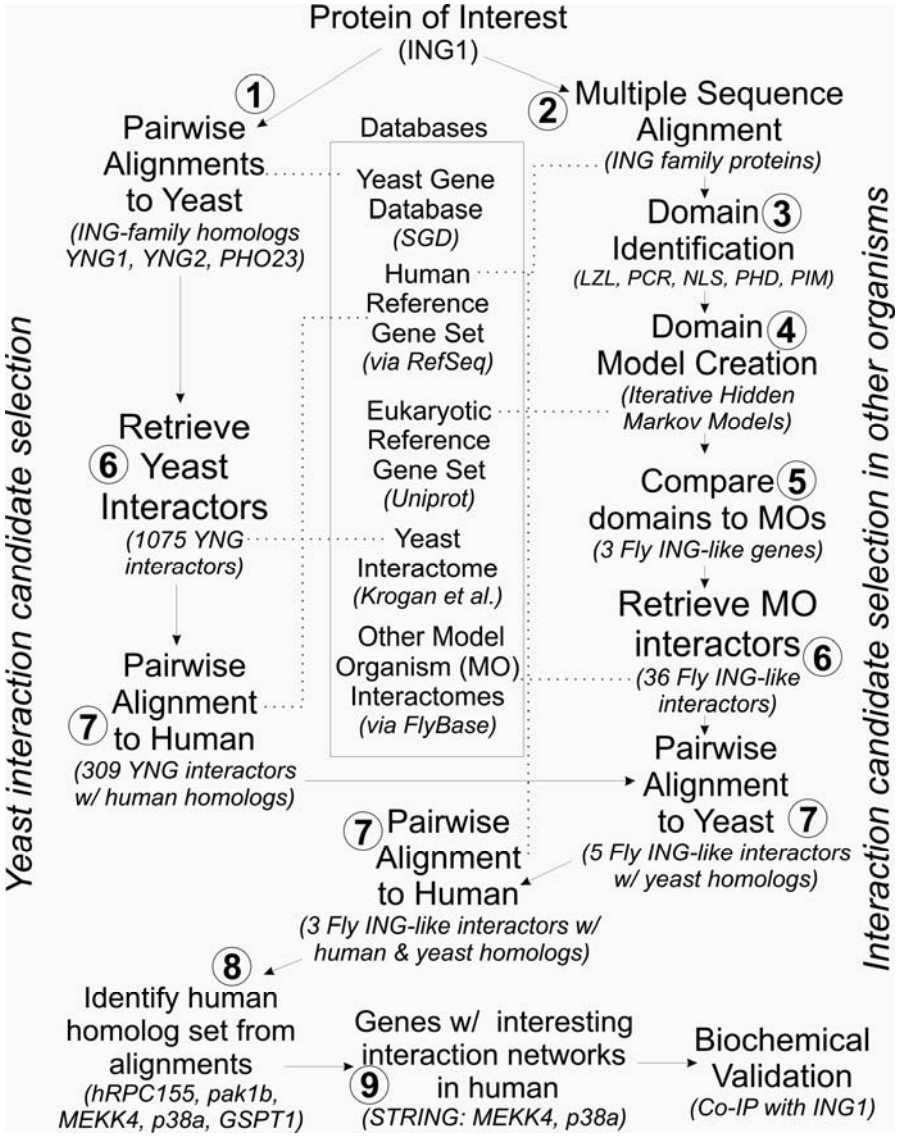

Protein of Interest
(ING1)
1
Pairwise Alignments to Yeast
(ING-family homologs YNG1, YNG2, PHO23)
2
Multiple Sequence Alignment
(ING family proteins)
Databases
Yeast Gene Database (SGD)
Human Reference Gene Set (via RefSeq)
Eukaryotic Reference Gene Set (Uniprot)
Yeast Interactome (Krogan et al.)
Other Model Organism (MO) Interactomes (via FlyBase)
Domain 3 Identification
(LZL, PCR, NLS, PHD, PIM)
Domain 4 Model Creation
(Iterative Hidden Markov Models)
Compare 5 domains to MOs
(3 Fly ING-like genes)
Retrieve MO interactors 6
(36 Fly ING-like interactors)
Retrieve 6 Yeast Interactors
(1075 YNG interactors)
Pairwise Alignment 7 to Human
(309 YNG interactors w/ human homologs)
Pairwise Alignment to Yeast 7
(5 Fly ING-like interactors w/ yeast homologs)
7 Pairwise Alignment to Human
(3 Fly ING-like interactors w/ human & yeast homologs)
8
Identify human homolog set from alignments
(hRPC155, pak1b, MEKK4, p38a, GSPT1)
Genes w/ interesting interaction networks in human 9
(STRING: MEKK4, p38a)
Biochemical Validation
(Co-IP with ING1)
Yeast interaction candidate selection
Interaction candidate selection in other organisms

Figure 14: ING domains conserved in different species

Domain consensus and phylogenetic distribution of ING domains, based on iteratively-built Hidden Markov Models, seeded with human ING proteins multiple sequence alignments**.** The height of letters represents the degree of conservation.

14

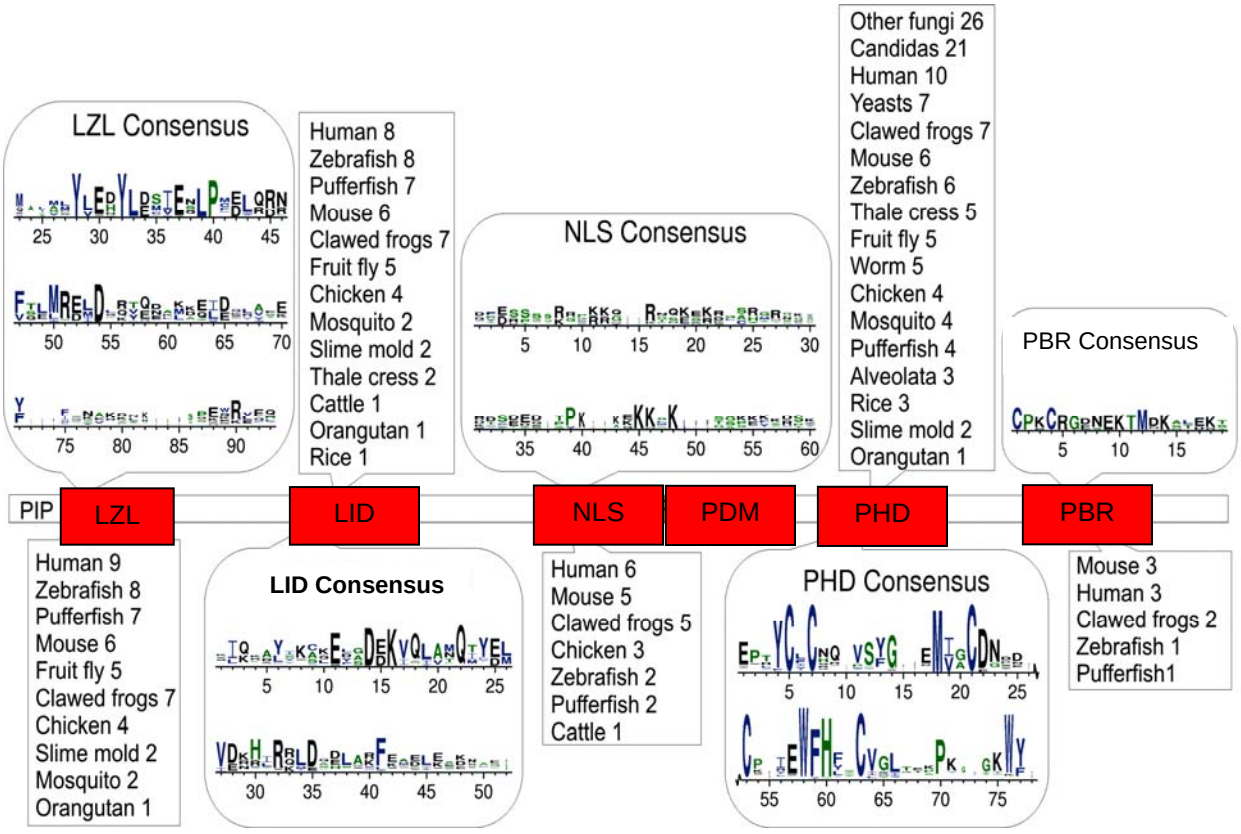

LZL Consensus
Human 8
Zebrafish 8
Pufferfish 7
Mouse 6
Clawed frogs 7
Fruit fly 5
Chicken 4
Mosquito 2
Slime mold 2
Thale cress 2
Cattle 1
Orangutan 1
Rice 1
NLS Consensus
Other fungi 26
Candidas 21
Human 10
Yeasts 7
Clawed frogs 7
Mouse 6
Zebrafish 6
Thale cress 5
Fruit fly 5
Worm 5
Chicken 4
Mosquito 4
Pufferfish 4
Alveolata 3
Rice 3
Slime mold 2
Orangutan 1
PBR Consensus
PIP
LZL
LID
NLS
PDM
PHD
PBR
Human 9
Zebrafish 8
Pufferfish 7
Mouse 6
Fruit fly 5
Clawed frogs 7
Chicken 4
Slime mold 2
Mosquito 2
Orangutan 1
LID Consensus
Human 6
Mouse 5
Clawed frogs 5
Chicken 3
Zebrafish 2
Pufferfish 2
Cattle 1
PHD Consensus
Mouse 3
Human 3
Clawed frogs 2
Zebrafish 1
Pufferfish1

Figure 15: Quantification of the degree of conservation of different ING proteins

The conservation of multiple ING domains amongst three different species increases the confidence in predicting conserved protein interactions. The plant homeodomain (PHD) is the most highly conserved region in the ING family, followed by the lamin-interaction domain (LID).

15

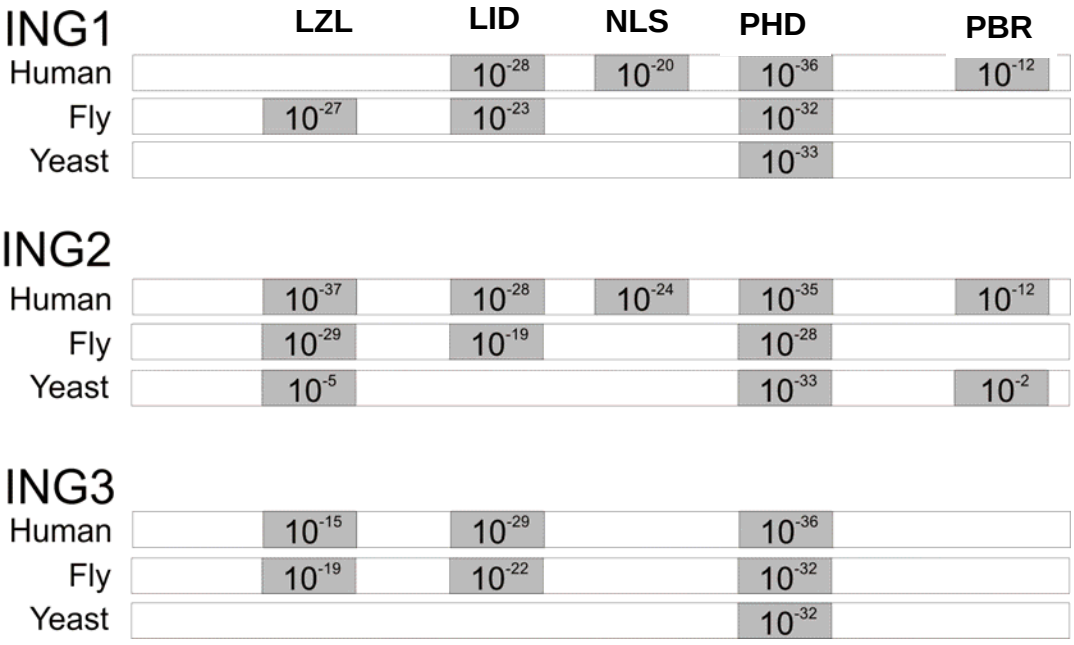

ING1
LZL
LID
NLS
PHD
PBR
Human
10^{-28}
10^{-20}
10^{-36}
10^{-12}
Fly
10^{-27}
10^{-23}
10^{-32}
Yeast
10^{-33}
ING2
Human
10^{-37}
10^{-28}
10^{-24}
10^{-35}
10^{-12}
Fly
10^{-29}
10^{-19}
10^{-28}
Yeast
10^{-5}
10^{-33}
10^{-2}
ING3
Human
10^{-15}
10^{-29}
10^{-36}
Fly
10^{-19}
10^{-22}
10^{-32}
Yeast
10^{-32}

Figure 16: Overlap of the ING interactome datasets for human, fly and yeast

Five potential interactions are shared between fly and yeast, and 3 of the fly interactions involve genes with good human homologs (central nodes). These interactions (3 in fly with 5 equivalents in yeast) are candidates for biochemical validation in human cells; hence the question marks, in humans (in the green circle).

16

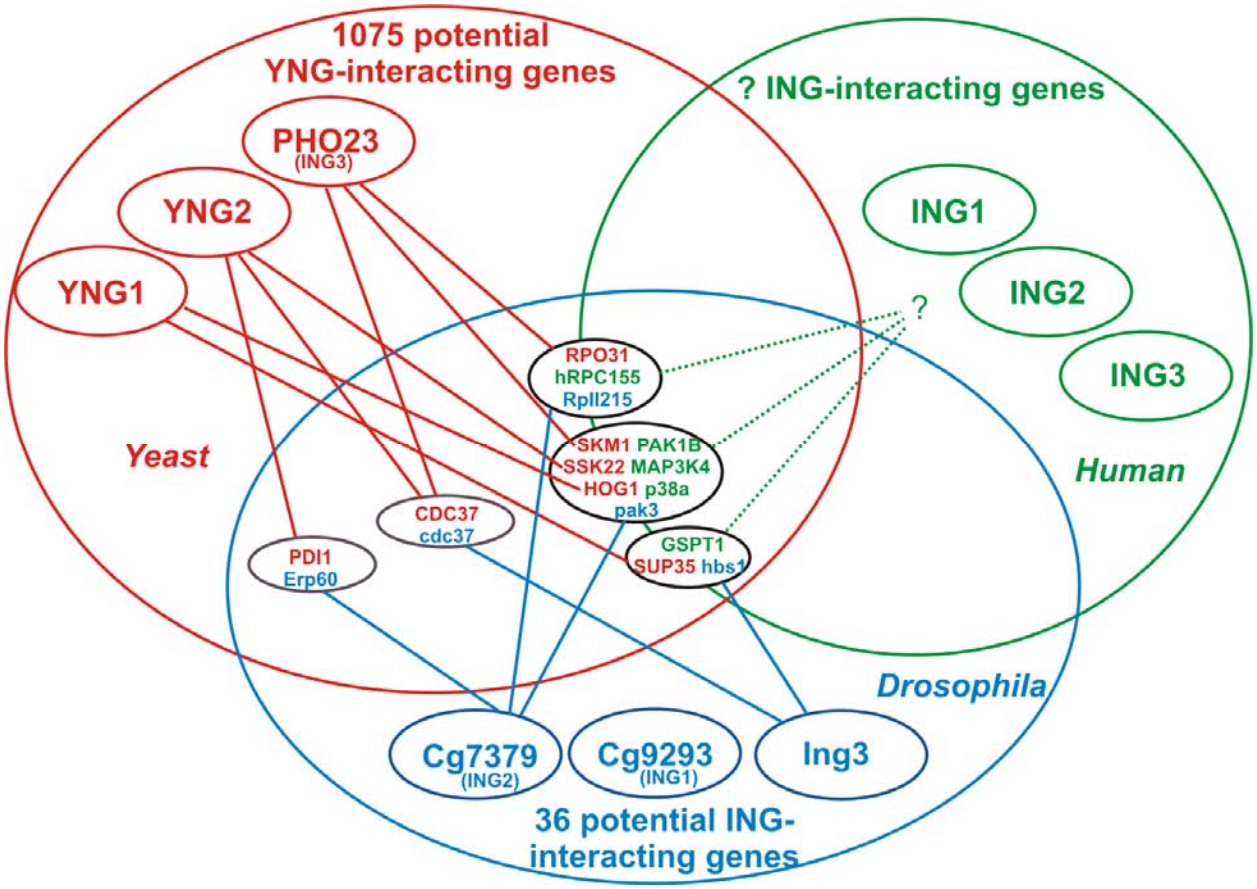

1075 potential
YNG-interacting genes
? ING-interacting genes
PHO23
(ING3)
YNG2
YNG1
ING1
ING2
ING3
?
RPO31
hRPC155
RpII215
SKM1 PAK1B
SSK22 MAP3K4
HOG1 p38a
pak3
CDC37
cdc37
GSPT1
SUP35 hbs1
PDI1
Erp60
Yeast
Human
Drosophila
Cg7379
(ING2)
Cg9293
(ING1)
Ing3
36 potential ING-
interacting genes

Figure 17: ING1 interacts with p38MAPK at the overexpression level

HEK 293 cells were transfected with the equal amounts of the indicated constructs and cell lysates were immunoprecipitated using anti-ING1 and then immunoblotted with anti-p38MAPK. The blot was reprobed with anti-ING1 to confirm equal IP efficiency. Anti-GST was used as a negative control. The lower panel represents input from cell lysates used in IPs and indicates similar expression levels of the indicated proteins.

