uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Rahin Farzadfar

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Biochemistry)

GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Identification of Artemis and PARP-1 Interacting Factors:
A Yeast Two Hybrid Screening Approach

TITRE DE LA THÈSE / TITLE OF THESIS

Robert Haché

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

Leonard Kleine

Alain Stintzi

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Identification of Artemis and PARP-1 interacting factors: A yeast two hybrid screening approach

By
Rahin Farzadfar

Thesis submitted to the Department of Biochemistry
In partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

University of Ottawa
Ottawa, Ontario, Canada
December 2009

Your file Votre référence
ISBN: 978-0-494-65503-0
Our file Notre référence
ISBN: 978-0-494-65503-0

ABSTRACT

The genomic integrity of the cell is under the constant threat of DNA damage from endogenous and exogenous sources. In response to DNA damage mammalian cells elicit a highly complex, yet coordinated series of cellular responses, which are triggered by the detection of DNA lesions and ultimately lead to cell survival (DNA repair) or apoptosis.

Artemis and PARP-1 are two proteins with significant implications in the various branches within the DNA damage response network. A series of yeast two hybrid screening (Y2H) experiments were conducted using both classical and interacting mating Y2H systems to find Artemis and PARP-1 interacting factors. The ATRX protein was identified as a putative interactor of Artemis, Caspase-7 and an "unkown" protein were identified as putative interactors of PARP-1 binding factors. In addition a rare case of false positives were identified and characterized. Furthermore, the strengths and limits of each Y2H system, as well as troubleshooting techniques developed to increase the efficiency of each system were thoroughly examined.

ACKNOWLEDGMENTS

There are a number of individuals I would like to thank whose support and guidance has made this work possible. First I would like to thank my supervisor Dr. Robert Haché for giving me the wonderful opportunity to be a part of his team. His teachings, patience and valuable guidance have made my studies possible. I would like to thank Dr. Sebastien Soubeyrand for his daily input, discussion, and direction throughout the course of this work. I am extremely grateful to have learned so much from Dr. Haché and Dr. Soueyrand's expertise.

I would like to also thank Louise Pope for her kind helping hand and her patience. Thank you to my friends and colleagues in the laboratory, in particular Kyna Caminiti, and Houssein Abdu-Salem for their support and for making each day brighter and more enjoyable. And last but not least, my father; for his example, enthusiasm, and encouragement throughout my studies.

Table Of Contents

LIST OF ABBREVIATIONS

3-AT.....3-amino-1,2,4-triazole
53BP1.....p53 binding protein 1
Ab.....Antibody
AD.....Activation Domain
Ade.....Adenine
AIF.....Apoptosis-Inducing Factor
A-T.....Ataxia telangiectasia
ATM.....Ataxia telangiectasia mutated
ATP.....Adenosine triphosphate
ATR.....Ataxia-telangiectasia and Rad3-related
ATR-X.....Alpha thalassemia/mental retardation syndrome, X-linked
BRCA1.....Breast cancer 1, early onset
BSA.....Bovine serum albumin
BLAST.....Basic Local Alignment Search Tool
cDNA.....Complementary DNA
Cdc25.....Cell division cycle 25
Cdk.....Cyclin-dependent kinases
Cfu.....Colony forming units
Chk-1 and -2.....Checkpoint kinases -1 and -2
Co-IP.....Co-immunoprecipitation
DBD.....DNA Binding Domain
DMSO.....Dimethyl sulfoxide
DNA.....Deoxyribonucleic Acid
dNTP.....Deoxynucleoside triphosphate
DSB.....Double-Strand Break
ECL.....Enhanced chemiluminescence
ECM.....Extracellular matrix
EDTA.....Ethylenediamine tetraacetic acid
E(z).....Protein Enancer Zeste E(z)
EZH2.....Enhancer of zeste homolog 2
SET.....Su(var)3-9, Enhancer-of-zeste, Trithorax
SSB.....Single-Strand Break
ds cDNA.....Double stranded cDNA
DNA-PK.....DNA-dependent protein kinase
DNA-PKcs.....Catalytic subunit of DNA-PK
Hsp90.....Heat shock protein 90
HEPES.....N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HP1.....heterochromatin protein 1
His.....Histidine
HB.....Human Bone Marrow
HRR.....Homologous recombination repair
ICAD/CAD.....Inhibitor of caspase-activated DNase/ caspase-activated DNase
Ig.....Immunoglobulin
IR.....Ionizing radiation

KAP-1.....Krüppel-associated box (KRAB)-associated protein 1
Leu.....Leucine
LiAc.....Lithium acetate
MCS.....Multiple Cloning Site
MEKK1.....MAP/ERK kinase kinase 1
MMLV.....Moloney Murine Leukemia Virus
mRNA.....Messenger RNA
NHEJ.....Non-homologous end joining
NLS.....Nuclear localization signal
ORF.....Open reading frame
PARP.....Poly(ADP-ribose) polymerase
PARP-1.....Poly(ADP-ribose) polymerase-1
PAR.....Poly(ADP-ribosyl)
PCR.....polymerase chain reaction
PIKK.....Phoshphoinositide-3-kinase-related protein kinase
PIAS1.....Protein inhibitor of activated STAT 1
ROS.....Reactive oxygen species
RSS.....Recombination signal sequences
RAG.....Recombination-Activating Genes
RNA.....Ribonucleic acid
RS-SCID.....Radiosensitive-SCID
RT.....Reverse Transcriptase
SCID.....Severe combined immunodeficiency disease
SV40.....Simian virus 40
SDS-PAGE.....Sodium dodecyl sulphate polyacrylamide gel-electrophoresis
SMART.....Switching Mechanism At 5' end of RNA Transcript
sscDNA.....single stranded cDNA
TA.....Transcription Activator
TCR.....T cell receptors
TdT.....Human terminal deoxynucleotidyltransferase
TEMED.....Tetramethylethylenediamine
TBS-T.....Tris-buffered saline with 0.1% tween-20 detergent
TRIS.....Tris(hydroxymethyl)aminomethane
Trp.....Tryptophan
Ura.....Uracil
UTR.....Untranslated region
UV.....Ultra violet
XRCC4.....X-ray repair cross complementing factor 4
XLF.....XRCC4-like factor
Y2H.....Yeast two hybrid
YPD.....Yeast extract peptone dextrose

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

DNA Damage Response and Repair Mechanisms

DNA Damage

It was once believed that it was the intrinsic stability of the DNA molecule which ensured genomic continuity into subsequent generations (1). Several decades of research have since altered this conceptual view, revealing the true dynamics of the DNA double helix within the cell. It is known today that the genomic integrity of the cell is under the constant threat of damage from endogenous and exogenous sources, and that the cell has evolved a remarkable capacity for DNA damage detection and repair (1,2). The DNA double helix continually endures various forms of damage, including base modifications, alterations of sugar-phosphate backbone, single-strand breaks (SSBs), and double-stand breaks (DSBs). The most lethal form of damage that the DNA molecule can endure is a DSB, where both strands of the DNA molecule are severed, leading to chromosomal breakage (3). DSBs can be introduced by an agent that is capable of severing the DNA double helix backbone. DSBs can be generated endogenously from products of cellular metabolism (e.g. reactive oxygen species or ROS), or by regulated cellular processes such as: V(D)J recombination, class-switch recombination, and meiosis (4). DSBs can also be generated exogenously from exposure to ionizing radiation (IR), x-rays, chemotherapeutic agents, and mutagenic chemicals (5). To properly protect the integrity of its genome, the cell must promptly detect DNA damage and elicit appropriate and coordinated cellular responses, which will ultimately lead to cell survival (DNA repair) or apoptosis. DSBs that are left un-repaired may lead to genome instability, chromosomal aberrations, and cell death (3, 6). Alternatively, inadequately repaired DSBs can lead to carcinogenic mutations, chromosomal rearrangements, and defects in the immune system of higher vertebrates (6).

Dynamics of the DNA Damage Response

The cellular response to DNA damage is a highly elaborate and coordinated network, triggered by the detection of DNA lesions (changes in chromatin structure). The cell propagates this initial signal into series of appropriate cellular responses, which collectively have been termed the DNA damage response. A highly branched biochemical phosphorylation cascade is triggered: signals are generated by sensors, amplified by transducers, and passed on to effectors which recruit and activate cell cycle checkpoint and DNA repair machinery (**Figure 1**) (7, 8). Among the initial sensors of DNA damage is the Mre11/Rad50/Nbs1 (MRN) complex (9). This highly conserved protein complex rapidly localizes to sites of damage upon generation of DNA lesions. Possessing both nuclease activity and DNA binding capability, the MRN complex is able to process DNA termini and stabilize DNA ends by bridging severed DNA ends together prior to repair (9, 10). The DNA damage signal is then passed down via activation of signal transducers such as members of the phoshphoinositide-3-kinase-related protein kinase (PIKK) family, including ATM (ataxia-telangiectasia-mutated), ATR (ataxia-telangiectasia and Rad3-related), and DNA-PKcs (catalytic subunit of DNA-dependent protein kinase) (11). ATR activation is mainly associated with SSBs and stalled activation forks, while ATM and DNA-PKcs are activated mainly in response to DSBs (11, 12). Upon activation, these kinases phosphorylate several shared substrates and initiate overlapping but not identical damage responses. ATM is the protein product of the gene mutated in the disorder ataxia-telangiectasia (AT), which is characterized by neuronal degeneration, immunodeficiency, chromosomal instability, and cancer predisposition. A nuclear protein with serine/threonine kinase activity, ATM is activated in response to DSBs and is one of the critical proteins involved in the maintenance