17

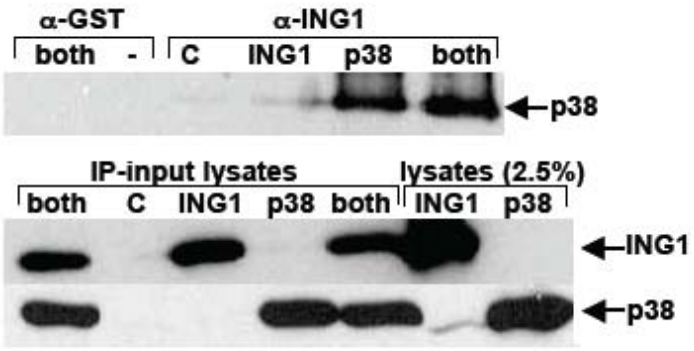

α-GST
α-ING1
both
-
C
ING1
p38
both
p38
IP-input lysates
lysates (2.5%)
both
C
ING1
p38
both
ING1
p38
ING1
p38

Figure 18: ING1 interacts with MAP3K4 (MEKK4) at the overexpression level

HEK 293 cells were transfected with the equal amounts of the indicated constructs and cell lysates were immunoprecipitated using anti-ING1 and then immunoblotted with anti-MEKK4. The blot was reprobed with anti-ING1 to confirm equal IP efficiency. Anti-GST was used as a negative control. The lower panel represents input from cell lysates used in IPs and indicates similar expression levels of the indicated proteins.

18

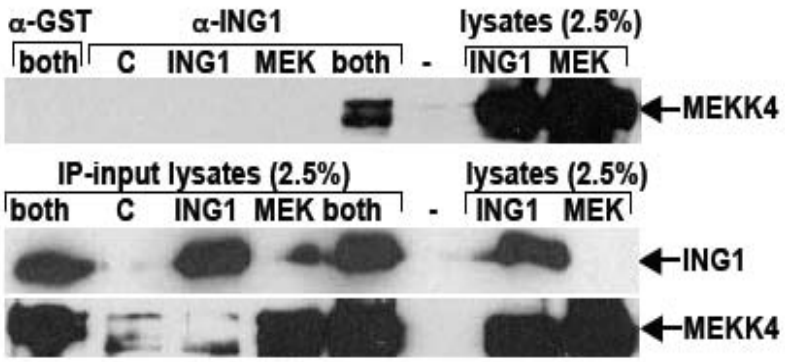

α-GST
α-ING1
lysates (2.5%)
both
C
ING1
MEK
both
-
ING1
MEK
MEKK4
IP-input lysates (2.5%)
lysates (2.5%)
both
C
ING1
MEK
both
-
ING1
MEK
ING1
MEKK4

Figure 19: ING1-p38MAPK and ING1-MEKK4 interaction at the endogenous levels

HEK293 cell lysates were immunoprecipitated using anti-ING and then immunoblotted with anti-p38MAPK. Interestingly, the interaction seems to be stronger under normal rather than stress conditions. The input shows equal amounts of cell lysates have been used. ING1 levels after reblotting of the same membrane with anti-ING1. ING1 levels were equal under both normal and stress conditions. The membrane was re-probed with anti-MEKK4 to detect ING1-MEKK4 interaction at the endogenous interaction.

19

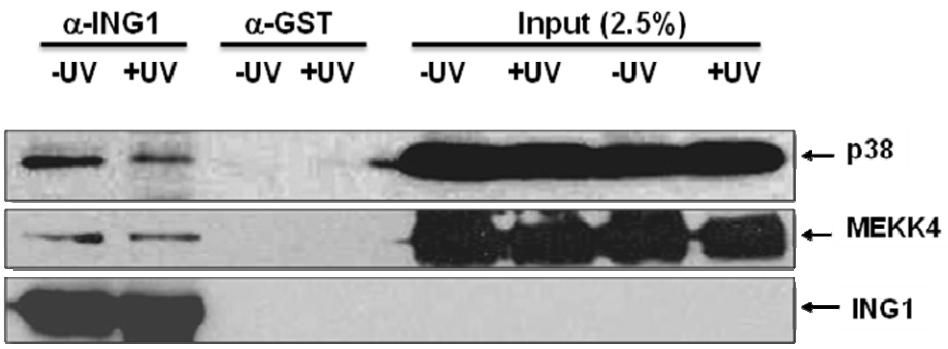


Figure 20: Merger of the interaction maps for the three human proteins ING1, MEKK4, and p38MAPK based on empirical data retrieved from the STRING database

Lines connecting p38MAPK and its interactors have been excluded for clarity, except where the interactor is directly shared with MEKK4 as part of a well-defined stress response pathway. MEKK4 and ING1 have relative few confirmed interactions, but the interactions predicted and confirmed in our study (shown as thick lines) are the first to directly tie ING1 to multiple key proteins in this stress response. Examining such interaction overlap graphs can help, in deciding which interaction predictions to validate biochemically, based on biological salience.

20

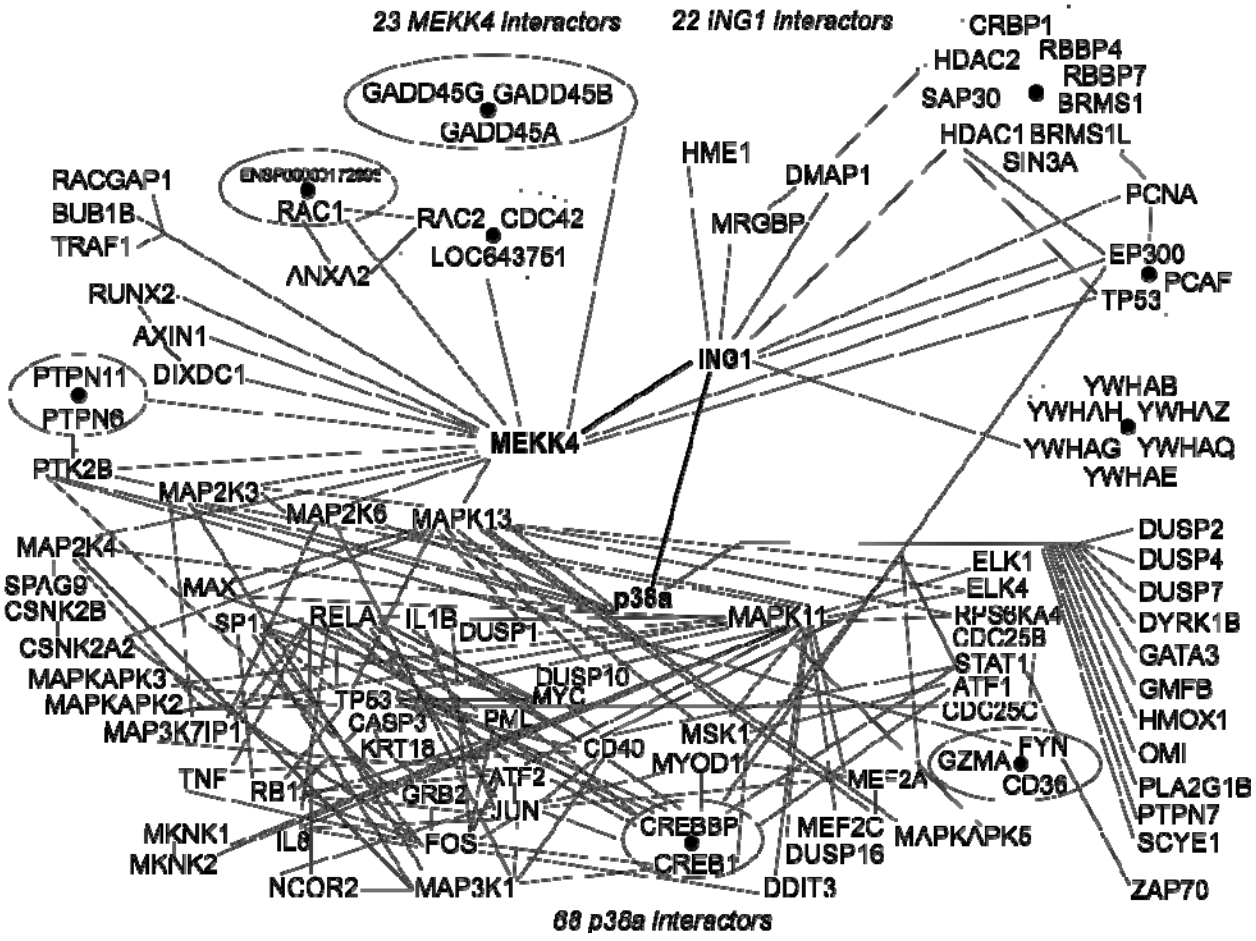

23 MEKK4 Interactors
22 ING1 Interactors
GADD45G
GADD45B
GADD45A
RACGAP1
BUB1B
TRAF1
RAC1
RAC2
CDC42
LOC643751
ANXA2
RUNX2
AXIN1
PTPN11
DIXDC1
PTPN6
MEKK4
PTK2B
MAP2K3
MAP2K6
MAPK13
MAP2K4
SPAG9
CSNK2B
CSNK2A2
MAPKAPK3
MAPKAPK2
MAP3K7IP1
MAX
SP1
RELA
IL1B
DUSP1
p38a
MAPK11
DUSP10
MYC
TP53
CASP3
KRT18
PML
CD40
MSK1
MYOD1
TNF
RB1
GRB2
ATF2
JUN
ME2A
MKNK1
MKNK2
IL6
FOS
CREBBP
CREB1
MEF2C
DUSP16
MAPKAPK5
NCOR2
MAP3K1
DDIT3
ING1
HME1
DMAP1
MRGBP
CRBP1
HDAC2
RBBP4
RBBP7
SAP30
BRMS1
HDAC1
BRMS1L
SIN3A
PCNA
EP300
PCAF
TP53
YWHAB
YWHAH
YWHAZ
YWHAG
YWHAQ
YWHAE
ELK1
ELK4
RPS6KA4
CDC25B
STAT1
ATF1
CDC25C
GZMA
FYN
CD36
DUSP2
DUSP4
DUSP7
DYRK1B
GATA3
GMFB
HMOX1
OMI
PLA2G1B
PTPN7
SCYE1
ZAP70
68 p38a Interactors

Table 5: list of the YNG interacting proteins which have conserved human counterparts

The list can be found at http://www.visualgenomics.ca/gordonp/S2.html. The probability of interaction scores were based on the Z-scores used in the Krogan *et al.* datasets. The Z-scores for protein identification by mass spectrometry were rescaled by the authors to a scale from 0 to 1 such that a score of 0.5 represents 70% confidence.

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Karyopherin (importin) alpha	SRP1	YNG2:0.981	Karyopherin alpha 5
Histone acetyltransferase (myst family)	ESA1	YNG2:0.968	HTATIP protein
Histone deacetylase complex SIN3 component	SIN3	PHO23:0.957, YNG2:0.024, YNG1:0.012	SIN3 homolog B transcription regulator
Karyopherin (importin) beta	KAP95	YNG2:0.369	Complex Of Ran With Importin Beta
Vacuolar H+-ATPase V1 sector subunit B	VMA2	YNG2:0.143	ATP6V1B2
Transcription factor involved in chromatin remodeling contains bromodomain	BDF1	YNG2:0.113, PHO23:0.012	RING3
Nop14-like family	NOP14	YNG2:0.105	Nucleolar protein 14
26S proteasome regulatory complex subunit RPN8/PSMD7	RPN8	YNG1:0.071	PSMD7
Uncharacterized conserved protein	DUS3	YNG2:0.071	Unnamed protein product
Histidyl-tRNA synthetase	HTS1	YNG1:0.024	Histidyl-tRNA synthetase 2
Histone H3 (Lys4) methyltransferase complex and rna cleavage factor II subunit SWD2	SWD2	YNG1:0.024	Histone H3 (Lys4) methyltransferase
RNA polymerase II transcription initiation/nucleotide excision repair factor tfiih subunit SSL1	SSL1	YNG2:0.024	Transcription factor IIH
26S proteasome regulatory complex subunit RPN6/PSMD11	RPN6	YNG1:0.024	26S proteasome subunit 9
rRNA processing protein Rrp5	RRP5	YNG2:0.024, PHO23:0.012, YNG1:0.012	RRP5 protein homolog (Programmed cell death protein 11)
DNA replication licensing factor MCM3 component	CDC47	YNG2:0.023	MCM7
WD40-repeat-containing subunit the 18S rRNA processing complex	DIP2	YNG2:0.023	WD repeat-containing protein 3
Golgi nucleoside diphosphatase	YND1	YNG2:0.023	KIAA0392
Beclin-like protein	VPS30	YNG2:0.023	Beclin 1
Histone acetyltransferase (myst family)	SAS3	YNG1:0.016	Tat interactive protein
Arp2/3 complex-interacting protein VIP1/Asp1 involved in regulation actin cytoskeleton	VIP1	YNG1:0.016, PHO23:0.012	KIAA0433
Phosphatidylinositol kinase and protein kinases the PI-3 family	TRA1	YNG2:0.016	Hypothetical protein
S-adenosylmethionine synthetase	SAM1	YNG1:0.015, YNG2:0.012	S-adenosylmethionine synthetase
S-adenosylmethionine synthetase	SAM2	YNG1:0.015, YNG2:0.014	Methionine adenosyltransferase I, alpha
Methionine aminopeptidase	MAP2	PHO23:0.015	unnamed protein

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Putative ubiquitin fusion degradation protein	UFD4	YNG1:0.015	Thyroid hormone receptor interactor 12
Ypt/Rab-specific GTPase-activating protein GYP1	GYP1	YNG2:0.014	TBC1 domain-containing protein
Phosphatidylserine decarboxylase	PSD1	YNG2:0.014	Phosphatidylserine decarboxylase
Histone deacetylase complex catalytic component RPD3	RPD3	YNG2:0.014, YNG1:0.014	Histone deacetylase 1
26S proteasome regulatory complex component	RPN5	YNG1:0.014, YNG2:0.012	Proteasome 26S non-ATPase subunit 12
Translation initiation factor 2B alpha subunit (eIF-2Balpha/GCN3)	GCN3	YNG1:0.014	Eukaryotic translation initiation factor 2B
Mitogen-activated protein kinase	SLT2	PHO23:0.014	Mitogen-activated protein kinase 7
Predicted ATPase	AFG1	YNG2:0.014	Lactation elevated 1
atcc	RRP12	YNG2:0.012, YNG1:0.014	Ribosomal RNA processing 12 homolog
Pleiotropic regulator 1	PRP46	YNG1:0.014	Pleiotropic regulator 1
Phosphoserine aminotransferase	SER1	YNG2:0.014	Phosphoserine aminotransferase
Vesicle coat complex copii subunit SEC13	SEC13	YNG1:0.014	SEC13 protein
26S proteasome regulatory complex subunit RPN3/PSMD3	RPN3	YNG1:0.014	unnamed protein
Serine/threonine specific protein phosphatase involved in cell cycle control PP2A-related	SIT4	YNG1:0.014	Serine/threonine protein phosphatase catalytic subunit
Mitogen-activated protein kinase	HOG1	YNG1:0.014	P38 MAPK alpha
Bifunctional leukotriene A4 hydrolase/aminopeptidase LTA4H	YNL045W	PHO23:0.014	Leukotriene A4 Hydrolase
RNA polymerase III large subunit	RPO31	PHO23:0.014	RNA polymerase III largest subunit
Translation initiation factor 2B beta subunit (eIF-2Bbeta/GCD7)	GCD7	YNG1:0.014	Eukaryotic translation initiation factor 2B
Histone deacetylase complex catalytic component HDA1	HDA1	YNG1:0.014	Hypothetical protein
Protein involved in turnover	MRT4	YNG1:0.014	Acidic ribosomal protein PO-like
Serine/threonine protein kinase	YPL150W	YNG2:0.012	MAP/microtubule affinity-regulating kinase 3
Myosin heavy chain	MYO1	YNG2:0.012	Myosin, heavy polypeptide 9
Actin-related protein Arp2/3 complex subunit ARPC1/p41-ARC	ARC40	YNG1:0.012	Actin related protein 2/3 complex subunit 1B
Fructose-2 6-bisphosphatase	PFK26	YNG2:0.012, YNG1:0.012	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
WD domain G-beta repeat	AMA1	YNG2:0.012	KIAA1242
export protein (contains WD40 repeats)	GLE2	YNG2:0.012	RAE1 (RNA export 1)
TPR-containing nuclear phosphoprotein that regulates K(+) uptake	CTR9	YNG1:0.012	KIAA0155
tpr Domain	CDC27	YNG1:0.012	Nuc2
5'-3' exonuclease	RAD27	YNG1:0.012	Flap structure-specific endonuclease 1
Predicted metal-dependent protease the PAD1/JAB1 superfamily	RPN11	PHO23:0.012, YNG1:0.012	26S proteasome-associated pad1 homologe
RNA polymerase I and III subunit	RPC40	YNG2:0.012	RNA polymerase 1C
Uncharacterized conserved protein	VAC14	YNG2:0.012, YNG1:0.012	Unnamed protein
Translation initiation factor 2 gamma subunit (eIF-2gamma GTPase)	GCD11	YNG1:0.012	Eukaryotic translation initiation factor 2
Chromatin remodeling complex wstf-iswi small subunit	STH1	YNG2:0.012	transcriptional activator hSNF2a
Serine/threonine protein kinase and endoribonuclease ERN1/IRE1 sensor the unfolded response pathway	IRE1	YNG2:0.012	Endoplasmic reticulum to nucleus signalling 2
RNA polymerase II second largest subunit	RPB2	YNG2:0.012	DNA-directed RNA polymerase II polypeptide B
Cullin a subunit E3 ubiquitin ligase	CDC53	YNG2:0.012	Cullin 2
Cysteinyl-tRNA synthetase	YNL247W	YNG1:0.012	Cysteinyl-tRNA synthetase isoform A
Serine/threonine protein kinase	KIN1	YNG1:0.012, YNG2:0.012	AMP-activated protein kinase
DNA polymerase zeta catalytic subunit	REV3	YNG2:0.012	REV3-like, catalytic subunit of DNA polymerase zeta
Mismatch repair MSH3	MSH3	YNG2:0.012	MutS homolog 3
Actin-related protein Arp2/3 complex subunit ARPC2	ARC35	YNG2:0.012	Actin related protein 2/3 complex, subunit 2
Subtilisin-like proprotein convertase	KEX2	YNG2:0.012	Proprotein convertase subtilisin/kexin type 7
Ribonucleotide reductase alpha subunit	RNR1	YNG1:0.012	Ribonucleotide reductase M1 subunit
NDR and related serine/threonine kinases	RIM15	YNG2:0.012	Serine/threonine kinase 38
RNA 3'-terminal phosphate cyclase	RCL1	YNG1:0.012	RNA terminal phosphate cyclase-like 1
20S proteasome regulatory subunit alpha type PSMA5/PUP2	PUP2	YNG2:0.012	Proteasome endopeptidase complex
Ctp synthase (UTP-ammonia lyase)	URA8	YNG2:0.012	Unnamed protein
Splicing factor ATP-dependent rna helicase	PRP43	YNG2:0.012	DEAH (Asp-Glu-Ala-His) box polypeptide 15
Serine/threonine-protein kinase involved in autophagy	GCN2	YNG2:0.012	EIF2AK4

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Cell division control protein negative regulator transcription	CDC39	YNG1:0.012, YNG2:0.012	KIAA1007
Glucosidase II catalytic (alpha) subunit	ROT2	YNG1:0.012	FLJ00088
Ubiquitin activating enzyme UBA1	UBA1	PHO23:0.012, YNG1:0.012	Ubiquitin activating enzyme E1
GTPase Rab1/YPT1 small G protein superfamily and related GTP-binding proteins	YPT1	YNG2:0.012	Rab1B
Superfamily II dna and rna helicases	SUB2	YNG2:0.012	Unnamed protein
Kinase A-anchor protein Neurobeachin and related beach WD40 repeat proteins	BPH1	YNG2:0.012	Unnamed protein
Replication factor C subunit RFC5	RFC3	YNG2:0.012	RFC5
5'-3' exonuclease	RAT1	YNG2:0.012	DHP protein
26S proteasome regulatory complex ATPase RPT3	RPT2	YNG2:0.012	26S protease (S4) regulatory subunit
Serine/threonine protein phosphatase 2A regulatory subunit	RTS1	YNG1:0.012	PP2A
Helicase conserved C-terminal domain	RAD5	YNG2:0.012	ATPase
26S proteasome regulatory complex ATPase RPT6	RPT3	YNG2:0.012	ATPase homolog
Predicted RNA-binding protein	YPL009C	YNG1:0.012	SDCCAG1 protein
Protein involved in the nuclear export pre-ribosomes	NOC3	YNG1:0.012	Nucleolar complex associated 3 homolog
Structural maintenance chromosome protein 3 (sister chromatid cohesion complex Cohesin subunit SMC3)	SMC3	YNG2:0.012	Structural maintenance of chromosomes 3
DNA/RNA helicase MER3/SLH1 DEAD-box superfamily	SLH1	YNG2:0.012	Unnamed protein
3-oxoacyl CoA thiolase	ERG10	YNG2:0.012	Acetyl-Coenzyme A acetyltransferase 1 precursor
Uncharacterized conserved protein	SHQ1	YNG2:0.012	SHQ1 homolog
2-polyprenylphenol hydroxylase	MCR1	YNG2:0.012	Cytochrome b5 reductase
Predicted membrane protein	YER140W	YNG2:0.012	TAPT1 protein
3-phosphoinositide-dependent protein kinase (PDK1)	PKH2	YNG2:0.012	PKB-like
WD40 repeat stress protein/actin interacting protein	AIP1	YNG2:0.012	Hypothetical protein
Ras-related GTPase	RAS1	YNG2:0.012	Ras family small GTP binding protein TC21
Oxysterol-binding protein	OSH3	YNG2:0.012	Oxysterol-binding protein-like protein 6
Nuclear pore complex subunit	NUP157	YNG2:0.012	Nucleoporin
Mannosyl oligosaccharide glucosidase	CWH41	YNG2:0.012	Glucosidase I
Uncharacterized conserved coiled-coil protein	MLP2	YNG2:0.012	CENP-E protein
Actin-related protein	ARP8	YNG1:0.012	Unnamed protein
Required LacZ RNA expression certain suppressor the Transcriptional T defect Hpr1 H by Overexpression O plays a role in transcription elongation polymerase II Involved Rlr1p RLR1 orf9 protein THO2	RLR1	YNG2:0.012, YNG1:0.012	THO complex 2
SEC7 domain proteins	SEC7	YNG1:0.012	Brefeldin A-inhibited

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
			guanine nucleotide-exchange protein 1
Guanine nucleotide binding protein MIP1	KOG1	YNG2:0.012	Raptor protein
P-type ATPase	NEO1	YNG2:0.012	KIAA0611
Ubiquitin carboxyl-terminal hydrolase	UBP15	YNG2:0.012	Ubp-Family Deubiquitinating Enzyme
DNA polymerase elongation subunit (family B)	POL2	YNG2:0.012	DNA polymerase epsilon catalytic subunit
atcc	NIP1	YNG1:0.012	eIF-3 p110 subunit
Chromosome condensation complex Condensin	YCS4	YNG2:0.012	unnamed protein
Superfamily II RNA helicase	SKI2	YNG1:0.012	SKI2W
DNA topoisomerase III alpha	TOP3	YNG2:0.012, YNG1:0.012	DNA topoisomerase III
Methionyl-tRNA synthetase	MES1	YNG2:0.012	MARS
Putative methionine aminopeptidase	MAP1	YNG1:0.012	Methionyl aminopeptidase 1
Transcription factor NF-X1 contains NFX-type Zn2+-binding and R3H domains	FAP1	YNG2:0.012, YNG1:0.012	Nuclear transcription factor, X-box binding 1
AAA+-type ATPase	SEC18	YNG2:0.012	N-ethylmaleimide-sensitive factor
Reductases with broad range substrate specificities	SPS19	YNG2:0.012	2,4-dienoyl CoA reductase 2
DNA damage-responsive repressor GIS1/RPH1 jumonji superfamily	RPH1	YNG2:0.012	Jumonji domain containing 2C
Golgi nucleoside diphosphatase	GDA1	YNG2:0.012	Ectonucleoside triphosphate diphosphohydrolase 6
Uncharacterized protein involved in cell differentiation/sexual development	CAF40	YNG1:0.012	required for cell differentiation1 (RCD1) homolog
Meiotic cell division protein Pelota/DOM34	DOM34	YNG2:0.012	Pelota homolog
RNA helicase	HCA4	YNG2:0.012, YNG1:0.012	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10
Uncharacterized conserved protein contains DEP domain	IML1	YNG2:0.012	DEP domain containing 5
Alpha subunit (small G protein superfamily)	GPA1	YNG2:0.012	Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1
ABC-type transport system involved in Fe-S cluster assembly permease and ATPase components	ATM1	YNG2:0.012	ABC transporter
DnaJ-class molecular chaperone with C-terminal Zn finger domain	SCJ1	YNG2:0.012	DnaJ subfamily A member 2
Ubiquitin-specific protease	UBP6	YNG2:0.012	ubiquitin specific protease 14
Chromatin remodeling complex SWI/SNF component	ISW1	YNG2:0.012,	SWI/SNF-related matrix-

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
SWI2 and related ATPases (DNA/RNA helicase superfamily)		YNG1:0.012	associated actin-dependent regulator of chromatin a5
WD40-repeat-containing subunit the 18S rRNA processing complex	PWP2	YNG2:0.012	periodic tryptophan protein 2
SMT3/SUMO-activating complex AOS1/RAD31 component	AOS1	YNG1:0.012	SUMO-1-activating enzyme E1 N subunit
Deadenylase subunit	POP2	YNG2:0.012	CAF2
RNA-binding protein required 60S ribosomal subunit biogenesis	SSF1	YNG1:0.012	SSF1/P2Y11 chimeric protein
Cdc4 and related F-box WD-40 proteins	CDC4	YNG2:0.012	Unnamed protein
Serine/threonine protein kinase	PKH1	YNG2:0.012	PKB-like
MEKK and related serine/threonine protein kinases	SSK22	YNG2:0.012	PRO0412
Transaldolase	TAL1	YNG2:0.012	TALDO1
60S ribosomal protein	RPL1A	YNG1:0.012	Ribosomal protein L10a
DNA primase small subunit	PRI1	YNG2:0.012	DNA primase polypeptide 1
Uncharacterized protein contains Trp-Asp (WD) repeat	YGR054W	YNG2:0.012	ukaryotic translation initiation factor 2A
Serine/threonine protein kinase	SNF1	YNG1:0.012	AMP-activated protein kinase
SNF2 family DNA-dependent ATPase domain-containing protein	MOT1	YNG2:0.012	BTAF1 RNA polymerase II, B-TFIID transcription factor-associated
Thymidylate synthase	CDC21	YNG2:0.012	Thymidylate synthase
40S ribosomal protein S24	RPS24B	YNG2:0.012	Hypothetical protein
MA3 domain	CWC22	YNG2:0.012	KIAA1604
DNA repair protein RAD50 ABC-type ATPase/SMC superfamily	RAD50	YNG2:0.012	RAD50
Predicted GTP-binding protein MMR1	LSG1	YNG2:0.012, YNG1:0.012	Unnamed protein
SRP40 C-terminal domain	SRP40	YNG1:0.012	Dentin phosphoryn
60S ribosomal protein L9	RPL9B	YNG1:0.012	Ribosomal protein L9
DNA polymerase III epsilon subunit and related 3'-5' exonucleases	PAN2	YNG2:0.012	PABP1-dependent poly(A) specific ribonuclease PAN2
Nucleic-acid-binding protein possibly involved in ribosomal biogenesis	MAK21	YNG1:0.012	CCAAT/enhancer binding protein zeta
Protein containing adaptin N-terminal region	GCN1	YNG2:0.012, YNG1:0.012	GCN1
DNA replication licensing factor MCM3 component	MCM2	YNG2:0.012	MCM2
RNA polymerase II transcription initiation/nucleotide excision repair factor tfiih subunit TFB2	TFB2	YNG2:0.012	Transcription factor IIH, polypeptide 4
Chromatin remodeling factor subunit and related transcription factors	RSC8	YNG2:0.012	SWI/SNF related, matrix associated, actin dependent regulator of

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
			chromatin
tRNA-splicing endonuclease positive effector (SEN1)	NAM7	PHO23:0.012, YNG1:0.012	Regulator of nonsense transcripts 1
TatD related DNase	YBL055C	YNG2:0.012	TatD DNase
Cell cycle control protein	CDC50	YNG2:0.012	Transmembrane protein 30A
Ribosome biogenesis protein	NOP58	YNG2:0.012	NOP5/NOP58
Vacuolar sorting protein PEP3/VPS18	PEP3	YNG2:0.012	Vacuolar protein sorting 18
Lysyl-tRNA synthetase (class II)	KRS1	YNG2:0.012	Lysyl-tRNA synthetase
GTPase-activating protein	MSB4	YNG2:0.012	GRTP1
Predicted metal-dependent hydrolases related alanyl-tRNA synthetase HxxxH domain	YNL040W	YNG2:0.012	Alanyl-tRNA synthetase domain containing 1
Dopey and related predicted leucine zipper transcription factors	DOP1	YNG2:0.012	Dopey family member 1
Vesicle coat complex copi alpha subunit	COP1	YNG2:0.012, YNG1:0.012	Coatomer protein complex
Peptide chain release factor 1 (eRF1)	SUP45	YNG1:0.012, YNG2:0.012	ETF1
Superfamily II DNA/RNA helicases SNF2 family	FUN30	YNG1:0.012	KIAA1122
Serine/threonine protein kinase	PKC1	YNG2:0.012	Protein kinase N2
Myosin heavy chain	MYO4	YNG1:0.012	KIAA1119
Molecular chaperones GRP170/SIL1 HSP70 superfamily	LHS1	YNG2:0.012	Hypoxia up-regulated 1 precursor
Superfamily II DNA and RNA helicases	DBP1	YNG2:0.012	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3
Isopeptidase T	UBP14	YNG2:0.012	USP5
ATP-dependent RNA helicase	DBP10	YNG2:0.012	DDX54
RNA polymerase I second largest subunit	RPA135	YNG2:0.012	RNA polymerase I polypeptide B
G-protein beta subunit	STE4	YNG2:0.012	Guanine binding protein
Clathrin coat binding protein/Huntingtin interacting protein HIP1 involved in regulation endocytosis	SET2	YNG1:0.012	unnamed protein
NAD-dependent aldehyde dehydrogenases	PUT2	YNG2:0.012	Aldehyde dehydrogenase 4
DNA replication licensing factor MCM4 component	MCM3	YNG2:0.012	Cervical cancer proto-oncogene 5
HSP90 co-chaperone CPR7/Cyclophilin	CPR6	YNG2:0.012	PPID
Microtubule binding protein YTM1 (contains WD40 repeats)	YTM1	YNG1:0.012	WD repeat domain 12
GTPase Rab5/YPT51 and related small G protein superfamily GTPases	YPT53	YNG2:0.012	RAB5C, member RAS oncogene family
Isoleucyl-tRNA synthetase	ILS1	YNG2:0.012	IARS
imp dehydrogenase/GMP reductase	IMD4	PHO23:0.012	IMP dehydrogenase type 1