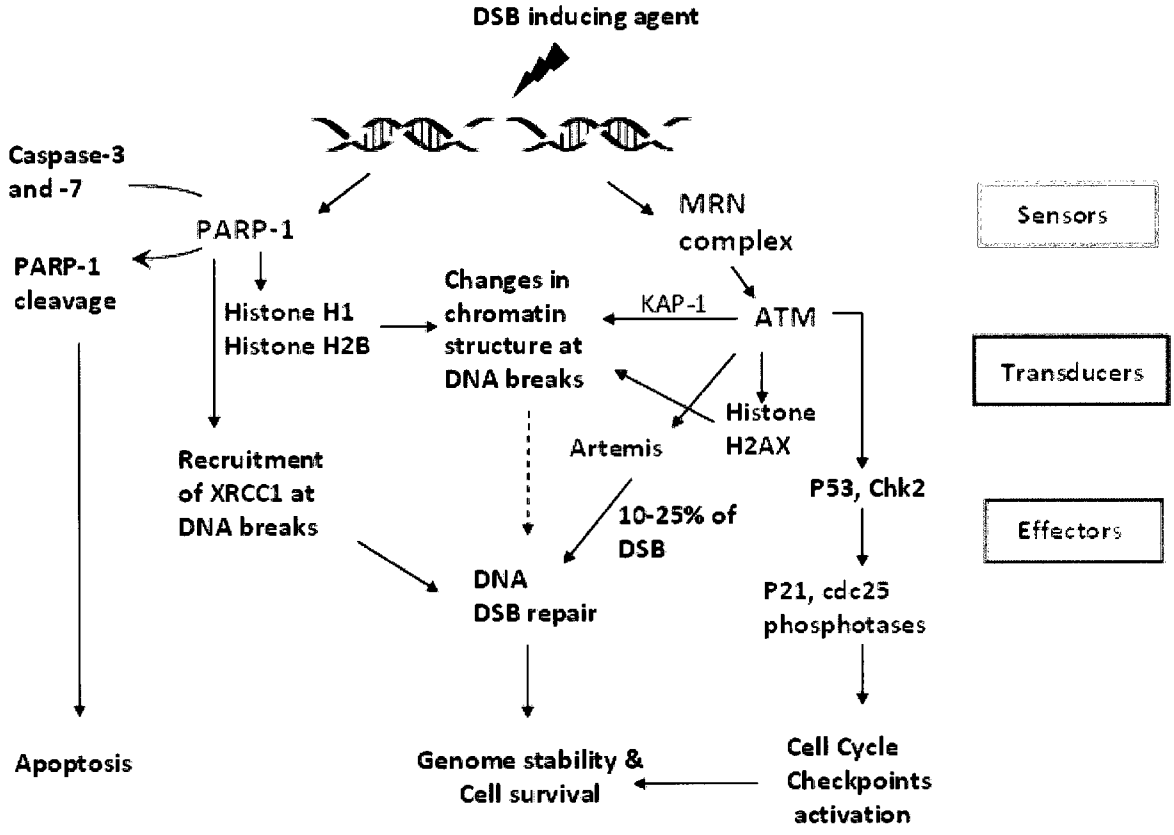

DSB inducing agent
Caspase-3 and -7
PARP-1
PARP-1 cleavage
Histone H1
Histone H2B
Changes in chromatin structure at DNA breaks
KAP-1
ATM
MRN complex
Sensors
Transducers
Effectors
Histone H2AX
Artemis
Recruitment of XRCC1 at DNA breaks
10-25% of DSB
P53, Chk2
DNA DSB repair
P21, cdc25 phosphotases
Apoptosis
Genome stability & Cell survival
Cell Cycle Checkpoints activation

Figure 1. A simplified representation of the DNA damage response network. A biochemical cascade is triggered with the detection of DNA lesions; signals are generated by sensors, amplified by ATM and PARPs, which promote chromatin re-modeling, and recruit and activate DNA repair and cell cycle checkpoint activation machinery (2, 8, 9).

of genomic stability (13). Recent studies have led researchers to believe that upon the generation of DSBs, the MRN complex recruits inactive ATM (dimer or multimer) to the site of the break (14). MRN triggers the auto- or trans- phosphorylation of ATM leading to the dissociation of inactive ATM complexes into catalytically active ATM monomers (14, 15, 16). ATM, a key signal transducer in this network, mediates pleotropic responses in various branches of the DNA damage response network, including DNA DSB repair, control of telomere length, activation of cell cycle checkpoints, and apoptosis (13). The DNA damage signal is amplified onto downstream effectors such as Artemis (end-processing nuclease), histone variant H2AX, p53, and checkpoint kinase (chk) 2 (7, 17, 18). In recent years, numerous studies have emerged proposing functional interactions between the chromatin 'remodeling' process and ATM/MRN complex, further highlighting the critical importance between the coordination of chromatin structural changes with DNA damage response (18, 19, 20, 21). The structure of highly condensed chromatin changes dynamically throughout different stages of cell cycle and also in response to DNA damage (21). Chromatin 'remodeling' occurs with post-translational modification of histones and recruitment of proteins which promote relaxation of the chromatin fiber allowing access of DNA repair machinery (18). A direct connection between chromatin alteration and ATM was identified when the transcriptional corepressor Krüppel-associated box (KRAB)-associated protein (KAP)-1 was recently identified as a substrate of ATM (21, 22). KAP-1 is a nuclear protein that is involved in heterochromatin formation. Upon phosphorylation by ATM (KAP-1's serine-824 residue), KAP-1 leads to transient relaxation of the chromatin fiber (20, 21). It is possible that in regions where nucleosome flexibility is constrained by heterochromatic factors such as KAP-1, repair factors are unable to access or manipulate DNA ends (21). ATM signaling is required to phosphorylate KAP-1, which transiently loosens the chromatin

structure, enabling repair proteins to adequately access and repair the DSBs. Moreover, histone H2AX, a variant of the H2A protein family, is phosphorylated by ATM (and DNA-PKcs) following formation of DSBs to generate γ-H2AX (23). It has been suggested that formation of γH2AX may serve as an epigenetic marker on DNA, inducing the relaxation of chromatin, and potentially recruiting and localizing DNA repair proteins (17, 24) and checkpoint protein complexes to the site of damage (25, 26). Other findings however suggest that γH2AX formation may solely be the consequence of free DNA ends, and indeed may not play a role in signal transduction or recruitment of DNA repair factors in response to DNA damage (23, 27, 28, 29, 30).

Eukaryotic Cell Cycle Regulation in response to DNA damage

A class of highly regulated protein kinases called cyclin-dependent kinases (Cdks) are the key molecules that coordinate progression throughout the eukaryotic cell cycle (31). Cdks are activated when bound to cyclin subunits, and phosphorylate various target proteins to mediate entry into the subsequent phase of the cell cycle. Each cell cycle phase is defined and regulated by a select set of cyclin-Cdk complexes; consequently, the cell cycle is driven forward by the periodic activation and inhibition of cyclin-Cdk complexes in a coordinated manner to ensure unidirectional signaling (31, 32). Activation of cyclin-cdk complexes is often achieved via dephosphorylation by members of the Cdc25 (cell division cycle 25) protein phosphatase family (33). Conversely, the inhibition of cyclin-cdk complexes is achieved by their phosphorylation or by the binding of inhibitory proteins (ie. p21), which delay Cdk activation and slow cell cycle progression (31). Additionally, the cell must coordinate its progression throughout the cell cycle with the DNA damage response and repair pathways to ensure the entry of accurate copies of the genome onto subsequent

generations (16). Cell cycle checkpoints are surveillance mechanisms evolved by eukaryotic cells to ensure that processes at each phase of the cell cycle have been accurately completed. When the cell senses DNA damage, the checkpoints are activated, and cell cycle arrests until the damage is repaired or the cell undergoes apoptosis. Damage response checkpoints have been identified at the G_1/S, intra S, and G_2/M levels (**Figure 2**) (34). The G_1 checkpoint is activated when DNA damage is detected in cells traversing G_1 (16). Upon activation, the signal is amplified by ATM via phosphorylation of Chk1/Chk2 (two structurally-unrelated kinases with various functions in checkpoint activation), which then phosphorylate two critical effectors: the p53 transcription factor and Cdc25C phosphatases (16). This leads to the transcriptional activation of the cyclinE/Cdk2 inhibitor, p21, which ultimately prevents entry from G1 into S phase. The S phase checkpoint is less understood, it has been proposed that its activation results merely in a transient, reversible delay in cell cycle progression by slowing down DNA replication (35). It is agreed upon however, that ATM-ATR-Chk1-mediated protein degradation of Cdc25A protein phosphatase and cdk2 is the main mechanism responsible for S-phase checkpoint activation. The G_2 checkpoint prevents cells from initiating mitosis when harboring DNA damage during G_2, or if they progress into G_2 with some un-repaired damage accumulated in S or G_1 phase. G_2/M arrest is a result of a combination of mechanisms involving post-transcriptional modification, via Chk2 and Chk1. The major downstream target of the G_2 checkpoint is cyclinB/Cdk1, whose activity promotes mitosis (16). Upon detection of DNA damage, the activation of cyclinB/Cdk1 is prevented through ATM/ATR- and Chk1/Chk2 phosphorylation and/or inhibition of the Cdc25 (responsible for activation of Cdk1 at the G2/M border under standard circumstances without DNA damage), thus preventing progression into mitosis until damage is repaired (16).

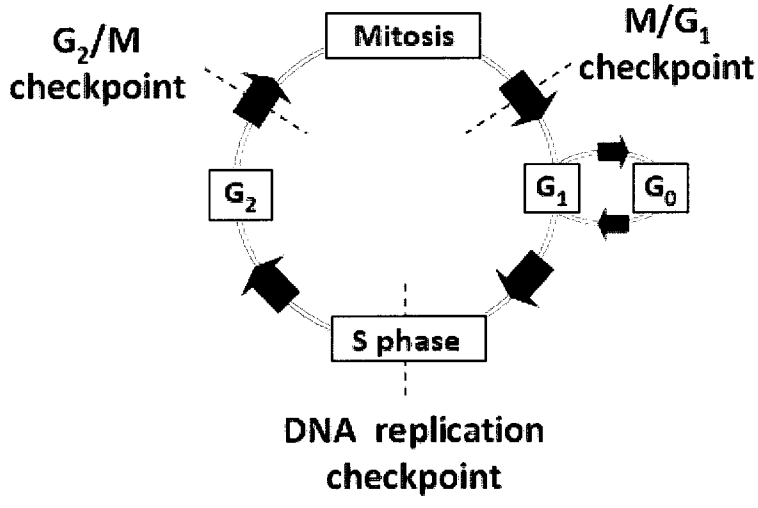

G_2/M
checkpoint
Mitosis
M/G_1
checkpoint
G_2
G_1
G_0
S phase
DNA replication
checkpoint

Figure 2. Schematic of the cell cycle and key cell cycle checkpoints. Cells activate checkpoints upon detection of DNA damage to provide time for the repair of DNA damage. G_1/S and intra-S (DNA replication checkpoint) are activated to allow repair of DNA damage prior to DNA replication, and G_2/M checkpoint allows repair of DNA damage prior to entry into mitosis [Adapted from (35)].