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Polypeptide release factor 3	SUP35	YNG1:0.012	G1/S phase transition 1
F-actin capping protein beta subunit	CAP2	YNG1:0.012	capping protein (actin filament)
Superfamily II DNA/RNA helicases SNF2 family	ISW2	YNG2:0.012	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin
Cysteine synthase	CYS4	YNG2:0.012	CBS
Uncharacterized conserved protein	NOC2	YNG1:0.012	Nucleolar complex associated 2
Vigilin	SCP160	YNG2:0.012	High density lipoprotein binding protein
PIK-related protein kinase and rapamycin target	TOR2	YNG2:0.012	FK506 binding protein 12-rapamycin associated protein 1
Transcriptional accessory protein	SPT6	YNG2:0.012	SUPT6H
ERCC4 domain	RAD1	YNG2:0.012	DNA repair endonuclease XPF
Threonine/serine dehydratases	CHA1	YNG2:0.012	Serine dehydratase
Vesicle coat complex AP-2 alpha subunit	APL3	YNG2:0.012	Adaptor-related protein complex 2
Helicase the dead superfamily	CHL1	YNG2:0.012	DDX11
Uncharacterized conserved protein	CDC123	YNG2:0.012	Chromosome 10 open reading frame 7
Conserved protein Mo25	HYM1	YNG2:0.012	Calcium binding protein 39-like isoform 2
Metalloendopeptidase family - mitochondrial intermediate peptidase	OCT1	YNG2:0.012	Mitochondrial intermediate peptidase
Nuclear exosomal rna helicase MTR4 DEAD-box superfamily	MTR4	YNG1:0.012	SKIV2L2
Mercaptopyruvate sulfurtransferase/thiosulfate sulfurtransferase	YOR251C	YNG2:0.012	Thiosulfate sulfurtransferase
SEC7 domain proteins	GEA1	YNG2:0.012	KIAA0248
Predicted sugar kinase	POS5	YNG2:0.012	FLJ13052
arm repeat	ECM29	YNG2:0.012	KIAA0368
DEAH-box RNA helicase	ECM16	YNG2:0.012, YNG1:0.012	DEAH (Asp-Glu-Ala-His) box polypeptide 37
Phosphatidylinositol kinase and protein kinases the PI-3 family	TOR1	YNG2:0.012	FK506 binding protein 12-rapamycin associated protein 1
Chromatin remodeling factor subunit and related transcription factors	SWI3	YNG2:0.012	SWI/SNF complex
Valyl-tRNA synthetase	VAS1	YNG2:0.012	Valyl-tRNA synthetase
Polyadenylate-binding protein (rrm superfamily)	PES4	YNG2:0.012	Poly A binding protein
Leucyl-tRNA synthetase	CDC60	YNG2:0.012	KIAA1352
Ribosomal protein S26	RPS26B	YNG1:0.012	Ribosomal protein S26

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Ubiquitin-specific protease	UBP13	YNG1:0.012	Ubiquitin-specific protease 12
Protein required biogenesis the 60S ribosomal subunit	NOP7	YNG1:0.012	Pescadillo homolog 1, containing BRCT domain
Glycosyltransferase	ALG5	YNG2:0.012	Asparagine-linked glycosylation 5
Predicted GTP-binding protein (odn superfamily)	YBR025C	YNG2:0.012, YNG1:0.012	GTP-binding protein
26S proteasome regulatory complex ATPase RPT6	RPT5	YNG2:0.012	Tat binding protein 1 (TBP-1)
Puf family RNA-binding protein	PUF6	YNG1:0.012, YNG2:0.012	RP11-526D20.2
Ubiquitin-protein ligase	CDC34	YNG2:0.012, PHO23:0.012, YNG1:0.012	Unnamed protein
RNA polymerase III second largest subunit	RET1	YNG2:0.012	RNA polymerase III
Uncharacterized conserved protein	TSR1	YNG1:0.012	KIAA1401
Ribosome-associated chaperone zuotin	ZUO1	YNG2:0.012	M-phase phosphoprotein 11
Ribosomal protein S6 kinase and related proteins	CBK1	YNG2:0.012	Serine/threonine kinase 38
Nucleolar GTPase	NOG2	YNG2:0.012	Guanine nucleotide binding protein-like 2
Proteins containing the fad binding domain	DLD3	YNG2:0.012	D-2-hydroxyglutarate dehydrogenase
ATP-dependent DNA ligase I	CDC9	YNG1:0.012	DNA Ligase I
Carbamoylphosphate synthase large subunit	URA2	YNG2:0.012	Multifunctional protein CAD
60s ribosomal protein L15	RPL15B	YNG2:0.012	60S ribosomal protein L15
ER-Golgi vesicle-tethering protein p115	USO1	YNG2:0.012, YNG1:0.012	Smooth muscle myosin heavy chain 11
Superfamily II helicase	BRR2	YNG2:0.012	Unnamed protein
Ubiquitin-protein ligase	UBC11	YNG2:0.012	Ubiquitin-conjugating enzyme E2C
Mitochondrial carrier protein	PET8	YNG2:0.012	Solute carrier family 25
Translation initiation factor 3 subunit a (eIF-3a)	RPG1	YNG2:0.012	EIF3A
Cytoplasmic Zn-finger protein BRAP2 (BRCA1 associated protein)	YHL010C	YNG2:0.012	Impedes mitogenic signal propagation mRNA
Oxidoreductin endoplasmic reticulum membrane-associated protein involved in disulfide bond formation	ERO1	YNG2:0.012	ERO1-like protein
Acetyl-CoA carboxylase	ACC1	YNG2:0.012, YNG1:0.012	Acetyl-CoA carboxylase 1
Multidrug resistance-associated protein/mitoxantrone resistance protein abc superfamily	YOR1	YNG2:0.012	Multidrug resistance-associated protein

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Transcription elongation factor	SPT5	YNG2:0.012, YNG1:0.012	DSIF p160
Phosphatidylinositol-4-phosphate 5-kinase	FAB1	YNG2:0.012	phosphatidylinositol-3-phosphate/phosphatidylinositol 5-kinase, type III
ATP-dependent rna helicase pitchoune	MSS116	YNG2:0.012	Unnamed protein
Uncharacterized conserved protein	DCS2	YNG2:0.012	Unnamed protein
Predicted methyltransferase	ABP140	YNG1:0.012	Methyltransferase-like protein 2
AAA+-type ATPase containing the peptidase M41 domain	AFG3	YNG2:0.012	AFG3 ATPase family gene 3-like 2
Histones H3 and H4	HHT2	YNG2:0.012, PHO23:0.012, YNG1:0.012	H3 histone, family 3A
Septin family protein (P-loop GTPase)	SHS1	PHO23:0.012	SEPT7
Endonuclease III	NTG1	YNG2:0.012	NTHL1
Phosphatidylinositol-4-phosphate 5-kinase	MSS4	YNG2:0.012	KIAA0589
U5 snRNP spliceosome subunit	PRP8	YNG2:0.012, YNG1:0.012	PRP8 pre-mRNA processing factor 8 homolog
S-adenosylhomocysteine hydrolase	SAH1	YNG2:0.012, YNG1:0.012	S-Adenosylhomocysteine Hydrolase
Nonsense-mediated decay 2 protein	NMD2	YNG2:0.012	KIAA1408
Glutamyl- and glutaminyl-tRNA synthetases	GLN4	YNG2:0.012	QARS
Ded_cyto Dedicator cytokinesis. family represents a conserved region approximately 200 residues long within number eukaryotic cytokinesis proteins. These are potential guanine nucleotide exchange factors which activate some small GTPases by exchanging bound gdp free gtp	YLR422W	YNG2:0.012	KIAA0299
Aicar transformylase/IMP cyclohydrolase/methylglyoxal synthase	ADE17	YNG1:0.012	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
Ribosome biogenesis protein RPF1 contains IMP4 domain	RPF1	YNG2:0.012	Hypothetical protein
Serine/threonine protein kinase	GIN4	YNG2:0.012	Putative serine/threonine protein kinase variant B3
Sister chromatid cohesion complex Cohesin subunit PDS5	PDS5	YNG2:0.012	PDS5, regulator of cohesion maintenance
Pyrophosphate-dependent phosphofructo-1-kinase	PFK2	YNG2:0.012	Phosphohexokinase
Putative cargo transport protein ERV29	ERV29	YNG1:0.012	Surfeit 4
Checkpoint kinase and related serine/threonine protein kinases	CHK1	YNG2:0.012	Checkpoint Kinase Chk1
tRNA cytosine-5-methylases and related enzymes the NOL1/NOP2/sun superfamily	NCL1	YNG2:0.012, YNG1:0.012	Unnamed protein

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Vacuolar H+-ATPase V0 sector subunit a	VPH1	YNG2:0.012	ATP6V0A1
Protein kinase-like (PK-like)	SCY1	YNG2:0.012	SCY1-like 2 protein
NMD protein affecting ribosome stability and decay	NMD3	YNG2:0.012	CGI-07 protein
Elongation factor 2	EFT2	YNG1:0.012	Eukaryotic translation elongation factor 2
Mitochondrial carnitine-acylcarnitine carrier protein	YMC2	YNG2:0.012	Unnamed protein
Formyltetrahydrofolate synthetase	MIS1	YNG1:0.012	Formyltetrahydrofolate synthetase
Predicted translation initiation factor related eIF-2B alpha/beta/delta subunits (CIG2/IDI2)	YPR118W	YNG2:0.012	Hypothetical protein MGC3207
N-terminal acetyltransferase	NAT1	YNG2:0.012	Transcriptional coactivator tubedown-100
Adenosine deaminase	YJL070C	PHO23:0.012	Adenosine monophosphate deaminase 2
UDP-glucose pyrophosphorylase	UGP1	YNG2:0.012, YNG1:0.012	UDP-glucose pyrophosphorylase 2
Thymidylate kinase/adenylate kinase	CDC8	YNG2:0.012	Thymidylate Kinase
HrpA-like helicases	PRP22	YNG2:0.012	DHX8
Alanyl-tRNA synthetase	ALA1	PHO23:0.012	Alanyl-tRNA synthetase
Casein kinase II beta subunit	CKB2	YNG2:0.012	Casein kinase II
Myosin-like protein	MLP1	YNG2:0.012	Translocated promoter region (to activated MET oncogene)
RNA polymerase I large subunit	RPA190	YNG2:0.012	POLR1A
Endonuclease III	NTG2	YNG2:0.012	NTHL1
20S proteasome regulatory subunit beta type PSMB2/PRE1	PRE1	YNG2:0.012	Proteasome beta 2 subunit
SDA1	SDA1	YNG1:0.012	SDAD1
Kinesin-like protein	CIN8	YNG2:0.012, YNG1:0.012	Kinesin family member 11
Translation elongation factors (GTPases)	MEF1	YNG1:0.012	G elongation factor
Adenylosuccinate synthase	ADE12	YNG2:0.012	Adenylosuccinate synthetase
RNA pseudouridylate synthases	RIB2	YNG2:0.012	RPUSD2
Structural maintenance chromosome protein 1 (sister chromatid cohesion complex Cohesin subunit SMC1)	SMC1	YNG2:0.012	SMC1
Translation initiation factor 2 alpha subunit (eIF-2alpha)	SUI2	YNG1:0.012	EIF2
Iron binding protein involved in Fe-S cluster formation	ISU2	YNG2:0.012	Iron-sulfur cluster assembly enzyme
Ribosome biogenesis protein	SIK1	YNG2:0.012	hNop56
26S proteasome regulatory complex subunit RPN10/PSMD4	RPN10	YNG2:0.012	Proteasome 26S non-ATPase subunit 4
ATP-dependent RNA helicase A	YLR419W	YNG2:0.012	Hypothetical protein
Threonyl-tRNA synthetase	THS1	YNG2:0.012	TARS

Description	Gene Name (Yeast)	YNG family binding affinity	Human Genbank
Predicted unusual protein kinase	YPL109C	YNG2:0.012	Putative ubiquinone biosynthesis protein AarF
Single-stranded DNA-binding replication protein A (rpa) large (70 kDa) subunit and related ssDNA-binding proteins	RFA1	YNG2:0.012, PHO23:0.012	RFA1
IkappaB kinase complex ikap component	IKI3	YNG2:0.012	Inhibitor of kappa light polypeptide gene enhancer in B-cells
RNB-like protein	SSD1	YNG2:0.012	Hypothetical protein
Translation initiation factor 2B epsilon subunit (eIF-2Bepsilon/GCD6)	GCD6	YNG1:0.012	eIF-2B epsilon
UDP-N-acetylglucosamine pyrophosphorylase	QRI1	YNG2:0.012	UDP-N-acteylglucosamine pyrophosphorylase 1
Importin beta-related nuclear transport receptor	CRM1	YNG2:0.012, YNG1:0.012	Hypothetical protein
Ferredoxin/adrenodoxin reductase	ARH1	YNG2:0.012	Adrenodoxin reductase
5'-3' exonuclease	DIN7	YNG2:0.012	Exonuclease 1a
Translation initiation factor 3 subunit b (eIF-3b)	PRT1	YNG2:0.012, YNG1:0.012	Eukaryotic translation initiation factor 3
H/ACA small nucleolar rnp component GAR1	GAR1	YNG1:0.012	Nucleolar protein family A
Mitochondrial processing peptidase alpha subunit	MAS2	YNG2:0.012	PMPCA
Predicted alpha/beta hydrolase BEM46	YNL320W	YNG2:0.012	Hydrolase domain containing 13
DNA replication licensing factor MCM4 component	CDC54	YNG2:0.012	MCM4
Pre-mRNA cleavage and polyadenylation specificity factor	CFT1	YNG2:0.012	cleavage and polyadenylation specific factor 1
Glycogen debranching enzyme	GDB1	YNG2:0.012	Amylo-1,6-glucosidase
Glycogen synthase	GSY1	YNG2:0.012	Glycogen synthase 2
Serine/threonine protein kinase	KIN4	YNG2:0.012	AMP-activated protein kinase alpha-1
Phospholipase D1	SPO14	YNG2:0.012	Phospholipase D2
Putative SAM-dependent rRNA methyltransferase SPB1	SPB1	YNG2:0.012	Hypothetical protein SB92
G-protein beta subunit-like protein (contains WD40 repeats)	LST8	YNG2:0.012	Unnamed protein
tRNA-splicing endonuclease positive effector (SEN1)	SEN1	YNG2:0.012	Ataxia/oculomotor apraxia protein 2
dna topoisomerase I	TOP1	YNG2:0.012	DNA Topoisomerase I
Oxoprolinase	YKL215C	YNG2:0.012	5-oxoprolinase

Table 6: Comparison of datasets, generated by highthrough-put large scale experiments, for ING protein interactions

	Reference	Total number of Interactions	ING1/YNG1 interactions	ING2/YNG2 interactions	ING3/PHO23 interactions	Interactions matching our predictions
Yeast	Krogan [164] *	393,878	237	772	349	5
	Krogan [164]	14,317	5	24	15	0
	Jansen [180]	49,640	0	0	0	0
	Ho [137]	8,118	0	0	0	0
	Gavin [138]	589~	0	0	3	0
	Ito [141]	4,549	2	1	2	0
	Uetz [135]	957	0	0	0	0
	Gavin [191]	491~	2	2	8	0
	von Mering [165] ^	N/A	11	19	17	0
Human	Rual [192]	6,726	0	0	0	0
	Stelzl [193]	5,749	0	0	0	0
	Rhodes [194]	39,816	11	0	0	0
	Lehner [195]	71,806	3	0	0	0
	von Mering [165] ^	N/A	22	3	18	0

***** raw dataset

~proteins purified

^medium confidence experimental reports

Table 7: Comparison of ING-interactome results for existing protein interaction prediction tools

Prediction Tool	Total number of Interactions	ING1/2/3 interactions	Interactions not in existing STRING datasets	Overlap with predictions from this work
I2D [186]	200,599	33	2	0
YeastNet [196] #	102,803	100	57	0
OPHID [197]	47,221	10	2	0
POINT [198]	45,378	49	26	1^
Ulysses [199]	26,797	22	15	1*
Predict [200]	20,088	5	5	1*

Yeast database

^ The concurring interactor, PDI1, was excluded from the 5 core predictions of this work due to weakness of the human homolog

***** The concurring interactor, CDC37, was excluded from the 5 core predictions of this work due to weakness of the human homolog

CHAPTER FOUR: DISCUSSION

ING1 isoforms differentially affect cell growth

Alternative splicing allows genes to encode multiple functionally distinct proteins to increase the coding capacity of DNA. It can generate different isoforms, via deletion or addition of certain domains. Using computational approaches, it is estimated that around 40-60% of all human genes have at least two splice variants. The biological importance of alternative splicing is illustrated by the fact that around 15% of human genetic diseases result from mutations in consensus splice sites. Splicing can be specific to certain stresses, tissues, developmental, and physiological states, such as aging. Aging and aging-related diseases are associated with modification and/or impairment in alternative splicing through changing the levels of splicing machinery factors. For example, the mRNA levels of the 65-kDa subunit of one of the splicing machinery component, U2 auxiliary factor (U2AF) has been found to increase over three-fold in the gastrocnemius muscle of aged mice [201] . Age-related diseases have also been linked to alterations in splicing machineries. Alzheimer's disease (AD), one of the most common neurodegenerative disorders of aging, is a famous example of misdirected splicing regulation of genes. Presenilins are components of the gamma-secretase complex, the mutation of which has been implicated in the accumulation of amyloid plaques in the brain of Alzheimer's patients. It has been shown that mutations within the presenilin 1 (PS1) fourth intron impaired PS1 splicing and caused early onset of Alzheimer's, while an exon 5 deficient splice variant of presenilin 2 pre-mRNA was found to be abundant in the brain of AD patients [202]. Advances in the molecular biology of aging will allow for screening of alternative splicing variants and changes in the levels pre-mRNA splicing factors associated with increased cell population doublings.

The INhibitor of Growth (ING) proteins represent a family of type II tumour suppressors. They are encoded by five genes, most of which encode multiple splice variants [3]. ING1, the first member to be identified, was discovered through PCR-mediated subtractive hybridization of normal epithelial mammary cells against seven breast cancer cell lines followed by screening of a senescent cell cDNA library and an *in vivo* functional screen. ING1 was found to be preferentially expressed in normal rather than cancer-derived epithelial cells [6]. The ING1 locus encodes differentially spliced mRNAs with distinct initiation sites and three exons. Upon translation, these mRNAs will give the different ING1

isoforms p47ING1a (422 aa), p33ING1b (279 aa), p24ING1c (210 aa) and the hypothetical protein p27ING1d (235 aa) [203], all of which share exon2 which contains the PHD domain. The two predominant ING1 isoforms are ING1a, which is a product of exon 2 and exon 1b, and ING1b which is a product of exon2 and exon 1a. ING1b appears to be the highest expressed form among the different ING1 isoforms in cultured fibroblasts and epithelial cells, thus, the majority of previous studies have focused on the function of ING1b. Both ING1a and ING1b appear to have opposing functions and different cellular partners. For example, ING1b binds to PCNA through the PCNA-interaction Protein (PIP) motif in response to UV and induces apoptosis while ING1a does not show any interaction with PCNA under stress conditions, as this isoform lacks the PIP motif [15] **.** In addition, microinjection of ING1b construct induces acetylation of both H3 and H4 histones while microinjection of ING1a inhibits the acetylation of both histones [18].

The major aim of my project was to determine the functional differences between ING1a and ING1b, and whether these two isoforms show altered expression during cell aging. In this study, we present evidence that the effects INGl has on growth arrest, apoptosis and senescence are mediated, in part, by the products of differential splicing. We have shown that the relative expression levels of the two major ING1 isoforms, ING1a and ING1b, change dramatically as cells enter replicative senescence, with the INGla protein being expressed at higher levels and INGlb protein at lower levels. This differential expression results in the gradual accumulation of an alteration greater than 30-fold in the INGla:INGlb mRNA ratio with increased cell passage. However, the difference in protein levels did not mirror exactly the changes in the mRNA expression levels. This agrees with the fact that for many genes the measurement of mRNA response is not predictive of the protein response [204].

An increase in ING1a levels with senescence suggests that this isoform may contribute to cell aging. Consistent with this idea, ING1, wrongly identified as ING1b before alternative splicing products of ING1 were identified, has been previously reported to be 8- to 10-fold higher in senescent than in young, proliferating human diploid fibroblasts [8]. Inhibiting ING1 gene expression by antisense RNA against the 942-1124 nucleotide region of the ING1 transcript allowed normal primary fibroblasts to undergo additional population

doublings when approaching senescence [8] similar to observations made when blocking the function of the p53 tumor suppressor [93]. Interestingly, the retroviral antisense construct was targeted to the C-terminal region of ING1, which corresponds to the common exon of the ING1 gene. This means that the antisense resulted in the inhibition of all ING1 isoforms, including ING1a, which were not fully characterized at the time of the study. Later on, it was discovered that the original ING1 sequence used in this study was a chimera of the common exon 2 and part of the N-terminus of ING1a splice variant. The sequence used lacked the PIP domain, which may explain the observation of induction of senescence rather than apoptosis, at least in human diploid fibroblasts, upon overexpression of the hybrid ING1 protein.

Like p53, ING1 has been shown to be a part of both HAT and HDAC complexes that mediate chromatin-remodeling and alter gene expression [3]. ING1a expression and association with the HDAC1 complex increase several-fold in senescent cells, and this is accompanied by an increase in ING1a-associated HDAC activity [18]. Since the protein levels of ING1a increase in senescent cells, the increased association of ING1a with HDAC1 may be due to a mass action effect, or alternatively to a preferential affinity of ING1a for the HDAC1 complex, or both. ING1b has been reported to bind to Sin3-HDAC1/2 complexes in young proliferating cells [47]. The conundrum may be explained in light of previous observations that in young cells, where the ING1b levels are higher than ING1a, ING1b can bind both HAT and HDAC complexes but with preferential binding affinity to different HAT complexes such as TRRAP, PCAF and CBP/p300 [18]. It is also possible that since ING1b and ING1a appear to act antagonistically in some assays, such as in association with PCNA and subsequent induction of apoptosis, ING1b may inhibit HDAC activity, which has recently been reported [205]. In fact, transient overexpression of ING1b induces the hyperacetylation of histones resulting in transactivation of cell-cycle regulatory and apoptosis-promoting genes such as p21 and Bax [13]. Furthermore, HDAC inhibitors, which can induce euchromatin formation, can also induce a senescence-like phenotype [133,206]. Moreover, HDAC1 has been linked to senescence in some strains of human fibroblasts [207]. Taken together, this suggests that a fine balance between HAT and HDAC localization and activity may maintain a chromatin environment that promotes the senescence phenotype.

The phenomenon of splice variants of certain genes associated with opposing biological effects is not uncommon. For example, p47, a splice isoform of p53 that lacks the N-terminal transactivation domain of p53, appears to functionally inhibit p53 activity, relieving its growth suppressive activity [208]. Another example is the tumor suppressor TID1. While the TID1L isoform induces apoptosis in human osteosarcoma cancer cells, another splicing isoform, TID1S, has been shown to inhibit apoptosis [209]. This raises the possibility that some splice variants, such as ING1a and ING1b, may actually act as dominant negatives for one another under certain physiological or stress conditions.

Overexpression of INGla induced a senescence-like phenotype over time as characterized by: cells displaying the formation of SAHF containing HP1γ, a flat morphology, large nuclei, expression of SA-β-gal, G1 phase arrest of the cell cycle, and induction of pl6 and Rb. The reason why ING1a appears to gradually cause G1 arrest in the cells over a 48 hr time period from 31% to 67% is unknown. A plausible explanation is the need to accumulate chromatin modifying complexes at specific gene promoters over time thereby inducing the observed effect. The regulation of chromatin structure is likely to be sensitive to the stoichiometry of components in the complexes involved.

ING1a binds to HDAC1, which is associated with gene transcription repression, but it induces pRb and p16. This appears paradoxical given that ING1b binds with the transcription co-activator CBP/p300. Moreover, ING1b has been shown to activate and repress similar numbers of genes [52], presumably through promoter-specific effects of associated HAT and HDAC complexes. Many reports have shown that genes encoding cell cycle regulatory proteins are repressed in senescent cells while other genes are upregulated [210,211]. The ING1a-HDAC1 complex may be localized to specific gene promoters that are associated with cell cycle progression. On the other hand, other HAT and/or HDAC complexes are competing for binding with other cell-cycle inhibitory gene promoters and activating their expression. Based upon this study, it is also tempting to speculate that ING1a and ING1b may compete for an overlapping set of targets on chromatin. Whereas chromatin-immunoprecipitation would provide the best link between the ING1 variants and specific gene expression, it was difficult to perform due to the lack of highly specific antibodies for

both isoforms. This project is currently being pursued and will serve as the basis for future students in the lab.

The effects of human ING1 isoforms on apoptosis and senescence are consistent with differential effects seen on histone acetylation [18] and gene expression [170]. However, in contrast to the observation that ING1a-but not INGlb-induces senescence, a recent report in which fragments of INGlb were fused to a DNA-binding domain has also implicated the ING1b isoform in growth arrest and the induction of senescence [212]. The authors used senescent-associated β-gal (SA-β-gal) staining as a measure of ING1b's ability to induce senescence. However, SA-β-gal can also be induced by other stimuli such as DNA damage, oncogene expression and cell confluence in culture, and, thus may also be a measure of ING1b-induced apoptotic stress. Consistent with this, we have also observed the ability of ING1b to induce low levels of SA-β-gal staining in pre-apoptotic cells, however the percentage of cells expressing SA-β-gal is considerably lower than those in response to ING1a (Figure 8). However, while the ING1b isoform can induce an acute G1 phase growth arrest, this arrest is quickly followed by apoptosis [7], and we show that ING1a can induce a permanent cell cycle arrest resembling senescence by modulating chromatin structure and inducing the formation of heterochromatic foci (Figure 7).

The effect of ING1a is similar to that seen in response to the related ING2 protein that also plays a role in regulating senescence [9], and suggests that the ING family of type II tumor suppressors may contribute to inducing senescence through several pathways involving histone acetylation as previously proposed [29]. Consistent with this idea and our current data, it was recently reported that deletion of the Ing1 locus from murine cells impairs Ras-induced senescence [213]. Hints as to how INGl, and the INGla splice isoform in particular, contribute to the establishment of a senescent state are provided by the ability of INGla to induce expression of some genes such as the pl6 and pRb. These two proteins are associated with the induction of senescence, upon ectopic expression. It is tempting to speculate that after initiation of a senescence program by telomere attrition, certain numbers of tumor suppressors, such as the INGl and pRb tumor suppressors, play a role in promoting cellular senescence by regulating the expression of specific subsets of genes (through both inhibitory and excitatory mechanisms) that govern cell growth and proliferation. Ultimately,

it would be valuable to produce knockouts of ING1a or its corresponding gene in multiple organisms in order to test whether ING1a is essential for senescence and how the deletion of ING1a can affect cell cycle gene expression. This, however, may be beyond the scope of our current study.

Although both ING1a and ING1b have a similar C-terminal structure, it is not clear why the ING1a isoform can specifically induce senescence. The unique N-terminus of ING1b has a PCNA-interaction protein motif that is not present in ING1a and is responsible for UV-induced ING1b-mediated apoptosis [15]. We tried different bioinformatics programs to search for potential domains in the ING1a unique N-terminal region, but we could not detect any. Using the PhosNetK phosphorylation-prediction tool, however, we found several potential phosphorylation sites in this region. Potential protein kinases responsible for ING1a N-terminus phosphorylation include protein kinase B and C (PKC and PKB), casein kinase II (CKII), and cyclin-dependent kinase 5 (cdk5). Since PKB/Akt has been shown to induce a senescence-like morphology, at least in epithelial cells [214], and the ING1a N-terminus has an Akt consensus R-x-R-x-x-S motif, it will be interesting to investigate whether ING1a is a downstream target of PI3K/Akt pathway, and whether the Akt-ING1a link can, at least in part, account for ING1a-induced senescence.

Although gene microarray experiments are less sensitive that individual direct assays such as real-time PCR, the list of genes up- and downregulated by ING1a fits with its proposed role in mediating senescence. For example, cyclin E and cdk8 are among the cell-cycle associated proteins that are inhibited by ING1a. Moreover, members of the mitogen activated kinase family, such as MAP4K1 and MAP4K5, are downregulated. The larger number of genes downregulated indicates a functional role of ING1a-HDAC association in gene silencing. The diversity of biological pathways affected by ING1a supports the notion that ING1a-induced senescence is mediated to a great extent through association with different histone-modifying complexes that modulate gene expression pattern, culminating in a senescence phenotype.

In summary, these data show that ING1a expression increases when cells approach senescence, while ING1b expression decreases, markedly altering the ratio of these functionally antagonistic isoforms. ING1a overexpression induced multiple markers of

senescence such as senescence-associated heterochromatic foci (SAHF), containing HP1γ, flat morphology, large nuclei, cell cycle arrest in the G1 phase, and an increase in p16 and pRb levels. We conclude that ING1a primarily contributes to the regulation of proliferative lifespan as previously reported for ING1 [8] and ING2 [9], while ING1b largely, or possibly exclusively, contributes to stress-induced apoptosis by differential regulation of HAT and HDAC complexes. This is consistent with the fact that suppression of the murine equivalent of ING1b affects apoptosis, but not senescence [5].

A novel method for identifying relative binding partners of ING

Protein-protein interactions are important determinants of cell functions. Most, if not all, of the biological processes are mediated through protein association and dissociation. The biological functions of a protein are largely determined by the repertoire of proteins it interacts with, and the signaling pathways these interactions mediate. In the past decade, a paradigm shift from descriptive to predictive models in molecular biology occurred with the generation of protein interaction networks (interactomes) for several model organisms such as yeast, fly and worms. A similar impact on our understanding of human biology will occur when the human interactome will be completely elucidated. Not only will new disease-associated proteins be identified, but also novel signaling pathways will be disclosed.