Poly(ADP-ribosyl)ation and PARP-1

Post-translational modifications of proteins are essential to virtually all cellular processes (36). The poly(ADP-ribosyl)ation of nuclear proteins in response to DNA damage in eukaryotic cells is an example of such post translational modification (36, 37, 38, 39). The poly(ADP-ribose) polymerase (PARP) family of proteins catalyze this modification. Polymers of ADP-ribose units (poly(ADP-ribose) or PAR) are synthesized from donor NAD^+ molecules, and are attached onto target nuclear proteins (such as core histones, p53, and PARP-1, itself) as linear or branched polymers (38, 39, 40). PARPs and poly(ADP-ribosyl)ation are proposed to be involved in the regulation of various cellular processes such as DNA repair, cell death, genomic stability (38). Poly(ADP-ribose) polymerase-1 (PARP-1) is the founding member of PARP family, which consists of 18 known members (in humans) and share a conserved catalytic domain (40). PARP-1 and PARP-2 are thus far the only two enzymes whose catalytic activity is stimulated in response to DNA damage (40). Poly(ADP-ribosyl)ation is initiated on glutamate residues of target proteins, and subsequent polymerization of ADP-ribose units extends from the site of initiation (**Figure 3**) (40). The result is the addition of a long negatively charged polymer (~50 to 200 units long) (38, 40). The consequences of this post-translational modification are profound; each residue in PAR contains an adenine moiety (capable of base stacking and hydrogen bonding), as well as two phosphate groups that harbor negative charges. Therefore, poly(ADP-ribosyl)ation substantially changes the biophysical and electrostatic properties of a protein, and thereby can alter DNA-binding functions, protein-protein interaction properties, and the cellular localization of target proteins 40).

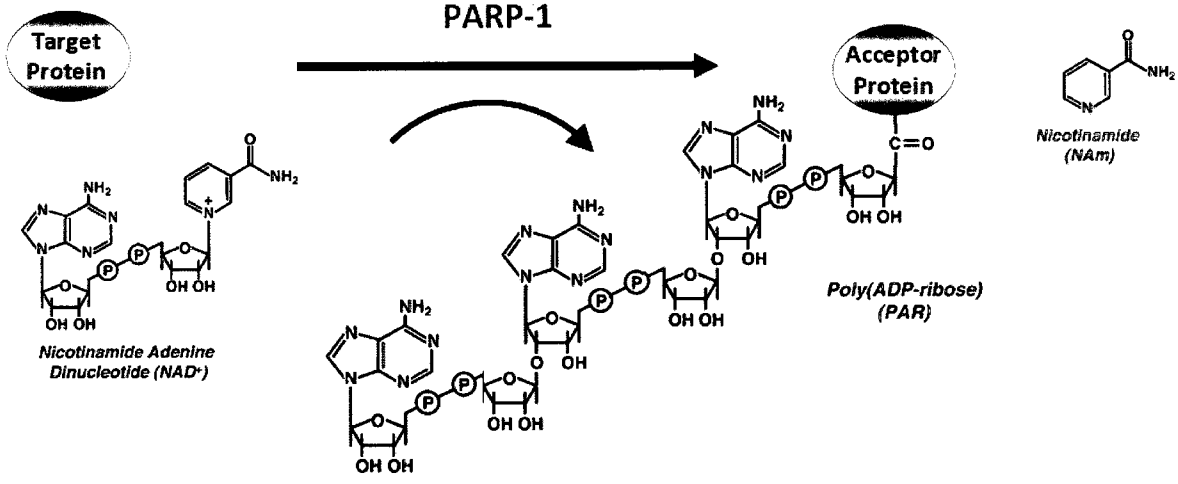

PARP-1
Target Protein
Acceptor Protein
Nicotinamide Adenine Dinucleotide (NAD+)
Poly(ADP-ribose) (PAR)
Nicotinamide (NAm)

Figure 3. PARP-1 catalyzes the synthesis of poly(ADP-ribose) units. PARP-1 catalyzes the synthesis of poly(ADP-ribose) units (poly(ADP-ribosyl)ation) onto acceptor protein from donor NAD^+ molecules [Adapted from (40)].

PARP-1 in the DNA Damage Response

The involvement and significance of PARP-1 in the DNA Damage response network became evident with the initial observation that immediately following DNA damage, PARP-1 binds to damaged DNA (40). PARP-1 is the most abundant PARP enzyme (~10^6 copies/cell) and accounts for at least 85% of PARP activity in living cells (absent in yeast); its enzymatic activity increases dramatically (up to 500 fold) upon the induction of DNA damage (40). It has been proposed that the initial binding of PARP-1 to damaged DNA may serve as an epigenetic marker that stabilizes severed DNA ends, and/or a surveillance protein that translates the occurrences of DNA strand breaks into an amplified nuclear signal (41). The primary *in vivo* target of poly(ADP-ribosyl)ation in response to DNA damage is PARP-1 itself. Upon binding to DNA at sites of damage, PARP-1 is activated and undergoes automodification leading to the accumulation of negative charge, and its subsequent dissociation from DNA. The PARylation of PARP-1, however, is short lived, as this modification is degraded by Poly(ADP-ribose) glycohydrolase (PARG) minutes after PAR synthesis. The auto-modification of PARP-1 alters its affinity for DNA and proteins with which it interacts. The result is PARP-1 shuttling on and then off of DNA, and likely serves to recruit DNA repair factors such as XRCC4 and DNA ligase III to the site of damage (39, 40). In addition, PARP-1 regulates chromatin architecture via poly(ADP-ribosyl)ation of core histones H1 and H2B, which destabilizes nucleosomes and leads to the relaxation of the chromatin fiber, thus increasing access of DNA repair machinery to damaged DNA termini (39, 40, 41). PARP-1 and PAR have been shown to associate with various DNA repair factors and checkpoint proteins (ie. p53 and p21), which likely contribute to the DNA damage response (42, 43).

PARP-1 in Cell Death: To repair or Die?

The involvement of PARP-1 in cell death pathways was initially proposed in 1985 with the observation that excessive DNA damage results in a drastic increase in the specific activity of PARP-1, followed by drastic alterations in the levels of intracellular intermediates (NAD^+ and ATP) (44, 45). It was proposed that massive PARP-1 activation was responsible for NAD^+ and ATP depletion, which would subsequently lead to low energy pools, signaling the activation of a "suicide" pathway (44, 45). Since this observation, various experiments were conducted using PARP inhibitors, $PARP^{-/-}$ cells, and PARP deficient mice to further address the role of PARP in cell death pathways. In short, findings suggested opposing roles for the enzyme: while certain studies indicated a role for PARP as a mediator in cell death, others indicated a survival role where PARP protected the cell from damage (39). It is now believed that PARP-1 plays a dual role in the DNA damage response: one as a survival factor and another in promoting cell death. Although the molecular determinants that result in the switch between cell cycle arrest/DNA repair and cell death are not yet fully understood, it is likely that the physiological conditions of the cell (ie. metabolic status) and the severity of DNA damage that determine cell viability. Under physiological conditions and limited DNA damage, PARP plays a role as a survival factor, promoting activation of DNA repair pathways. On the other hand, under "non-physiological" conditions, PARP over-activation by excessive damage may promote activation of cell death pathways (46, 47, 48, 49). The apoptotic pathway is an intra-molecular signal cascade that is induced by external stimuli (such as appearance and disappearance of signals from surrounding cells), and is controlled by a set of genes that encode death receptors, adaptors, and proteolytic enzymes (49). PAR is synthesized from cytocolic NAD+, triggering the release of the flavoprotein Apoptosis-

Inducing Factor (AIF) from the mitochondrion in to the nucleus (49). AIF is a powerful stimulant of apoptosis, which induces DNA fragmentation and chromatin condensation leading to caspase-independent cell death. PARP-1 is then cleaved by either one of two pro-apoptotic cysteine proteases: caspase-3 or capase-7 (39). This cleavage leads to the subsequent inactivation of PARP-1 enzymatic activity in response to DNA damage, which may serve to prevent additional PARP-1 activation and any further depletion of ATP levels, thus preventing the activation of necrosis. The PAR polymer therefore, is a key player in both the life and death of the cell, and PARP-1 activity is likely an indication of DNA damage severity to determine the fate of the cell: to repair or die? (40).

DNA Repair Pathways

All eukaryotic cells, from yeast to man, possess evolutionarily conserved mechanisms to counteract the potentially deleterious effects of DNA damage (50). Genomic instability arises from various forms of damage and can be described from various different angles; here the repair pathways that are activated in response to DNA DSBs will be reviewed. Eukaryotic cells repair DSBs primarily by two main DNA repair pathways: non-homologous end joining (NHEJ) and homologous recombination repair (HRR). DNA lesions associated with replication fork collapse are rapidly repaired by HRR by synthesizing DNA using an identical sequence as a template, such as a sister chromatid. Since the two sisters are localized within close proximity, the two sequences are aligned and a crossover between the aligned DNA strands occurs, resulting in the breaking and repairing of DNA to produce an exchange of material between the strands (50). The NHEJ pathway, on the other hand, repairs the DSB by directly ligating the DNA ends without the need for a homologous template to guide repair. The contribution of HRR and NHEJ to DNA repair in eukaryotic

cells is a highly regulated process that shifts during the various stages of the cell cycle and is also contingent on the nature of the break (51). While the HRR pathway can only be activated during the S/G_2 phases of the cell cycle when a sister chromatid (or homologous chromosome) is available to serve as a template, NHEJ is functional in all phases of the cell cycle (52). Furthermore, the DSBs that are produced by replication fork collapse are one-ended (have no second end to rejoin with) and are therefore repaired primarily (and potentially exclusively) by HRR (4). In the case of an "accidental" DSB that arises from ionizing radiation or enzymatic cleavage in a packed chromatin structure, NHEJ is the predominant repair pathway in higher eukaryotes. A DSB of such nature would not be able to interact with an intact homologous sequence (or even sister-chromatid), because the highly compact structure of the intact chromosome significantly separates the two sister chromatids (53). Extensive chromosome condensation in higher eukaryotic cells makes homology search difficult in the G2 and G_1 phase; therefore, DSBs in higher eukaryotes are predominately repaired by NHEJ, while HRR repair dominates in yeast (which have a less compartmentalized genome). Furthermore, it is possible that higher eukaryotes have evolved a smaller reliance on HRR due to the fact that a large number of sequence repeats in mammalian cells (compared to yeast) could potentially lead to undesirable inter-chromosome interactions though HRR (54).