The number of human protein-protein interaction networks is estimated to be around 200,000 with all possible combinations of the 25,000 or so human proteins, not to mention their splice variants and post translational modifications that greatly increase the expected PPI number. Increasing numbers of *in silico* approaches are being developed for predicting novel PPIs. These approaches rely on different factors such as similarity in protein structure, protein sequence, gene neighborhood, and phylogenetic profile (see introduction for further details). The abundance of interaction data from yeast, and to a lesser extent from the fly and worm, encouraged researchers to use lower organism interactomes to predict novel PPIs in humans. Since conserved proteins tend to have conserved interactions, and since yeast is the organism whose interactome has been best surveyed in deep detail, a number of computational tools have been developed to utilize published yeast interactome data to search for conserved “interlogs” in humans. It must be noted that the quality of the predicted

interaction data depends greatly on the quality of the biochemical method used to generate a specific yeast interactome.

In this study, we asked whether we could use the available yeast interactome to predict novel human ING protein interactors. We first investigated the similarities between the three yeast INGs (YNG1, YNG2, and YNG3/Pho23) and human INGs (ING1-5). Using several bioinformatics alignment tools, we found that ING1 and 2 are more similar to YNG1, ING4 and 5 show higher similarities to YNG2, and YNG3/Pho23 is the closest homologue to human ING3. Although the difference among the YNGs and INGs homology scores was close, this classification is in line with previous reports categorizing human ING into three groups according to their phylogenetic relationships [1] and functional association with different HAT and HDAC complexes [30]. We then examined the conservation of human ING domains in yeast and two other species, namely fly and worms, as the interactomes of the three species have been analyzed by more than one group and are considered to have the most consolidated lists of PPIs. We found a high degree of similarities among human, yeast and fly ING domains. The PHD showed the highest degree of conservation followed by Novel Conserved Region (NCR, now known as lamin-interaction domain or LID). To our surprise, *C. elegans* had only one clear ING-like protein, with a PHD domain that is highly similar to human ING. This result is surprising given that the *C. elegans* is phylogenetically closer to humans than yeast. This may be due to the use of human domain protein sequences to examine the degree of similarity with worm domains. Although domains are more or less similar in different species, our approach may have excluded worm ING domains that show similar conserved core residues but different flanking regions. Indeed, using YOGY, a web-based database that allows for searching of orthologues for human proteins, we identified only one protein based upon the ING PHD domain.

We then used the yeast interactome published by Krogan *et al.* to identify novel ING interacting partners in yeast. The yeast interactome was used for three reasons: 1) yeast is a prototypic eukaryotic organism that can act as a model for human cells, 2) many yeast proteins and protein complexes have homologes in humans, and 3) the yeast interactome has been extensively investigated by more than one group. This has given us more flexibility in comparing and validating different published interactomes. Although many reports have

investigated yeast PPIs, the Krogan *et al.* data represents the first genome-wide screen of an organism. *S. cerevisiae* strains were generated with in-frame insertions of TAP tags individually introduced by homologous recombination of each predicted open reading frame (ORF), i.e. each protein was expressed from its natural promoter, and proteins were purified from yeast cultures under native conditions. Finally the identities of the co-purifying proteins were determined using two complementary techniques. Each purified protein sample containing the protein of interest and its interacting partners was then electrophoresed using SDS-PAGE, stained with silver, and visible bands were removed and identified by trypsin digestion and peptide mass-fingerprinting using matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry. In parallel, another aliquot of each purified protein preparation was digested in solution and the peptides were separated and sequenced by data-dependent liquid chromatography tandem mass spectrometry (LC-MS/MS). Among 4,087 different proteins identified with high confidence by mass spectrometry from 2,357 successful purifications, the core data set comprises 7,123 PPIs involving 2,708 proteins with high confidence (99%) by MALDI–TOF mass spectrometry and/or LC-MS/MS, altogether corresponding to 72% of the predicted yeast proteome. The PPIs were organized into 547 complexes, with an average of 5 subunits per complex. The authors' success rate for identifying proteins was 94% and 89% for nuclear and cytosolic proteins, respectively, and at least 70% of the proteins in most subcellular compartments were successfully purified.

The reason for our using the Krogan *et al.* data was because it has many advantages over other published yeast interactomes. Compared to other TAP-tagging/MS approaches previously used [137], yeast proteins were expressed and purified under natural conditions avoiding the caveats of overexpression. The use of two mass spectrometry approaches greatly increased the confidence and the reliability of their interactome data. A major strength of this study over the other reports [191] was that we were able to access the raw affinity-mass spectrometry data before data curation by implementing a cut-off value. This allowed us to search for weak, yet valid interactions and investigate whether any of the transient yeast ING interactors have strong orthologes in humans. The Krogan *et al*. raw data contained around 1,075 yeast ING interactors [164] making it a very fertile ground for data mining as compared to Gavin *et al*. with only 12 ING interactions [191]. The significant

difference between the two genome-wide affinity/MS reports may be due to the fact that different complexes are recovered even when the same tagged protein is purified repeatedly. For example, when 139 purifications were repeated, only 70% were common to the two runs [215]. Another possibility is that one tagged protein may pull down several independent complexes that appear as one large complex. These differences support the notion that raw interaction data from different interactome projects should be released, besides the filtered high-confidence data. In this context, weak interactions conserved among different species can define real from false interactions [143]

A major reason why we did not use PPI datasets generated by the Y2H approach is the conflicting results of Y2H results, where the overlap between the first two Y2H screen yeast interactomes was only 7% [135,136,141]. This may be due to the use of PCR-generated ORFs that may bear mutations in one screen and produce proteins unable to interact with their physiological targets. Besides, false positives are not uncommon in Y2H biological screens due to the sensitivity of the technique and the different mechanisms that the organism may have for promoting survival. In addition, each of the Y2H screens used a different and unique plasmid construct which may have affected protein folding. Improperly folded proteins may interact with the wrong proteins or may not interact at all. A final observed caveat is that systematic large-scale Y2H studies may fail to recapitulate many of yeast interactions identified by conventional methods, which may be due to the large-scale nature of the approach, or may be inherent in the technique itself. We found very few yeast ING interactors in datasets produced by Y2H screens as indicated in Table 6.

In order to filter the yeast ING 1075 interactors to those having homologues in human, we used MAGPIE (multipurpose automated genome project investigation environment), a system designed to assign many functions to the sequence data. It shows a great degree of flexibility as it is for the analysis of datasets in a variety of formats, and the generated inputs of which can be accessed through WWW browsers (see Table 5). We used only the taxonomic tool of MAGPIE, which has allowed us to align each of the yeast ING interacting proteins against the human proteome in a short time. We found around 380 proteins that interacted with yeast ING and have homologues in humans. This number includes proteins that may interact with human INGs directly or indirectly, or may even be

subunits in large complexes containing ING. In order to narrow down the list to interactions that may have functional significance, we re-filtered the human ING interacting protein list against the fly database, a metazoan organism that has ING-like proteins with high similarity to human and yeast ING proteins and has available interactome datasets. We found 3 fly ING-like proteins interacting with around 36 proteins. CG7379, the human ING2-like homologue in fly, interacts with 26 fly proteins. CG9293, the ING1-like protein in fly has 4 interacting partners and CG6632, the ING3-like fly protein shows 6 PPIs. Although five fly ING interacting proteins have counterparts in yeast, the three ING potential interacting proteins showed a high degree of conservation in humans, yeast, and fly. The three genes have 5 potential homologues in humans ranging from cytoplasmic to nuclear proteins as indicated in Figure 16.

This novel *in silico* approach allowed us to predict new protein-protein interactions for the human INGs with a high degree of success, and confirmed many previously elucidated interactions, such as those with p21, Karyopherin, HAT/HDAC proteins and histone H3 (Table 5). Our findings suggest that ING family proteins are involved in a more diverse array of biological processes than are presently suspected from the current literature and some of the interactions suggest possible additional mechanisms that might underlie their tumor suppressor capabilities. The two new interactions we have elucidated and biochemically confirmed here, p38MAPK and MEKK4, further confirm ING1 role in DNA damage/stress response pathways.

In an attempt to understand the connection between the ING, MEKK4, and p38MAPK protein interaction networks, we generated a merged interaction graph (Figure 20). Several reports have indicated that different forms of stress, such as UV, chemotherapeutic agents and hypoxia affect the function of the ING proteins [12,15,216]. The mammalian JNK/p38MAP kinase kinase kinase (MEKK4) and the *S. cerevisiae* Ssk2p protein are homologous, with MEKK4 being able to replace all of the known functions of Ssk2p in yeast. The stress-activated mitogen-activated protein kinase (SAPK) pathways are integral components of diverse stress signaling pathways such as UV, hypoxia, heat, osmotic shock, pH, oxidative damage, cytokines, pheromones and others [190,217]. The fact that ING1 can interact with both MEKK4 and p38MAPK is not surprising given the facts that all

three proteins are evolutionarily conserved, bear common links to several different signaling pathways, both ING1 and MEKK4 bind to GADD45 [16,218] and both MEKK4 and p38MAPK are in a well defined stress response pathway [217]. This observation is also consistent with the observation that ING proteins affect transcription factor activity [13] since the MEKK4/p38MAPK stress activated kinase cascade culminates in the regulation of various transcription factors, some of which are outlined in Figure 20. Analysis of the effects of altering ING1 activity on MEKK4/p38MAPK signaling under different conditions of extracellular and intracellular stress should serve to better clarify the roles that physical interaction of ING1 with these proteins play in the mammalian stress response cascades. It will be interesting to identify interaction regions in ING, p38MAPK, and MEKK4. Since several other ING-interacting partners showed similar degrees of interaction, it is tempting to speculate that further examination of the additional candidate ING-interacting pathways we have identified in multiple model organisms, and particularly in yeast, will reveal more regarding the function of the ING family of chromatin regulators.

In summary, recent screens in yeast and other organisms have generated considerable data regarding protein interactions, but it is frequently difficult to predict what interactions are biologically significant and will stand the scrutiny of biochemical validation. Thus, there are considerable amounts of weak interaction data from large-scale interaction datasets, in particular, those are currently underutilized and might be useful in predicting PPI with a high degree of confidence. The ING family of proteins is involved in many cellular activities; for a small number of cases, however, the proteins with which they interact are not well described, and so the biochemical pathways in which they are involved remain to be well defined. We developed a method using comparisons in different organisms in which homologs exist, to predict with a high degree of certainty what particular protein interactions found in unfiltered data may occur *in vivo* and contribute to the activities of the ING1 gene products. This cross-species (yeast, fly, and human) bioinformatics-based approach was used to identify potential human ING1 interacting proteins with higher probability and accuracy than approaches based on screens in a single species. We confirm the validity of this screen and show that ING1 interacts specifically with two proteins tested: p38MAPK, and MEKK4 at the endogenous and overexpression levels. These novel ING interacting proteins further

link ING proteins to cell stress and the DNA damage signaling, providing previously unknown upstream links to DNA damage response pathways in which ING1 participates.

Conclusion and Significance

ING1 was initially discovered by screening for genes whose expression was suppressed in breast cancer cells. Knocking down ING1 expression with antisense mRNA promotes cellular growth and transformation of mouse mammary epithelial cells *in vitro* and tumour formation *in vivo.* These results suggest an important role of ING1 in promotion of growth arrest and tumour suppression. The two predominant ING1 isoforms are ING1a and ING1b. Both ING1a and ING1b appear to have opposing functions and different cellular partners. Based on our data, ING1a may play a unique role in the establishment of stable cell cycle arrest and is transcriptionally activated in a senescence-specific manner in contrast to ING1b which has been shown in many reports to induce transient cell cycle arrest followed by apoptosis. The overall goal of this research is to investigate the role of ING1a in promotion of senescence which will give a better idea about the different mechanisms by which ING1 isoforms can perform their tumor suppressor functions. Four specific aims were designed as follows: 1) determine whether ING1a is upregulated in senescent human fibroblasts, 2) determine whether ING1a can induce senescence-like features, 3) examine the molecular mechanism mediating ING1a growth arrest, and 4) compare the altered gene expression profile of ING1a infected cells with normal replicative senescent cells. It is hoped that this study will help to define the antagonistic biochemical roles proposed for the two major ING1 isoforms. Moreover, it may provide a better understanding of organismal aging since cell senescence is considered to be an underlying mechanism of organismal aging.

Since some experimental studies showed that downregulation and upregulation of ING1 levels can enhance or suppress tumorigenesis respectively, understanding the mechanisms by which ING1 isoforms carry out their tumor suppressor functions will give insights to design better cancer treatments. Indeed, it now becomes apparent that senescence can be readily induced in tumor cells by genetic manipulation or different methods of cancer treatment such as chemotherapeutics and radiation. Clinical studies showed that expression of different senescence-regulatory genes not only has promising prognostic applications but

also can induce senescence in tumour cell lines. Not surprising, many of these genes have a role in senescence pathways in normal fibroblasts, the model in which cellular senescence has been most extensively studied. This raises the hope that ING1a can be used to induce senescence in specific tumors to prevent premalignant lesions from progressing to malignant phases or to be used as a prognostic tool in patient receiving chemotherapeutics acting through senescence-induction such as Hydroxyurea and Doxorubicin.

In collaboration with the local bioinformatics group, we focused on designing a cross-species bioinformatics-based approach to identify potential human ING interacting proteins with higher probability and accuracy than approaches based on screens in a single species. Our data suggest a high degree of conservation of the ING proteins exists between human and yeast. Using TAP-tagged proteins and different types of Mass Spectrometry, a large consortium found over 1,000 proteins that interacted with yeast INGs, and of these proteins, we found that 381 had identified homologs in human cells. The data presented in this study 1) highlight the fact that many potential and novel interactions may occur between analogous proteins in different species, 2) show a novel *in silico* approach we designed, which allowed us to predict new protein-protein interactions for the human INGs with a high degree of success, 3) confirm many previously elucidated ING-interactions such as with p21, Karyopherin, HAT/HDAC proteins and histone H3, and 4) suggests that ING family proteins are involved in a more diverse array of biological processes than are presently suspected from the current literature. We trust that examination in greater detail of the interactions discovered in this study will be useful in understanding the contributions of the ING family of epigenetic regulators to the complex processes of cell senescence and tumorigenesis.

PUBLICATIONS PRODUCED DURING THE COURSE OF THIS THESIS

Manuscripts (Published/Accepted)

1- **Soliman MA**. and Riabowol K. (2007) After a decade of study-ING, a PHD for a versatile family of proteins. **Trends in Biochemical Sciences.** 32, 509-519**.**

2- Tallen G., **Soliman MA.** and Riabowol K. (2007) The Cancer-Aging interface and the significance of telomere dynamics in cancer therapy. **Rejuvenation Research.** 10, 387-95.

3- Russell MW., **Soliman MA.**, Schriemer D., and Riabowol K. (2008) Targeting of ING1 to the nucleus by karyopherins is necessary for activation of p21. **Biochemical and Biophysical Research Communications.** 374(3), 490-495.

4- **Soliman MA.** *, Berardi P. *, Pastyryeva S., Bonnefin P., Feng X., Colina A., Young D. and Riabowol K. (2008) ING1a expression increases during senescence and induces a senescent phenotype. **Aging Cell.** *In press.* (First two authors contributed equally to the work).

5- Gordon PM. *, **Soliman MA.** *, Bose P., Trinh Q., Sensen CW. and Riabowol K. (2008) Interspecies data mining to predict novel ING-protein interactions in human. **BMC Genomics.** *In press* (First two authors contributed equally to the work).