Homologous Recombination Repair

The genetic framework that was established for studying HRR initially came from identifying relevant genes whose mutations led to recombinational defects or whose gene products possessed the ability to rescue mutants whose phenotypes are suggestive of such defects (54). HRR operates in two distinct pathways that have been identified thus far: a

RAD51-dependent pathway (DSB repair) and a RAD51-independent pathway (SSB repair) (54). The RAD51-dependent pathway provides "error-free" removal of damage present in replicated DNA, and thus functions in a highly regulated matter in coordination with the S and G_2 checkpoint machinery before cell division occurs (54). The majority of RAD51-dependent HRR involves the re-synthesis of a small amount of homologous sequence (gene conversion) from the undamaged homologous sequence (**Figure 4**). DNA ends are processed to create single-strand overhangs, which invade base-paired strands of a homologous DNA molecule in an ATP-dependent reaction mediated by the MRN complex, Rad51, and Rad52, followed by the formation of a cross-stranded intermediate by the damaged and undamaged strands (Holliday junction). The Holliday junction is then often resolved by excision at sites of intersection or in the outer strands, and the DSB is repaired using template guided DNA repair (54).

Non-homologous end joining

In the last decade, countless studies have been carried out to clarify our understanding of NHEJ in DSB repair. Several proteins have been identified as essential factors in NHEJ (Ku70/Ku80, DNA-PKcs, XRCC4, DNA ligase IV, Cernunnos/XLF and Artemis) and progressively, a clearer picture of this pathway has come into focus (5). NHEJ is potentially a less accurate form of DSB repair, where the broken termini of the DNA molecules must be processed to form compatible ends that are suitable for ligation (5). Under certain conditions, processing of non-compatible ends may lead to loss of nucleotides and potentially loss of genome integrity (5). The NHEJ pathway is initiated by the binding of the Ku70/80 heterodimer in a sequence specific manner to both termini of the severed DNA molecule (5). The Ku70/80 heterodimer is able to translocate from the DNA end in an

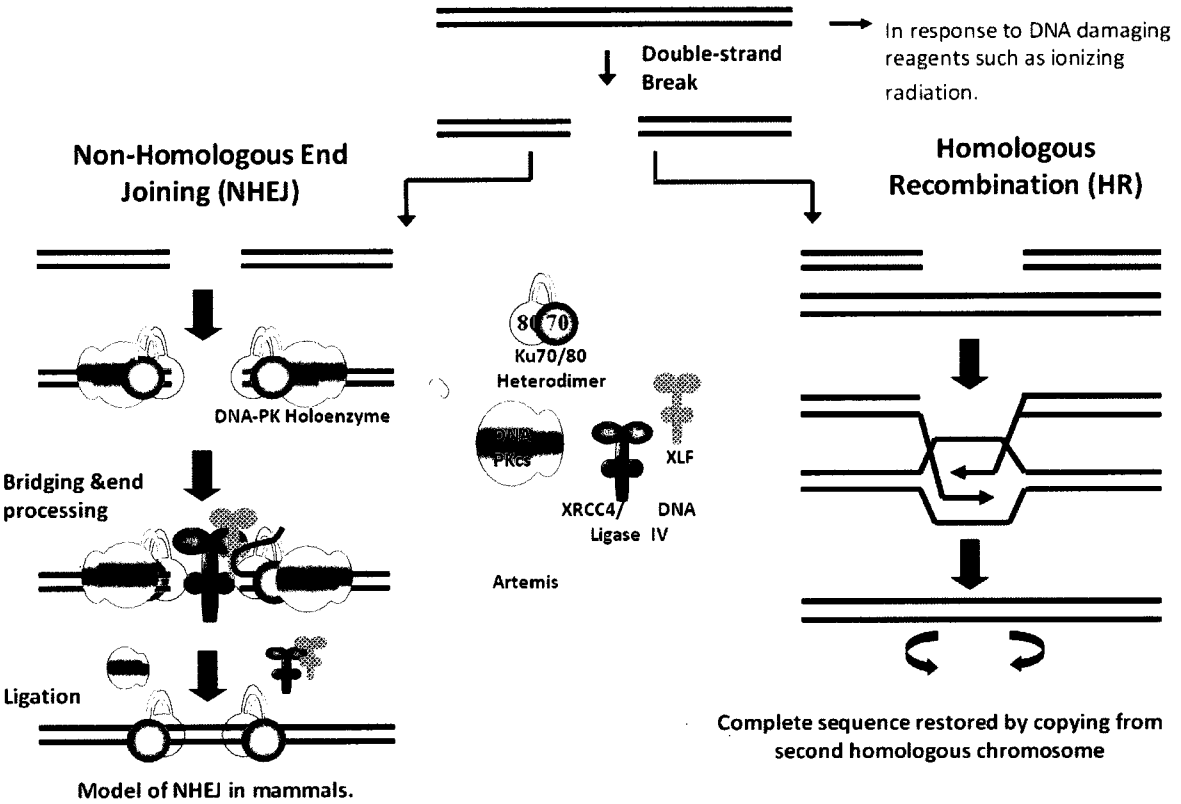

In response to DNA damaging reagents such as ionizing radiation.
Double-strand Break
Non-Homologous End Joining (NHEJ)
Homologous Recombination (HR)
Ku70/80 Heterodimer
DNA-PK Holoenzyme
PKcs
XLF
XRCC4/ Ligase IV
DNA
Artemis
Bridging &end processing
Ligation
Model of NHEJ in mammals.
Complete sequence restored by copying from second homologous chromosome

Figure 4. Depiction of NHEJ and HRR. NHEJ and HRR are two mechanisms that repair DSBs. HRR requires the presence of an identical or a nearly identical sequence to be used as a template for repair of the break. NHEJ involves the direct ligation of both broken DNA ends which are captured by the DNA-PK complex, a molecular bridge is formed and non-compatible termini are processed to form ligatable DNA ends, followed by repair of the break by the ligase IV/XRCC4 complex [Adapted from (5)].

ATP-independent manner (Figure 4) (51). The binding of Ku70/80 is thought to recruit DNA-PKcs to the DSB to form the "DNA-PK holoenzyme" complex (consisting of two DNA ends, two Ku70/80and two DNA-PKcs molecules) (5). DNA-PKcs is an extremely large complex (465-kDa) with weak DNA binding capability and serine/threonine kinase activity. The Ku 70/80 heterodimer directs DNA-PKcs to DNA, stabilizing its association with DNA, and ensuring efficient activation (5, 55). Evidence shows that the DNA-PK complex on each DNA end serves as a structural scaffold that tethers the two broken DNA ends together and serves as a docking site to which other repair factors are recruited. Essential for efficient NHEJ repair is the autophosphorylation activity of DNA-PKcs (5, 55). Upon phosphorylation, DNA-PKcs induces a conformational change at the synaptic complex that likely alters the accessibility of processing enzymes and ligases to the site of the break (5). Furthermore, the DNA-PK holoenzyme complex has been shown to phosphorylate Ku70/Ku80, XRCC4, Artemis, in addition to DNA-PKcs itself (56, 57). Although it has been well established that the activity of DNA-PKcs is essential for efficient DSB repair, the precise mechanism of its function and regulation remain unclear (5). Once the DNA ends have been bridged together, non-compatible DNA overhangs (present in the majority of DSBs) must be processed or removed prior to ligation (blunt end termini would not require processing and would simply be ligated together). Several players in this step are recruited to the site of damage to process DNA termini. These enzymes either remove or fill-in the single stranded overhangs, add 5'-phosphate groups, or remove 3'-phosphate or 3'-phosphoglycolate groups from DNA termini prior to ligation (5). It is also possible that the DNA contiguous to the DSB may endure additional forms of DNA damage, producing what is referred to as 'complex' or 'clustered' lesions (58, 59). It is possible that Artemis plays a role in the repair of such complex lesions, which will be discussed further detail below.

Several enzymes capable of synthesizing complementary strands of DNA have been proposed to be involved in filling in single stranded overhangs, including human terminal deoxynucleotidyltransferase (TdT) (5). To date, the only enzyme that has been demonstrated capable of restricting single stranded overhangs during NHEJ is the endonuclease Artemis. Intrinsically, Artemis exhibits 5′→3′ exonuclease activity (60). However, it has been proposed that upon association with and/or phosphorylation by DNA-PKcs, Artemis adopts endonuclease activity which is necessary for processing of a subset of DNA termini during repair of radiation-induced DSBs. Artemis is also capable of removing 5'-phosphoglycolate groups from DNA termini *in vitro* (60). Although it is unclear at which stage of the NHEJ pathway Artemis is initially involved, it has been suggested that its association with DNA-PKcs may also play a role in the early recruitment of the kinase to sites of DNA damage (61). Numerous studies, including those conducted within our own laboratory, have demonstrated that Artemis is phosphorylated by DNA-PKcs following exposure of cells to DSB-inducing agents (61, 62, 63), and that Artemis and DNA-PKcs form a stable complex *in vitro* (61, 64). In an elegant study conducted by Goodarzi *et al*, the endonuclease activity of Artemis was assayed, showing that the DNA-PK complex and ATP were sufficient to activate the endonuclease activity of Artemis at physiological salt concentrations (100mM NaCl) *in vitro* (61). Prior to ligation of the DNA ends, additional processing takes place if necessary. For example, addition of 5' phosphate groups is catalyzed by the mammalian polynucleotide kinase (PNK), the removal of 3' phosphoglycolates is catalyzed by endonuclease APE1, the tyrosyl-DNA phosphodiesterase, TDP1, polynucleotide kinase PNK, and Artemis (5). Once the two DNA ends have been tethered together and non-compatible termini have been processed into 5' phosphorylated ligatable ends, ligation is mediated by the ligase IV/XRCC4 complex. A novel NHEJ

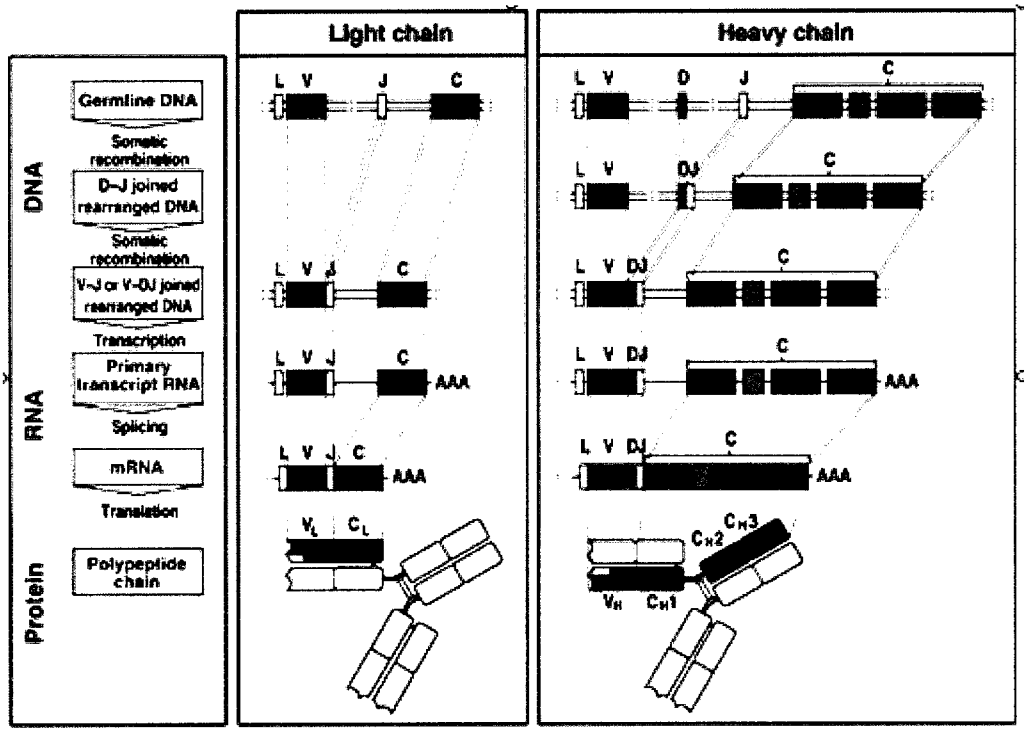
factor, XLF/Cernunnos factor (XRCC4-like factor), was recently discovered (65). The XLF/Cernunnos factor interacts with the ligase IV/XRCC4 complex, suggesting that the ligation reaction may be enhanced or modulated by the presence of the XLF/Cernunnos protein (5).

The V(D)J recombination connection: Immunology meets DNA repair

DSBs are not only generated in response to mutagenic agents, they are also continually generated and repaired (via NHEJ) in the developing B- and T- lymphocytes of the vertebrate immune system (66). V(D)J recombination is a site specific recombination process that rearranges and assembles the antigen-specific gene segments of B and T lymphocytes to generate a vast repertoire of diversity during the development of the vertebrate immune system (66, 67, 70). The immune system detects a wide variety of pathogens through the antigen binding receptors of B and T lymphocytes: immunoglobulins (Igs) and T-cell receptors (TCRs), respectively. Since each receptor chain variant cannot be encoded in the genome, V(D)J recombination generates highly variable proteins by recombining and assembling scattered gene segments to form a complete Variable (V) region sequence (70). Within the light chain of an immunoglobulin molecule, the V region is encoded by two separate gene segments. The first segment is termed the V region and encodes the first 95-101 amino acids of the light chain, while the second encodes the remainder of the V region (up to 13 amino acids) and is termed the joining (J) region (**Figure 5A**) (70). The heavy chain of an immunoglobulin is encoded by three gene segments: V, J and also diversity (D) segments. The rearrangement of gene segments is initiated by the lymphoid-specific factors *Rag1* and *Rag2* (Recombination-Activating Genes). *Rag1* and

Rag2 recognize conserved non-coding DNA sequences called recombination signal sequences (RSS) found adjacent to all V, D, and J gene segments (67). These RSSs consist of a conserved block of seven nucleotides, followed by a non-conserved spacer, either 12 nucleotides or 23 nucleotides long, called 12-RSS or 23-RSS respectively, followed by a second conserved sequence of nine nucleotides (68). A fundamental rule in the regulation of V(D)J recombination is the 12/23 rule, which ensures DNA rearrangements take place at the correct locations relative to the V,D or J gene segment coding regions (70). The gene segments recombine according to the 12/23 rule, such that an RSS with a 12 spacer is joined only to one flanked by a 23 spacer (**Figure 5B**) (68). *Rag1* and *Rag2* introduce a SSB at the site of the RSS flanking gene segments, which are brought together by a complex of enzymes that recognize the spacers and enforce the 12/23 rule (67, 68). Via nucleophilic attack of the complementary strand of DNA, the SSB ultimately results in the formation of covalently sealed hairpin structure, while the intervening DNA is looped out. The DNA hairpin is resolved, and the resulting endogenously generated DSBs are repaired via NHEJ forming the signaling end (67). The V and the J segments remain on the chromosome and form the covalently sealed (hairpin) coding ends (70). Coding joints must be joined in order to form the mature Ig or TCR gene, and thus their characteristic hairpin structure must be cleaved before end joining can proceed (4). The opening (processing) of the coding end hairpin structures is mediated by the essential endonucleolytic activity of Artemis. The V(D)J recombination process is concluded with the joining of the processed coding ends and unprocessed signal joint by the NHEJ machinery (69). This process is outlined in **Figure 6**. The critical roles of the Rag proteins and the NHEJ machinery in V(D)J recombination have been illustrated via gene targeting experiments in mice. The targeted selective inactivation of these genes leads to either embryonic lethality or severe combined immunodeficiency

A


Light chain
Heavy chain
DNA
RNA
Protein
Germline DNA
Somatic recombination
D–J joined rearranged DNA
Somatic recombination
V–J or V–DJ joined rearranged DNA
Transcription
Primary transcript RNA
Splicing
mRNA
Translation
Polypeptide chain
L V J C
L V D J C
L V DJ C
AAA

B

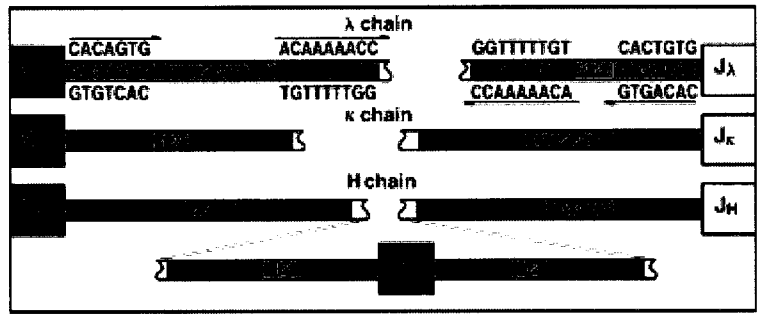

λ chain
CACAGTG
ACAAAAACC
GGTTTTTGT
CACTGTG
GTGTCAC
TGTTTTTGG
CCAAAAACA
GTGACAC
Jλ
κ chain
Jκ
H chain
JH

Figure 5. Representation of the recombination process that generates a complete V region from several separate gene segments of immunoglobulin heavy or light chain. **(A)** Light chain V regions are constructed by joining V and J gene segments from genomic DNA. The process of recombination that generates a heavy chain occurs in two stages. First a D segment is joined to a J gene segment. Then a V segment rearranges to a DJ to generate a complete V-region sequence adjacent to the constant (C) region gene of the immunoglobulin **(B)** 12-RSS or 23-RSS flank the gene segments encoding the V regions of heavy (H) and light (κ and γ) immunoglobulin chains. Joining between the V, (D) and J segments involves the 12/23 rule (70).

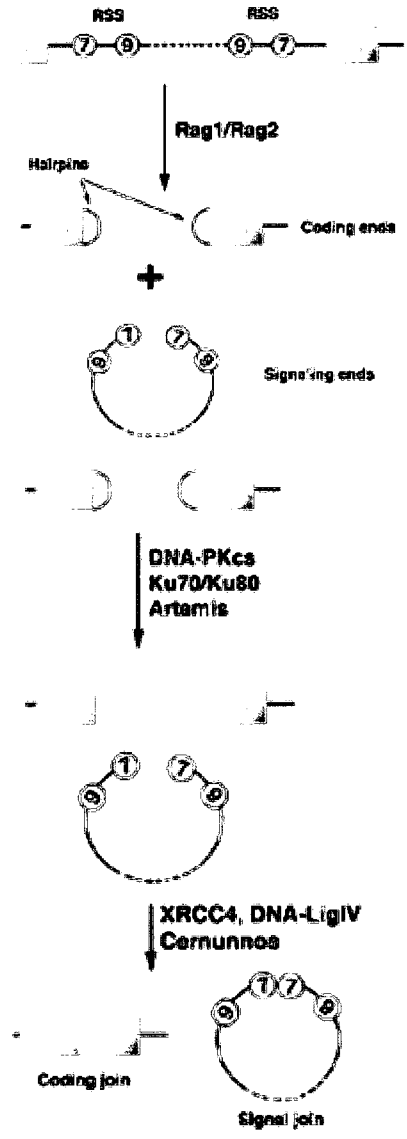

RSS
RSS
Rag1/Rag2
Hairpins
Coding ends
Signaling ends
DNA-PKcs
Ku70/Ku80
Artemis
XRCC4, DNA-LigIV
Cernunnos
Coding join
Signal join

Figure 6. Simplified representation of V(D)J recombination at the heavy chain locus of an immunoglobulin. During the initial step, *Rag1* and *Rag2* recognize and introduce a DSB at the border of RSSs creating hairpin-sealed coding ends and blunt signal ends. The NHEJ machinery is recruited to the site: Artemis is phosphorylated by the Ku/DNA-PKcs complex, and resolves the hairpins through its endonuclease activity. The XRCC4/Cernunnos/DNA-LigaseIV complex seals the coding and signal joints [Adapted from (67)].