6- Han X., Feng X., Rattner JB., Smith H., Bose P., Suzuki K., **Soliman MA**., Scott M., Burke BE. and Riabowol K. (2008) Tethering of the ING1 PHD protein to lamin A is required for ING1 stability and function in apoptosis. **Nature Cell Biology.** *In press.*

Published Abstracts

Soliman MA., Berardi P., Feng X. and Riabowol K. **2008**
The ING1 Tumor Suppressors: Link-ING Chromatin Remodeling to Cellular Senescence. **AACR meeting**, Boston, Massachusetts, USA

Soliman MA., Gordon PM., Bose P, Sensen CW. and Riabowol K. **2007**
Interspecies data mining to predict novel protein-protein interactions in human
Cold Spring Harbor Laboratory/Welcome Trust Conference on Genome Informatics**,**
CSHL, New York, USA

Soliman MA., Berardi P. and Riabowol K. **2007**
ING1a as a novel senescence marker
Gordon Research Conference, Les Diablerets conference centre, Switzerland

Soliman MA., Berardi P. and Riabowol K. **2007**
Differential splicing of ING1 tumour suppressor in aging and senescence
Keystone Symposium, Whistler resort, British Columbia, Canada

Posters and Abstracts, Local Conferences

Soliman MA., Berardi P., Feng X. and Riabowol K. **2008**
ING1 mediates cellular senescence through chromatin remodeling
Alberta Genomic Instability and Aging Conference, Alberta, Canada

Soliman MA., Berardi P., Feng X., and Riabowol K. **2008**
The tumor suppressor ING1 links senescence to chromatin remodelling
The Canadian Society of Biochemistry and Molecular & Cellular Biology meeting,
Banff, Alberta, Canada

Soliman MA., Berardi P. and Riabowol K. **2007**
ING1 splicing isoforms are differentially regulated during Senescence.
Alberta Cancer Board Meeting, Banff, Alberta, Canada

Reference List

1 He,G.H. *et al.* (2005) Phylogenetic analysis of the ING family of PHD finger proteins. *Mol. Biol. Evol.* 22, 104-116

2 Campos,E.I. *et al.* (2004) Biological functions of the ING family tumor suppressors. *Cell Mol. Life Sci.* 61, 2597-2613

3 Soliman,M.A. and Riabowol,K. (2007) After a decade of study-ING, a PHD for a versatile family of proteins. *Trends Biochem. Sci.* 32, 509-519

4 Kichina,J.V. *et al.* (2006) Targeted disruption of the mouse ing1 locus results in reduced body size, hypersensitivity to radiation and elevated incidence of lymphomas. *Oncogene* 25, 857-866

5 Coles,A.H. *et al.* (2007) Deletion of p37Ing1 in mice reveals a p53-independent role for Ing1 in the suppression of cell proliferation, apoptosis, and tumorigenesis. *Cancer Res.* 67, 2054-2061

6 Garkavtsev,I. *et al.* (1996) Suppression of the novel growth inhibitor p33ING1 promotes neoplastic transformation. *Nat Genet.* 14, 415-420

7 Helbing,C.C. *et al.* (1997) A novel candidate tumor suppressor, ING1, is involved in the regulation of apoptosis. *Cancer Res.* 57, 1255-1258

8 Garkavtsev,I. and Riabowol,K. (1997) Extension of the replicative life span of human diploid fibroblasts by inhibition of the p33ING1 candidate tumor suppressor. *Mol. Cell Biol.* 17, 2014-2019

9 Pedeux,R. *et al.* (2005) ING2 regulates the onset of replicative senescence by induction of p300-dependent p53 acetylation. *Mol. Cell Biol.* 25, 6639-6648

10 Garkavtsev,I. *et al.* (1998) The candidate tumour suppressor p33ING1 cooperates with p53 in cell growth control. *Nature* 391, 295-298

11 Shinoura,N. *et al.* (1999) Adenovirus-mediated transfer of p33ING1 with p53 drastically augments apoptosis in gliomas. *Cancer Res.* 59, 5521-5528

12 Nagashima,M. *et al.* (2001) DNA damage-inducible gene p33ING2 negatively regulates cell proliferation through acetylation of p53. *Proc. Natl. Acad. Sci. U. S. A* 98, 9671-9676

13 Kataoka,H. *et al.* (2003) ING1 represses transcription by direct DNA binding and through effects on p53. *Cancer Res.* 63, 5785-5792

14 Gonzalez,L. *et al.* (2006) A functional link between the tumour suppressors ARF and p33ING1. *Oncogene* 25, 5173-5179

15 Scott,M. *et al.* (2001) UV-induced binding of ING1 to PCNA regulates the induction of apoptosis. *J. Cell Sci.* 114, 3455-3462

16 Cheung,K.J., Jr. *et al.* (2001) The tumor suppressor candidate p33(ING1) mediates repair of UV-damaged DNA. *Cancer Res.* 61, 4974-4977

17 Simpson,F. *et al.* (2006) The PCNA-associated factor KIAA0101/p15(PAF) binds the potential tumor suppressor product p33ING1b. *Exp. Cell Res.* 312, 73-85

18 Vieyra,D. *et al.* (2002) Human ING1 proteins differentially regulate histone acetylation. *J. Biol. Chem.* 277, 29832-29839

19 Shiseki,M. *et al.* (2003) p29ING4 and p28ING5 bind to p53 and p300, and enhance p53 activity. *Cancer Res.* 63, 2373-2378

20 Warbrick,E. (2000) The puzzle of PCNA's many partners. *Bioessays* 22, 997-1006

21 Gozani,O. *et al.* (2003) The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. *Cell* 114, 99-111

22 Kaadige,M.R. and Ayer,D.E. (2006) The polybasic region that follows the plant homeodomain zinc finger 1 of PF1 is necessary and sufficient for specific phosphoinositide binding. *J. Biol. Chem.*

23 Jones,D.R. *et al.* (2006) Nuclear PtdIns5P as a transducer of stress signaling: an in vivo role for PIP4Kbeta. *Mol. Cell* 23, 685-695

24 Shi,X. *et al.* (2006) ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* 442, 96-99

25 Pena,P.V. *et al.* (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100-103

26 Martin,D.G. *et al.* (2006) The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. *Mol. Cell Biol.* 26, 7871-7879

27 Palacios,A. *et al.* (2006) Solution structure and NMR characterization of the binding to methylated histone tails of the plant homeodomain finger of the tumour suppressor ING4. *FEBS Lett.* 580, 6903-6908

28 Huang,W. *et al.* (2007) Stabilized phosphatidylinositol-5-phosphate analogues as ligands for the nuclear protein ING2: chemistry, biology, and molecular modeling. *J. Am. Chem. Soc.* 129, 6498-6506

29 Feng,X. *et al.* (2002) Different HATS of the ING1 gene family. *Trends Cell Biol.* 12, 532-538

30 Doyon,Y. *et al.* (2006) ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol. Cell* 21, 51-64

31 Kuo,W.H. *et al.* (2007) The ING1b tumor suppressor facilitates nucleotide excision repair by promoting chromatin accessibility to XPA. *Exp. Cell Res.* 313, 1628-1638

32 Wang,J. *et al.* (2006) The novel tumor suppressor p33ING2 enhances nucleotide excision repair via inducement of histone H4 acetylation and chromatin relaxation. *Cancer Res.* 66, 1906-1911

33 Pena,P.V. *et al.* (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100-103

34 Li,H. *et al.* (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442, 91-95

35 Wysocka,J. *et al.* (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86-90

36 Bannister,A.J. *et al.* (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120-124

37 Lachner,M. *et al.* (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116-120

38 Cao,R. *et al.* (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043

39 Czermin,B. *et al.* (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185-196

40 Barlev,N.A. *et al.* (2001) Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Mol. Cell* 8, 1243-1254

41 Monte,M. *et al.* (2006) MAGE-A tumor antigens target p53 transactivation function through histone deacetylase recruitment and confer resistance to chemotherapeutic agents. *Proc. Natl. Acad. Sci. U. S. A* 103, 11160-11165

42 Wagner,M.J. and Helbing,C.C. (2005) Multiple variants of the ING1 and ING2 tumor suppressors are differentially expressed and thyroid hormone-responsive in Xenopus laevis. *Gen. Comp Endocrinol.* 144, 38-50

43 Unoki,M. *et al.* (2006) Novel splice variants of ING4 and their possible roles in the regulation of cell growth and motility. *J. Biol. Chem.* 281, 34677-34686

44 Wang,Y. *et al.* (2006) Leucine zipper-like domain is required for tumor suppressor ING2-mediated nucleotide excision repair and apoptosis. *FEBS Lett.* 580, 3787-3793

45 Pena,P.V. *et al.* (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature* 442, 100-103

46 Bienz,M. (2006) The PHD finger, a nuclear protein-interaction domain. *Trends Biochem. Sci.* 31, 35-40

47 Kuzmichev,A. *et al.* (2002) Role of the Sin3-histone deacetylase complex in growth regulation by the candidate tumor suppressor p33(ING1). *Mol. Cell Biol.* 22, 835-848

48 Scott,M. *et al.* (2001) UV induces nucleolar translocation of ING1 through two distinct nucleolar targeting sequences. *Nucleic Acids Res.* 29, 2052-2058

49 Russell,M. *et al.* (2006) Grow-ING, Age-ING and Die-ING: ING proteins link cancer, senescence and apoptosis. *Exp. Cell Res.* 312, 951-961

50 Gong,W. *et al.* (2006) Subcellular targeting of p33ING1b by phosphorylation-dependent 14-3-3 binding regulates p21WAF1 expression. *Mol. Cell Biol.* 26, 2947-2954

51 Gong,W. *et al.* (2005) Function of the ING family of PHD proteins in cancer. *Int. J. Biochem. Cell Biol.* 37, 1054-1065

52 Feng,X. *et al.* (2006) HSP70 induction by ING proteins sensitizes cells to tumor necrosis factor alpha receptor-mediated apoptosis. *Mol. Cell Biol.* 26, 9244-9255

53 Leung,K.M. *et al.* (2002) The candidate tumor suppressor ING1b can stabilize p53 by disrupting the regulation of p53 by MDM2. *Cancer Res.* 62, 4890-4893

54 Tsang,F.C. *et al.* (2003) ING1b decreases cell proliferation through p53-dependent and -independent mechanisms. *FEBS Lett.* 553, 277-285

55 Vaziri,H. *et al.* (2001) hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. *Cell* 107, 149-159

56 Serrano,M. *et al.* (1997) Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* 88, 593-602

57 Vieyra,D. *et al.* (2002) ING1 isoforms differentially affect apoptosis in a cell age-dependent manner. *Cancer Res.* 62, 4445-4452

58 Irvine,R.F. (2003) Nuclear lipid signalling. *Nat. Rev. Mol. Cell Biol.* 4, 349-360

59 Papayannopoulos,V. *et al.* (2005) A polybasic motif allows N-WASP to act as a sensor of PIP(2) density. *Mol. Cell* 17, 181-191

60 Jenuwein,T. and Allis,C.D. (2001) Translating the histone code. *Science* 293, 1074-1080

61 Martin,C. and Zhang,Y. (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838-849

62 Flanagan,J.F. *et al.* (2005) Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 438, 1181-1185

63 Pray-Grant,M.G. *et al.* (2005) Chd1 chromodomain links histone H3 methylation with. *Nature* 433, 434-438

64 Eberharter,A. *et al.* (2004) ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. *EMBO J.* 23, 4029-4039

65 Li,B. *et al.* (2007) Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* 316, 1050-1054

66 Shi,X. *et al.* (2007) Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J. Biol. Chem.* 282, 2450-2455

67 Taverna,S.D. *et al.* (2006) Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol. Cell* 24, 785-796

68 Vakoc,C.R. *et al.* (2005) Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell* 19, 381-391

69 Sims,R.J., III and Reinberg,D. (2006) Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev.* 20, 2779-2786

70 Huang,Y. *et al.* (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science* 312, 748-751

71 Wysocka,J. *et al.* (2005) WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121, 859-872

72 Xin,H. *et al.* (2004) Components of a pathway maintaining histone modification and heterochromatin protein 1 binding at the pericentric heterochromatin in Mammalian cells. *J. Biol. Chem.* 279, 9539-9546

73 Bunce,M.W. *et al.* (2006) Stress-ING out: phosphoinositides mediate the cellular stress response. *Sci. STKE.* 2006, e46

74 Ruthenburg,A.J. *et al.* (2007) Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol. Cell* 25, 15-30

75 Hayflick,L. and MOORHEAD,P.S. (1961) The serial cultivation of human diploid cell strains. *Exp. Cell Res.* 25, 585-621

76 Campisi,J. (2005) Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors. *Cell* 120, 513-522

77 Blackburn,E.H. (2001) Switching and signaling at the telomere. *Cell* 106, 661-673

78 Griffith,J.D. *et al.* (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97, 503-514

79 Bodnar,A.G. *et al.* (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science* 279, 349-352

80 Deng,Y. *et al.* (2008) Telomere dysfunction and tumour suppression: the senescence connection. *Nat. Rev. Cancer* 8, 450-458

81 Harley,C.B. *et al.* (1990) Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458-460

82 Di,L.A. *et al.* (1994) DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts. *Genes Dev.* 8, 2540-2551

83 Dumont,P. *et al.* (2000) Induction of replicative senescence biomarkers by sublethal oxidative stresses in normal human fibroblast. *Free Radic. Biol. Med.* 28, 361-373

84 Mooi,W.J. and Peeper,D.S. (2006) Oncogene-induced cell senescence--halting on the road to cancer. *N. Engl. J. Med.* 355, 1037-1046

85 Campisi,J. (2003) Analysis of tumor suppressor gene-induced senescence. *Methods Mol. Biol.* 223, 155-172

86 Ben-Porath,I. and Weinberg,R.A. (2005) The signals and pathways activating cellular senescence. *Int. J. Biochem. Cell Biol.* 37, 961-976

87 Kinzler,K.W. and Vogelstein,B. (1997) Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* 386, 761, 763

88 Campisi,J. (2003) Cancer and ageing: rival demons? *Nat Rev. Cancer* 3, 339-349

89 Lundberg,A.S. *et al.* (2000) Genes involved in senescence and immortalization. *Curr. Opin. Cell Biol.* 12, 705-709

90 Hahn,W.C. *et al.* (1999) Creation of human tumour cells with defined genetic elements. *Nature* 400, 464-468

91 Vaziri,H. *et al.* (1997) ATM-dependent telomere loss in aging human diploid fibroblasts and DNA damage lead to the post-translational activation of p53 protein involving poly(ADP-ribose) polymerase. *EMBO J.* 16, 6018-6033

92 Atadja,P. *et al.* (1995) Increased activity of p53 in senescing fibroblasts. *Proc. Natl. Acad. Sci. U. S. A* 92, 8348-8352

93 Gire,V. and Wynford-Thomas,D. (1998) Reinitiation of DNA synthesis and cell division in senescent human fibroblasts by microinjection of anti-p53 antibodies. *Mol. Cell Biol.* 18, 1611-1621

94 Sherr,C.J. (1998) Tumor surveillance via the ARF-p53 pathway. *Genes Dev.* 12, 2984-2991

95 Kamijo,T. *et al.* (1997) Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* 91, 649-659

96 Munro,J. *et al.* (1999) Role of the alternative INK4A proteins in human keratinocyte senescence: evidence for the specific inactivation of p16INK4A upon immortalization. *Cancer Res.* 59, 2516-2521

97 Dimri,G.P. *et al.* (2000) Regulation of a senescence checkpoint response by the E2F1 transcription factor and p14(ARF) tumor suppressor. *Mol. Cell Biol.* 20, 273-285

98 Tahara,H. *et al.* (1995) Increase in expression level of p21sdi1/cip1/waf1 with increasing division age in both normal and SV40-transformed human fibroblasts. *Oncogene* 10, 835-840

99 Noda,A. *et al.* (1994) Cloning of senescent cell-derived inhibitors of DNA synthesis using an expression screen. *Exp. Cell Res.* 211, 90-98