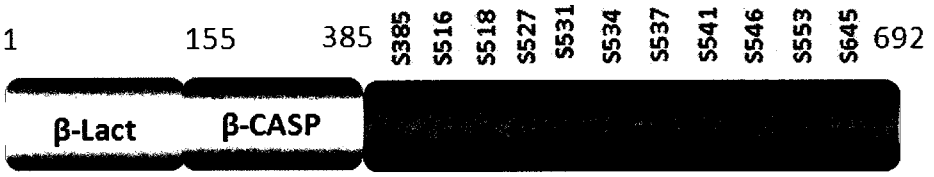
(SCID) disease (71, 72, 73, 74, 75). SCID is a primary immune deficiency that is marked by a defect in the development of the B and T lymphocyte systems, which compromises the adaptive immune system, and therefore leaves the host susceptible to infectious pathogens. There are several forms of SCID, including RS-SCID (radiosensitive severe combined immunodeficiency), and Athabascan SCID (A-SCID or SCIDA) to name a few (69). Approximately 20% of human SCID patients harbour a B^-T^- Nk^+ immunological phenotype: natural killer cells are present, whereas mature B and T lymphocytes are absent in the circulation (76). For the first time in 1993, with a series of elegant experiments, Taccioli *et al* demonstrated the link between NHEJ and V(D)J recombination (77). To determine if DNA repair pathways and V(D)J recombination share common factors, several cell lines defective in various DNA repair factors were assayed for the ability to rearrange the V(D)J recombination substrates. This led to the subsequent identification of the essential roles of Ku70, Ku80, DNA-PKcs, XRCC4 and DNA ligase IV in both NHEJ and V(D)J recombination (77). It was determined that a mutation in any of the core NHEJ factors leads to a dual SCID phenotype: absence of circulating B and T lymphocytes, coupled with increased sensitivity of fibroblasts and bone marrow precursor cells to IR (77). In 1998 Nicolas and colleagues characterized an abnormally radiosensitive subset of B^-T^- Nk^+ SCID patients (71). These radiosensitive SCID (RS-SCID) patients were genetically distinct from "normally" radiosensitive SCID patients in the sense that the V(D)J recombination/NHEJ defect of these RS-SCID could not be rescued by Rag1/Rag2 complementation. Furthermore, while the cell lines derives from these patients were severely impaired in V(D)J recombination capability, they were only mildly impaired in NHEJ (ability to rejoin the blunt ended signal junctions) (78). Additionally, a role for all other known factors involved in the NHEJ/ V(D)J recombination pathway was excluded. This led to the speculation for the

existence of an additional factor involved in both NHEJ and V(D)J recombination that remained un-described (69, 71). This distinct subset of RS-SCID was prevalent amongst Athabascan-speaking Native Americans, allowing Moshous and colleagues to identify and localize the disease-related locus to the short arm of chromosome 10p in a 6.5cM long region (78). In 2001, the gene was successfully cloned and identified as Artemis, an essential factor in V(D)J recombination/DNA repair factor, whose mutation led to RS-SCID (69, 78).

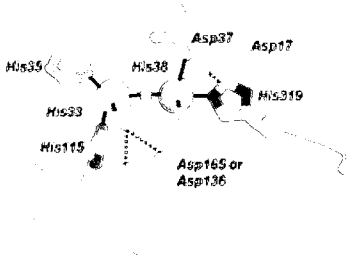
Artemis: Structure and Function

Artemis/SNM1C is a nuclear phosphoprotein, which is comprised of 692 amino acids, and translates into an approximately 77-kDa protein that is organized into three domains (**Figure 7**) (69). The N terminus of Artemis (aa1-385) is encoded by 13 exons and constitutes its catalytic core which consists of two domains: a β-lactamase domain (aa1-155) and a β-CASP domain (aa156-385) (69, 79). The β-lactamase domain classifies Artemis as a member of the metallo-β-lactamase superfamily of enzymes possessing a wide variety of substrates. The β-CASP (metallo-β-lactamases-associated CPSF, Artemis, Snm1, PSO2) domain does not exist as an isolated domain, but is always appended to the metallo-β-lactamase domain. β-CASP is unique to several enzymes within the metallo-β-lactamase superfamily capable of processing nucleic acids(80, 81). The β-lactamase/ β-CASP domain of Artemis is essential for its enzymatic activity (82). The C-terminal domain (aa386-692) is encoded by a single exon and is non-conserved between members of the SNM1 family of proteins. This domain contains three basal phosphorylation sites (Ser-385, Ser-516, Ser-518) as well as numerous other sites that are phosphorylated in response to DNA damage by members of the PIKK family of kinases (83). Although the C-terminus of Artemis is not directly involved in its catalytic activity, it has been proposed to play an important regulatory

A



B



β-CASP

Unique only to members of metallo-β-lactamase that process nucleic acid:

CPSF, ARTEMIS, SNM1/ PSO2

C

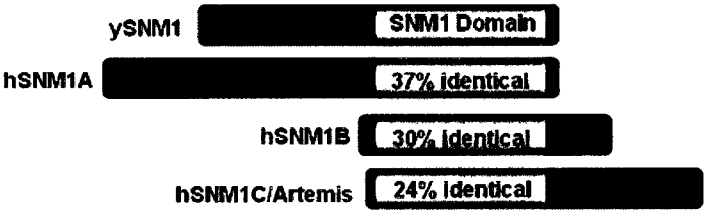


Figure 7. Structural and functional organization of Artemis. **(A)** Artemis is a modular protein with a molecular weight of 77-kDa and is comprised of 692 amino acids. The metallo-β-lactamase domain spans the first 155 amino acids which contains four out of the five conserved catalytic residues. The β-CASP domain (amino acids 155-385) contains the fifth conserved residue of the metallo-β-lactamase domain, and defines the catalytic site of Artemis. **(B)** β-CASP containing proteins are a subfamily of the metallo-β-lactamase domain that are involved in the metabolism of nucleic acids, including: the cleavage and polyadenylation specific factor (CPSF), Artemis, the mouse DNA repair factor SNM1 and its yeast counterpart PSO2. The C-terminal domain of Artemis is non-conserved (aa386-692) and contians a cluster of SQ motifs that are phosphorylated in response to DNA damage by members of the PIKK family of kinases [Adapted from (82, 84)]. **(C)** The SNM1 family of proteins include hSNM1A, hSNM1B and hSNM1C which share 37%, 30% and 24% homology respectively, to the metallo-β-lactamase domain of the yeast SNM1 protein [Adapted from (85)].

role in Artemis activity (63). Sequence database searches of mammalian genes originally identified three members of the metallo-β-lactamase which shared homology with the yeast SNM1 (sensitivity to nitrogen mustard)/PSO2 gene: hSNM1A, hSNM1B and hSNM1C/Artemis. The more distantly related cleavage and polyadenylation specificity factor (CPSF) was later identified. All SNM1 homologs share a homologous region of approximately 300 amino acids (metallo-β-lactamase domain) but are unique outside of this domain and have distinct functions (86).

Artemis Expression

Artemis expression was examined via Northern blot and PCR analysis using a range of hematopoietic and non-hematopoietic tissues (69). Artemis was shown to be ubiquitously expressed at low levels in both hematopoietic and non-hematopoietic tissues: Artemis RNA levels were not increased in bone marrow tissue, the site of V(D)J recombination (69).

Artemis: Phosphorylation and Endonucleolytic activity

Like other members of the β-lactamase family, Artemis was predicted to possess nucleolytic capability. In 2002 Ma *et al* demonstrated the 5'→ 3' exonucleolytic activity of Artemis on single stranded DNA (61). Upon association with the DNA-PKcs, the exonucleolytic activity of Artemis is modified to endonucleolytic activity, allowing Artemis to process 5' and 3' overhangs as well as cleave hairpin junctions (61). This observation further solidified the role of Artemis in V(D)J recombination. The Artemis/DNA-PKcs complex was shown to resolve hairpins generated by RAG proteins *in vitro* (61, 87). Numerous studies have shown that all three PIKK kinases (ATM, ATR and DNA-PKcs)

have the capability of phosphorylating Artemis in proliferating cells and that Artemis is hyper-phosphorylated in response to DSB inducing agents (61, 62, 63, 64). The individual *in vivo* contribution of each kinase, the nature and regulation of Artemis phosphorylation, and the activation of its endonucleolytic activity have remained the subjects of study and debate in the past recent years (61, 62, 63, 94, 95, 99). PIKK kinases preferentially phosphorylate serine or threonine followed by glutamine (SQ motifs) residues. Artemis contains a cluster of 10 SQ motifs, eight of which are located in the C-terminal 200 amino acids (64). The phosphorylation of recombinant Artemis by DNA-PK was assayed in our laboratory; where we observed that the truncation of C-terminus Artemis containing clusters of SQ motifs (removing amino acid residues 445–647) reduced Artemis phosphorylation by 80% compared to full length Artemis (without disturbing association of Artemis with DNA-PK) (63). Furthermore, six phosphorylation sites on Artemis clustered in the C-terminal region were identified (S516, S531, S537, S541, S546 and S645) and were shown to contribute to 70% of the total basal phosphorylation of Artemis in HEK293T cells (63). DNA-PK phosphorylation of recombinant Artemis containing serine-to-alanine substitutions in all of these six serine residues revealed an 85% reduction in Artemis phosphorylation compared to wild type Artemis (63). In addition, basal phosphorylation of S516 and S645 residues was confirmed *in vivo* in a DNA-PKcs dependent manner (via anti-phosphorserine antibodies in unirradiated HEK293T cells). Interestingly, following treatment of cells with the DSB-inducing drug belomycin, phosphorylation of S516 and S645 proceeded (phosphorylation enhanced 10- to 15- fold) without association of Artemis with DNA-PK. This observation is consistent with the involvement of additional kinases in the DNA damage response-dependent phosphorylation of Artemis (63). In support of this finding, Goodarzi *et al* showed that ATM hyper-phosphorylation of Artemis occurs at S516 *in vitro* and S645 *in*

vivo in response to DNA damage (64). Another study found that S534 and S538, in addition to S516 and S645 were also phoshorylated *in vivo* primarily by ATM (88). Despite their previous report suggesting that ATM signaling regulates the activity of Artemis by phosphorylation (93), Goodarzi *et al* (using multisite phosphorylation mutants) showed that Artemis hyper-phosphorylation by ATM was in fact not essential for its endonuclease activity *in vitro*, nor for DSB repair and V(D)J recombination *in vivo* (64, 89). It has been agreed upon however, that the association of Artemis with DNA-PKcs is essential to 'support' Artemis activity (89). It is unclear however, whether DNA-PKcs is required to phosphorylate Artemis for activation of its endonucleolytic activity, or whether the auto-phosphorylation of DNA-PKcs itself is required for 'remodeling' of the DNA-PK holoenzyme to facilitate physical access of Artemis to the DNA termini (5, 64, 89, 90). Inconsistent with this model, our laboratory showed that an Artemis mutant defective in interaction with DNA-PKcs was surprisingly proficient in the joining of DSBs created during V(D)J recombination (63). In addition, Goodarzi et al (2006) showed that mutation of DNA-PKcs/ATM phosphorylation sites in Artemis has no effect on its endonuclease activity *in vitro* (64).

How essential is the endonuleolytic activity of Artemis?

The activity of Artemis during V(D)J recombination is essential: its deficiency leads to a complete absence of circulating B and T lymphocytes, since processing of the coding end is a necessary reaction in V(D)J recombination (69). The precise function of Artemis in NHEJ however is not essential and is less understood. Artemis-deficient cells are not, in fact, significantly defective in DNA repair. Following exposure to IR, Artemis-deficient cells retain 10-25% of unrepaired DSBs at later timepoints after IR (21). Results from *in*

vivo studies have suggested that Artemis might be involved in the processing of a subset of non-ligatable ends that are more complex in nature and thereby require additional processing (21). However, in contrast to its "less essential" role in NHEJ, Artemis-deficiency in patients and mice confers genomic instability similar to other essential NHEJ factors (91). In addition, Artemis-deficient patients display B cell lymphomas and aberrant chromosome rearrangements, which have led to the speculation that Artemis may also play a more central role as a genomic caretaker (92).

Artemis links ATM to DSB repair

ATM and DNA-PKcs are both major kinases involved in various branches of the DNA damage response network, including DSB repair. While ATM plays a less understood and "less essential" role, DNA-PKcs plays an essential and distinct role within the NHEJ pathway. Although the majority of DSB repair is ATM independent, ATM is required for the repair of a subset of DBSs (~10-25%), in a thus far uncharacterized pathway via the phosphorylation of the Artemis (93) and DNA-PKcs (21, 94). This repair defect, although subtle, leads to profound radiosensitivity in ATM deficient (A-T) cells, identical to that which is observed in Artemis-deficient cells. Given that Artemis has been identified as a substrate of ATM and double mutations of ATM and Artemis do not lead to an increased repair defect, ATM and Artemis are proposed to be involved in the same DSB repair pathway and/or are required for the repair of the same subset of DSBs (21). In 2006 Chen *et al* showed that that the ATM-dependent phosphorylation of DNA-PKcs at threonine-2609 is required for efficient activation of DNA-PKcs (94). Similarly, DNA-PKcs is essential in stimulating Artemis activity. Thereby, it is possible that ATM is involved indirectly in regulating Artemis activity via activation of DNA-PKcs.

It has been speculated that this uncharacterized ATM/Artemis-dependent repair pathway may involve the repair of a subset (~10-25%) of DSBs which are more complex in nature and may require additional factors or time for processing prior to ligation (21). This proposal is supported by the finding that ATM and Artemis are not essential for the repair of etoposide induced DSBs; however, reliance upon ATM and Artemis increases for the repair of more complex lesions that were generated by densely ionizing α particles (21). More recently, it was proposed that this subset of complex DSBs that require ATM signaling may not necessarily correlate with damage complexity, but rather with heterochromatinization (21). For the first time, Goodarzi *et al* demonstrated that the subset of DSBs that persist in the absence of ATM signaling localize to heterochromatin, that heterochomatinized chromosomes accumulate increased breaks in the absence of ATM activity, and that cells with relaxed heterochromatin have a smaller reliance on ATM signaling for DSB repair (21).

Artemis in cell cycle checkpoint arrest

In 2004 a study conducted by Zhang *et al* reported a novel function for Artemis as a protein involved in IR-induced cell cycle checkpoint arrest, rather than its presupposed role as an NHEJ factor. It was proposed that Artemis-deficient cells exhibited radiosensitivity due to a defect in the G_2/M DNA damage checkpoint, rather than a defect in the NHEJ pathway (95). An *in vitro* DNA end-joning assay using human cell extracts depdendent on DNA-PK, XRCC4, and DNA ligase IV was used as a measure of NHEJ repair. To determine the contribution of Artemis in NHEJ, a DNA repair assay was preformed *in vivo* after exposure to IR, with and without immunodeleption of Artemis (95). In contrast to DNA-PKcs-deficient cells, Artemis-deficient cells did not exhibit a significant defect in DSB repair. Intriguingly, human cancer cell lines depleted of Artemis irradiated with IR

did not maintain arrest at the G_2/M checkpoint after IR, implicating a novel function for Artemis in checkpoint arrest (95). An additional study was conducted by the same laboratory in 2007, suggesting that Artemis functions as a negative regulator of recovery from G_2/M checkpoint (88). Geng *et al* reported defective recovery from the G_2/M cell cycle checkpoint with mutant Artemis (S516/ S645 substituted with alanine residues), proposing that the phosphorylation of Artemis by ATM in these latter residues has an inhibitory effect that activates the Cdk1-cyclinB complex upon completion of DNA repair (88). By contrast, another study conducted by Krempler *et al,* found that Artemis-defective cells were indeed proficient in checkpoint arrest, and maintained a prolonged G_2/M arrest after IR due solely to DSB repair defect, thus exhibiting a cooperative phenotype between checkpoint and repair functions (96). They elucidated that the former study conducted by Zheng *et al* (2004) was carried out using human tumour cells for the siRNA knock down of Artemis, explaining that such cells have abnormal levels of Chk1/Chk2 which lead to abnormal cell cycle checkpoint regulation (96). Intriguingly, Krempler *et al* found similar levels of Artemis/ATM-dependent DSB repair in all stages of the cell cycle (prevailing evidence suggests that NHEJ is operates predominately in G_0 and G_1) (97). This finding could represent two possibilities: Artemis is indeed involved in a DSB repair pathway beyond NHEJ, or perhaps NHEJ operates in all cell cycle phases; in either case HRR cannot compensate for Artemis deficiency in G_2 phases of the cell cycle (96). There is also increasing evidence emerging in favour of the existence of an alternative end joining pathway capable of directly ligating DNA ends in the absence of NHEJ (59, 98, 99, 100, 101, 102).

Artemis: To Be Continued...

Despite the extensive research and significant insight that has been gained on the role of Artemis in NHEJ, cell cycle checkpoint activation, ATM/Artemis association, Artemis/DNA-PKcs association, and Artemis regulation, further elucidation is required to define: (A) the specific role/requirement of Artemis in DSB repair, (B) the role of Artemis in cell cycle checkpoint activation and genomic maintenance, (C) the cell's choice of Artemis hyperphosphorylating kinase(s), (D) the contribution of Artemis hyperphosphorylating kinase(s), (E) the functional significance of these phosphorylations, (F) the precise mode/mechanism of Artemis activation, and (G) the uncertain function of the highly phosphorylated Artemis C-terminal region.

PARP-1

PARP-1 has been studied extensively since it was first identified in 1963 as the enzyme responsible for poly(ADP-ribosyl)ation and has been a protein of recent interest in our laboratory. PARP-1 is a modular protein and is composed of 1014 aa which translates into a 113-kDa protein. PARP-1 has a highly conserved structural organization that is organized into three domains (**Figure 8**). PARP-1 has an N terminal DNA-binding domain (DBD) (aa1-372), which contain two Cys-Cys-His-Cys zinc finger motifs (FI and FII) that act as DNA damage "sensors" recognizing altered DNA structures. PARP-1 contains a nuclear localization signal (NLS) (aa202-231) within its DBD domain which targets PARP-1 to the nucleus. The NLS comprises the DEVD cleavage site, which is recognized by caspases-3 and -7 and cleaves PARP-1 into approximately 25-kDa and 85-kDa fragments (103). PARP-1 also possess a central auto-modification domain containing a BRCA1 C-terminus like (BRCT) motif (aa372-524), that is essential for mediating interactions with protein partners and forming homo- and heterodimers with PARP-1 monomers (40). This

auto-modification domain also contains several glutamic acid residues for the addition of PAR. The C terminus of PARP-1 encodes its catalytic domain (aa524-1014), which contains the "PARP signature" motif (aa859-908) (40). It has been shown that both Ku70 and DNA-PKcs are poly(ADP-ribosyl)ated by PARP-1, and that PARP-1 is phosphorylated by DNA-PK (38, 104, 105). Furthermore, various studies have reported increased levels of DSB in PARP-1 deficient cells (106, 107). Intriguingly, PARP-1-deficient cells are not defective in DSB repair. However, cells that are defective in NHEJ show delay in DSB repair, yet are still capable of rejoining the majority of DSBs (108, 109). Therefore, it is possible that PARP-1 is involved in the previously proposed alternative "backup" NHEJ pathway, which is DNA-PK- independent, displays slower repair kinetics, and prevails when DNA-PK dependent NHEJ is compromised (110, 111). Although PARP-1 has been extensively researched in the past 45 years, many questions remain unanswered about the regulation of PAR synthesis and degradation, and the biochemistry of the various pathways in which PARP-1 has been implicated. How the length and complexity of PAR polymers are regulated remains unknown. PARP-1 has been implicated in DNA DSB repair and cell cycle progression, yet its specific role in these processes remains to be elucidated (40).