100 Kim,W.Y. and Sharpless,N.E. (2006) The regulation of INK4/ARF in cancer and aging. *Cell* 127, 265-275

101 Wong,H. and Riabowol,K. (1996) Differential CDK-inhibitor gene expression in aging human diploid fibroblasts. *Exp. Gerontol.* 31, 311-325

102 Hara,E. *et al.* (1996) Regulation of p16CDKN2 expression and its implications for cell immortalization and senescence. *Mol. Cell Biol.* 16, 859-867

103 Brenner,A.J. *et al.* (1998) Increased p16 expression with first senescence arrest in human mammary epithelial cells and extended growth capacity with p16 inactivation. *Oncogene* 17, 199-205

104 Bracken,A.P. *et al.* (2007) The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes Dev.* 21, 525-530

105 Gil,J. *et al.* (2004) Polycomb CBX7 has a unifying role in cellular lifespan. *Nat. Cell Biol.* 6, 67-72

106 Jacobs,J.J. *et al.* (1999) The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. *Nature* 397, 164-168

107 Itahana,K. *et al.* (2003) Control of the replicative life span of human fibroblasts by p16 and the polycomb protein Bmi-1. *Mol. Cell Biol.* 23, 389-401

108 Narita,M. *et al.* (2003) Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 113, 703-716

109 Funayama,R. *et al.* (2006) Loss of linker histone H1 in cellular senescence. *J. Cell Biol.* 175, 869-880

110 Stein,G.H. *et al.* (1999) Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts. *Mol. Cell Biol.* 19, 2109-2117

111 Beausejour,C.M. *et al.* (2003) Reversal of human cellular senescence: roles of the p53 and p16 pathways. *EMBO J.* 22, 4212-4222

112 Collado,M. *et al.* (2005) Tumour biology: senescence in premalignant tumours. *Nature* 436, 642

113 Zhu,J. *et al.* (1998) Senescence of human fibroblasts induced by oncogenic Raf. *Genes Dev.* 12, 2997-3007

114 Lin,A.W. *et al.* (1998) Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes Dev.* 12, 3008-3019

115 Collado,M. and Serrano,M. (2006) The power and the promise of oncogene-induced senescence markers. *Nat. Rev. Cancer* 6, 472-476

116 Braig,M. *et al.* (2005) Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* 436, 660-665

117 Michaloglou,C. *et al.* (2005) BRAFE600-associated senescence-like cell cycle arrest of human naevi. *Nature* 436, 720-724

118 Chen,Z. *et al.* (2005) Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* 436, 725-730

119 Janzen,V. *et al.* (2006) Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* 443, 421-426

120 Chen,Q.M. *et al.* (2000) Apoptosis or senescence-like growth arrest: influence of cell-cycle position, p53, p21 and bax in H2O2 response of normal human fibroblasts. *Biochem. J.* 347, 543-551

121 Jackson,J.G. and Pereira-Smith,O.M. (2006) p53 is preferentially recruited to the promoters of growth arrest genes p21 and GADD45 during replicative senescence of normal human fibroblasts. *Cancer Res.* 66, 8356-8360

122 Marcotte,R. *et al.* (2004) Senescent fibroblasts resist apoptosis by downregulating caspase-3. *Mech. Ageing Dev.* 125, 777-783

123 Zhang,H. *et al.* (2003) Senescence-specific gene expression fingerprints reveal cell-type-dependent physical clustering of up-regulated chromosomal loci. *Proc. Natl. Acad. Sci. U. S. A* 100, 3251-3256

124 Dimri,G.P. *et al.* (1995) A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc. Natl. Acad. Sci. U. S. A* 92, 9363-9367

125 Lee,B.Y. *et al.* (2006) Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. *Aging Cell* 5, 187-195

126 Macieira-Coelho,A. (1980) Implications of the reorganization of the cell genome for aging or immortalization of dividing cells in vitro. *Gerontology* 26, 276-282

127 Wilson,V.L. and Jones,P.A. (1983) DNA methylation decreases in aging but not in immortal cells. *Science* 220, 1055-1057

128 Vertino,P.M. *et al.* (1994) Stabilization of DNA methyltransferase levels and CpG island hypermethylation precede SV40-induced immortalization of human fibroblasts. *Cell Growth Differ.* 5, 1395-1402

129 Young,J.I. and Smith,J.R. (2001) DNA methyltransferase inhibition in normal human fibroblasts induces a p21-dependent cell cycle withdrawal. *J. Biol. Chem.* 276, 19610-19616

130 Chan,H.M. and La Thangue,N.B. (2001) p300/CBP proteins: HATs for transcriptional bridges and scaffolds. *J. Cell Sci.* 114, 2363-2373

131 Yao,T.P. *et al.* (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* 93, 361-372

132 Bandyopadhyay,D. *et al.* (2002) Down-regulation of p300/CBP histone acetyltransferase activates a senescence checkpoint in human melanocytes. *Cancer Res.* 62, 6231-6239

133 Ogryzko,V.V. *et al.* (1996) Human fibroblast commitment to a senescence-like state in response to histone deacetylase inhibitors is cell cycle dependent. *Mol. Cell Biol.* 16, 5210-5218

134 Lagger,G. *et al.* (2002) Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *EMBO J.* 21, 2672-2681

135 Uetz,P. *et al.* (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. *Nature* 403, 623-627

136 Ito,T. *et al.* (2000) Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl. Acad. Sci. U. S. A* 97, 1143-1147

137 Ho,Y. *et al.* (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415, 180-183

138 Gavin,A.C. *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141-147

139 Tong,A.H. *et al.* (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* 295, 321-324

140 Tong,A.H. *et al.* (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294, 2364-2368

141 Ito,T. *et al.* (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U. S. A* 98, 4569-4574

142 Fields,S. and Song,O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340, 245-246

143 Hart,G.T. *et al.* (2006) How complete are current yeast and human protein-interaction networks? *Genome Biol.* 7, 120

144 Rigaut,G. *et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030-1032

145 Bader,G.D. *et al.* (2003) BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* 31, 248-250

146 Steen,H. and Mann,M. (2004) The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* 5, 699-711

147 Stark,C. *et al.* (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34, D535-D539

148 Chatr-Aryamontri,A. *et al.* (2007) MINT: the Molecular INTeraction database. *Nucleic Acids Res.* 35, D572-D574

149 Dandekar,T. *et al.* (1998) Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem. Sci.* 23, 324-328

150 Marcotte,E.M. *et al.* (1999) Detecting protein function and protein-protein interactions from genome sequences. *Science* 285, 751-753

151 Aloy,P. and Russell,R.B. (2002) The third dimension for protein interactions and complexes. *Trends Biochem. Sci.* 27, 633-638

152 Ma,B. *et al.* (2003) Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc. Natl. Acad. Sci. U. S. A* 100, 5772-5777

153 Aloy,P. and Russell,R.B. (2003) InterPreTS: protein interaction prediction through tertiary structure. *Bioinformatics.* 19, 161-162

154 Ogmen,U. *et al.* (2005) PRISM: protein interactions by structural matching. *Nucleic Acids Res.* 33, W331-W336

155 Sonnhammer,E.L. *et al.* (1997) Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 28, 405-420

156 Han,D.S. *et al.* (2004) PreSPI: design and implementation of protein-protein interaction prediction service system. *Genome Inform.* 15, 171-180

157 Martin,S. *et al.* (2005) Predicting protein-protein interactions using signature products. *Bioinformatics.* 21, 218-226

158 Pitre,S. *et al.* (2006) PIPE: a protein-protein interaction prediction engine based on the re-occurring short polypeptide sequences between known interacting protein pairs. *BMC. Bioinformatics.* 7, 365

159 Smith,T.F. and Waterman,M.S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.* 147, 195-197

160 Nash,R. *et al.* (2007) Expanded protein information at SGD: new pages and proteome browser. *Nucleic Acids Res.* 35, D468-D471

161 Thompson,J.D. *et al.* (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680

162 Schwarz,E.M. *et al.* (2006) WormBase: better software, richer content. *Nucleic Acids Res.* 34, D475-D478

163 Crosby,M.A. *et al.* (2007) FlyBase: genomes by the dozen. *Nucleic Acids Res.* 35, D486-D491

164 Krogan,N.J. *et al.* (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* 440, 637-643

165 von Mering C. *et al.* (2007) STRING 7--recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res.* 35, D358-D362

166 Altschul,S.F. *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402

167 Quevillon,E. *et al.* (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* 33, W116-W120

168 Garate,M. *et al.* (2008) NAD(P)H quinone oxidoreductase 1 inhibits the proteasomal degradation of the tumour suppressor p33(ING1b). *EMBO Rep.* 9, 576-581

169 Ikura,T. *et al.* (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102, 463-473

170 Nourani,A. *et al.* (2003) Opposite role of yeast ING family members in p53-dependent transcriptional activation. *J. Biol. Chem.* 278, 19171-19175

171 Dulic,V. *et al.* (1993) Altered regulation of G1 cyclins in senescent human diploid fibroblasts: accumulation of inactive cyclin E-Cdk2 and cyclin D1-Cdk2 complexes. *Proc. Natl. Acad. Sci. U. S. A* 90, 11034-11038

172 Loewith,R. *et al.* (2000) Three yeast proteins related to the human candidate tumor suppressor p33(ING1) are associated with histone acetyltransferase activities. *Mol. Cell Biol.* 20, 3807-3816

173 Needleman,S.B. and Wunsch,C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48, 443-453

174 Notredame,C. *et al.* (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205-217

175 Gandhi,T.K. *et al.* (2006) Analysis of the human protein interactome and comparison with yeast, worm and fly interaction datasets. *Nat. Genet.* 38, 285-293

176 Kelley,B.P. *et al.* (2004) PathBLAST: a tool for alignment of protein interaction networks. *Nucleic Acids Res.* 32, W83-W88

177 Mika,S. and Rost,B. (2006) Protein-protein interactions more conserved within species than across species. *PLoS. Comput. Biol.* 2, e79

178 Sharan,R. *et al.* (2005) Conserved patterns of protein interaction in multiple species. *Proc. Natl. Acad. Sci. U. S. A* 102, 1974-1979

179 Shoemaker,B.A. and Panchenko,A.R. (2007) Deciphering protein-protein interactions. Part II. Computational methods to predict protein and domain interaction partners. *PLoS. Comput. Biol.* 3, e43

180 Jansen,R. *et al.* (2003) A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* 302, 449-453

181 Ramani,A.K. and Marcotte,E.M. (2003) Exploiting the co-evolution of interacting proteins to discover interaction specificity. *J. Mol. Biol.* 327, 273-284

182 Pazos,F. and Valencia,A. (2002) In silico two-hybrid system for the selection of physically interacting protein pairs. *Proteins* 47, 219-227

183 Marcotte,C.J. and Marcotte,E.M. (2002) Predicting functional linkages from gene fusions with confidence. *Appl. Bioinformatics.* 1, 93-100

184 Espadaler,J. *et al.* (2005) Prediction of protein-protein interactions using distant conservation of sequence patterns and structure relationships. *Bioinformatics.* 21, 3360-3368

185 Mathivanan,S. *et al.* (2006) An evaluation of human protein-protein interaction data in the public domain. *BMC. Bioinformatics.* 7 Suppl 5, S19

186 Brown,K.R. and Jurisica,I. (2007) Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol.* 8, R95

187 O'Brien,K.P. *et al.* (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res.* 33, D476-D480

188 Wheeler,D.L. *et al.* (2007) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 35, D5-12

189 von Mering C. *et al.* (2007) STRING 7--recent developments in the integration and prediction of protein interactions. *Nucleic Acids Res.* 35, D358-D362

190 Bettinger,B.T. and Amberg,D.C. (2007) The MEK kinases MEKK4/Ssk2p facilitate complexity in the stress signaling responses of diverse systems. *J. Cell Biochem.* 101, 34-43

191 Gavin,A.C. *et al.* (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631-636

192 Rual,J.F. *et al.* (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437, 1173-1178

193 Stelzl,U. *et al.* (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122, 957-968

194 Rhodes,D.R. *et al.* (2005) Probabilistic model of the human protein-protein interaction network. *Nat. Biotechnol.* 23, 951-959

195 Lehner,B. and Fraser,A.G. (2004) A first-draft human protein-interaction map. *Genome Biol.* 5, R63

196 Lee,I. *et al.* (2007) An Improved, Bias-Reduced Probabilistic Functional Gene Network of Baker's Yeast, Saccharomyces cerevisiae. *PLoS. ONE.* 2, e988

197 Brown,K.R. and Jurisica,I. (2005) Online predicted human interaction database. *Bioinformatics.* 21, 2076-2082

198 Huang,T.W. *et al.* (2004) POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome. *Bioinformatics.* 20, 3273-3276

199 Kemmer,D. *et al.* (2005) Ulysses - an application for the projection of molecular interactions across species. *Genome Biol.* 6, R106

200 Liu,Y. *et al.* (2005) Inferring protein-protein interactions through high-throughput interaction data from diverse organisms. *Bioinformatics.* 21, 3279-3285

201 Wang,Z. *et al.* (1995) Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA.* 1, 21-35

202 Sato,N. *et al.* (1999) A novel presenilin-2 splice variant in human Alzheimer's disease brain tissue. *J. Neurochem.* 72, 2498-2505

203 Saito,A. *et al.* (2000) p24/ING1-ALT1 and p47/ING1-ALT2, distinct alternative transcripts of p33/ING1. *J. Hum. Genet.* 45, 177-181

204 Griffin,T.J. *et al.* (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. *Mol. Cell Proteomics.* 1, 323-333

205 Binda,O. *et al.* (2008) SIRT1 negatively regulates HDAC1-dependent transcriptional repression by the RBP1 family of proteins. *Oncogene* 27, 3384-3392

206 Munro,J. *et al.* (2004) Histone deacetylase inhibitors induce a senescence-like state in human cells by a p16-dependent mechanism that is independent of a mitotic clock. *Exp. Cell Res.* 295, 525-538

207 Place,R.F. *et al.* (2005) HDACs and the senescent phenotype of WI-38 cells. *BMC. Cell Biol.* 6, 37

208 Courtois,S. *et al.* (2002) DeltaN-p53, a natural isoform of p53 lacking the first transactivation domain, counteracts growth suppression by wild-type p53. *Oncogene* 21, 6722-6728

209 Syken,J. *et al.* (1999) TID1, a human homolog of the Drosophila tumor suppressor l(2)tid, encodes two mitochondrial modulators of apoptosis with opposing functions. *Proc. Natl. Acad. Sci. U. S. A* 96, 8499-8504

210 Shelton,D.N. *et al.* (1999) Microarray analysis of replicative senescence. *Curr. Biol.* 9, 939-945

211 Narita,M. *et al.* (2006) A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. *Cell* 126, 503-514

212 Goeman,F. *et al.* (2005) Growth inhibition by the tumor suppressor p33ING1 in immortalized and primary cells: involvement of two silencing domains and effect of Ras. *Mol. Cell Biol.* 25, 422-431

213 Abad,M. *et al.* (2007) Ing1 mediates p53 accumulation and chromatin modification in response to oncogenic stress. *J. Biol. Chem.*

214 Oyama,K. *et al.* (2007) AKT induces senescence in primary esophageal epithelial cells but is permissive for differentiation as revealed in organotypic culture. *Oncogene* 26, 2353-2364

215 Goll,J. and Uetz,P. (2006) The elusive yeast interactome. *Genome Biol.* 7, 223

216 Ozer,A. *et al.* (2005) The candidate tumor suppressor ING4 represses activation of the hypoxia inducible factor (HIF). *Proc. Natl. Acad. Sci. U. S. A* 102, 7481-7486

217 Kyriakis,J.M. and Avruch,J. (2001) Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev.* 81, 807-869

218 Miyake,Z. *et al.* (2007) Activation of MTK1/MEKK4 by GADD45 through induced N-C dissociation and dimerization-mediated trans autophosphorylation of the MTK1 kinase domain. *Mol. Cell Biol.* 27, 2765-2776