Protein-protein Interactions

DNA encodes all of the information required for the assembly and function of a cell (112). The human genome consists of approximately 20-25,000 genes, which give rise to the dynamic, multi-dimensional architecture of the cell through transcription of genes, translation of RNAs, post-translational modifications, and gene splicing mechanisms to give rise to $1.0x10^6$ proteins (112, 113). Although some of these proteins function as monomeric units, the majority of proteins function in association with other proteins, factors, or as

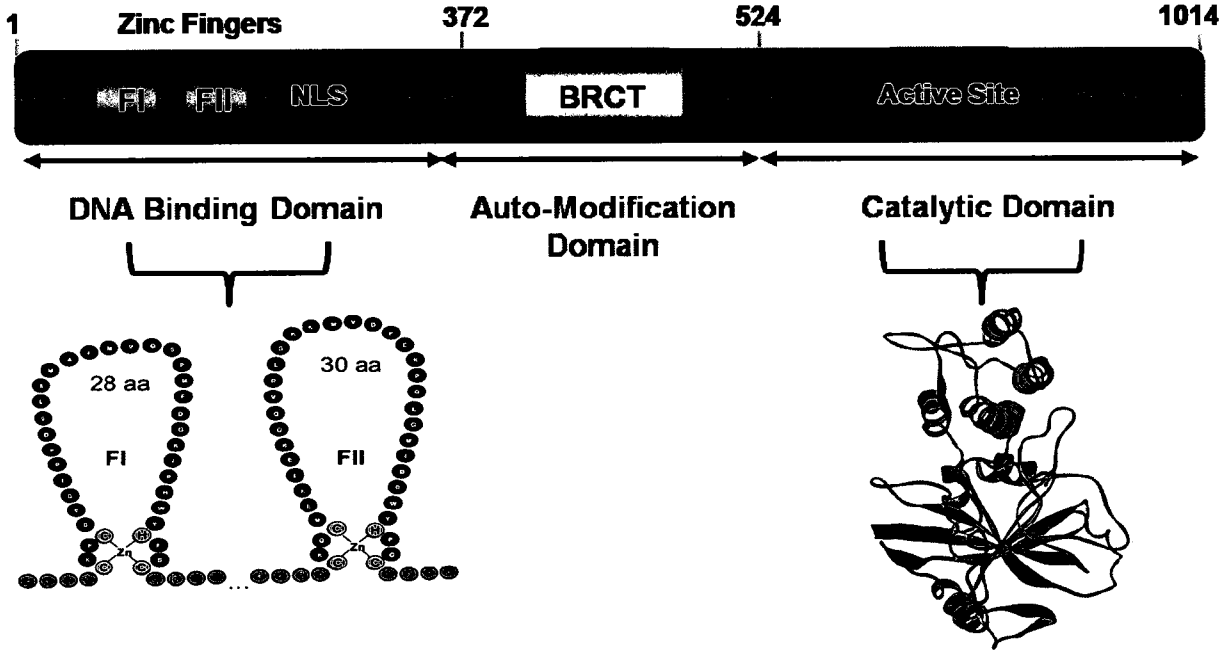

1
Zinc Fingers
372
524
1014
FI
FII
NLS
BRCT
Active Site
DNA Binding Domain
Auto-Modification Domain
Catalytic Domain
28 aa
30 aa
FI
FII
Zn
Zn

Figure 8. The domain architecture of human PARP-1. PARP-1 is a modular protein comprised of 1014 aa and translates into a 113-kDa protein. PARP-1 consists of an N-terminal DNA-binding domain with two Cys-Cys-His-Cys zinc finger motifs (FI and FII), a nuclear localization signal, a central automodification domain containing a BRCT protein–protein interaction motif, a C-terminal catalytic domain with a contiguous 50-amino-acid sequence, the "PARP signature" motif, that encodes the active site [Adapted from (40)].

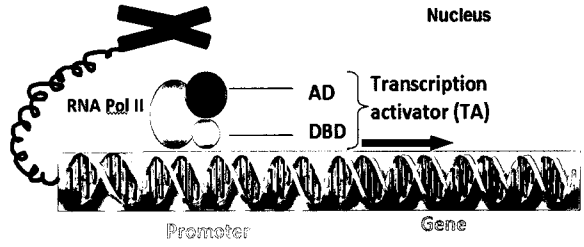
components of large molecular assemblies (112). Since the emergence of the Human Genome Project, identifying and characterizing protein-protein interactions within the interactome has become the milestone to understanding the biochemistry of the cell (112). Protein-protein interactions can be classified according to their stability and their mode of interaction; interactions can be stable or transient in nature, both of which can be either strong or weak (114). Stable interactions are those associated with proteins that are purified as multi-subunit complexes. The subunits of the complex can be identical or different. Stable interactions are best studied by co-immunoprecipitation, pull-down or far-western methods (114). Proteins involved in transient interactions are often regulatory proteins, and are expected to regulate or modify the majority of cellular processes (114). Transient interactions are short-lived, or on-and-off in nature, and serve mainly to confer temporary, but regulated physiological properties to the proteins or core complexes with which they interact (114). For example, if two proteins interact during a specific stage of the cell cycle or in response to a signal, the interaction is likely transient in nature (114). Transient interactions can also be strong or weak in nature, and kinetically, fast or slow. Transient interactions are best captured by Y2H screening, cross-linking or label transfer methods (115).

Yeast Two Hybrid Assay

Equally important to understanding and analyzing protein-protein interactions, has been the development of new techniques to experimentally analyze such interactions. A major advance in the study of protein-protein interactions was the development of the yeast two hybrid screening system by Fields and Song in 1989 (116). The yeast two-hybrid system (Y2H) is an *in vivo* method that can be used to identify novel protein interactions,

analyze known protein-protein interactions, or to examine the interactions between specific protein domains in the yeast *Saccharomyces cerevisiae* (116). The basis of the two-hybrid system was initially developed from studies conducted in 1987 by Ptashne and colleagues, on the modular nature of eukaryotic transcription factors (117). The Y2H system exploits the modular nature of eukaryotic transcription factors which consist of two separate domains, a DNA binding domain (DBD) and an activation domain (AD). The key to the Y2H screen is that the two transcription domains need not be on the same protein, and are capable of activating transcription if they are brought within close proximity of one another (116, 118, 119). Fields and Song demonstrated that transcription can be activated via independent expression of two functional domains of the yeast Gal4 activator protein as chimeric fusions with proteins that interact *in vivo* (116). In this system, two separate hybrids are constructed: the protein of interest (X) is fused to DBD and if it does not activate transcription of reporter genes on its own, it is used to screen a library of cDNA clones (Y) that are fused to an AD. These two hybrids are transformed into a yeast host strain that contains at least one nutritional reporter gene (auxotrophic marker(s)) integrated into its genome and *lacZ* (β-galactosidase) (118). The regulatory regions of these reporter genes contain upstream activating sequences (UAS) cognate to the DBD vector and are activated by the binding of reconstituted transcription factor (119). If protein X interacts with protein Y in the nucleus, the AD is brought within physical proximity of the DBD, which reconstitutes transcriptional activation and results in the expression of the reporter gene (**Figure 9**). Positive interactions are then detected by selecting on media lacking the appropriate auxotrophic marker(s), followed by second screen for β-galactosidase activity.

A


Nucleus
RNA Pol II
AD
DBD
Transcription activator (TA)
Promoter
Gene

B

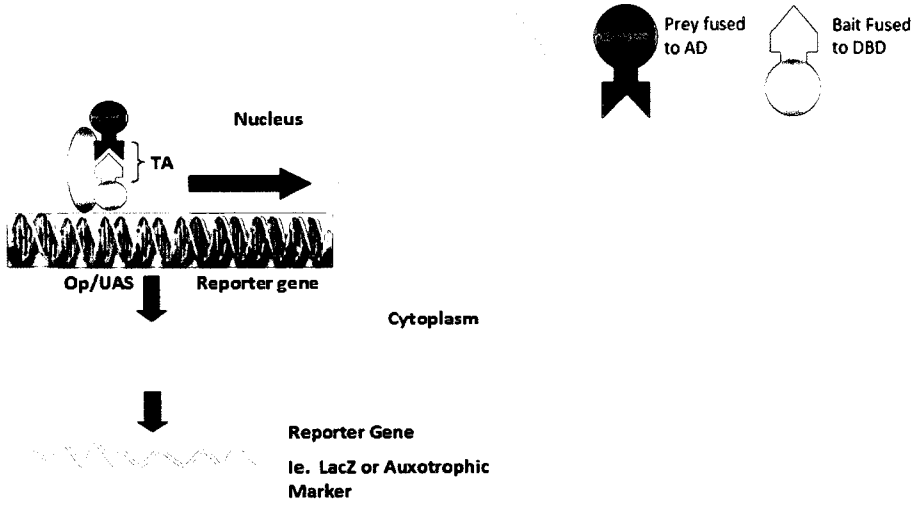

Prey fused to AD
Bait Fused to DBD
Nucleus
TA
Op/UAS
Reporter gene
Cytoplasm
Reporter Gene
Ie. LacZ or Auxotrophic Marker

Figure 9. Basis of yeast two hybrid principle. **(A)** Yeast two hybrid principle is based on the fact that transcription factors are modular in nature and consist of two domains: a DNA binding domain (DBD) and an activation domain (AD). **(B)** Schematic of the basic yeast two-hybrid system to detect interactions between two proteins in a yeast cell. Two hybrids are constructed: a DBD-fused bait protein of interest interacts with an AD-fused prey protein, either known or selected from a cDNA library. The interaction reconstitutes transcriptional activation which binds a specific sequence motif [op (operator) or UAS (upstream activating sequence) depending on whether LexA or GAL4 is used as DBD], which activates transcription of two (or more) reporter genes which produces visible phenotypes. Positive interactions can be detected by selected on media lacking the appropriate nutritional marker(s), followed by an additional screen for expression of the β-galactosidase gene [Adapted from (125)].

Advantages and Limitations of the Yeast Two Hybrid System

The Y2H system has proven to be a powerful technique for efficient genome wide screening and offers particular advantages over classical biochemical and genetic approaches (120, 121). First, interactions are assayed *in vivo* using the eukaryotic yeast as a host. Using live yeast cells as a test tube allows for the genome wide detection of interactions within the native environment of the cell eliminating the need for biochemical purification. Additionally, this allows for the assessment of protein-protein interactions in a more physiological framework compared to *in vitro* based techniques (120). Second, very high numbers of potential coding sequences are assayed and can be scaled to high-through put screenings. A significant virtue of this highly sensitive system is the ability to detect interactions that are weak or transient in nature and may not be detected by other *in vitro* methods (122). Furthermore, no previous characterization of the interacting proteins is necessary, and the sequence of the identified putative interactor is immediately available (120). Another strength of genome wide screening using theY2H system is that all combinations of protein-protein interactions are assayed. However, this can lead to the enrichment of false positives since an interaction may occur between two proteins that may never be present in the same cell type or the same sub- cellular compartment, or they may never be expressed at the same time or exist within close proximity to each other (120, 121). Indeed, the biggest disadvantage of Y2H screening is the high incidence of false positives and false negatives. The majority of putative interactors identified via Y2H screening are biologically irrelevant, known as false positives; conversely, physiological protein-protein interactions may not be detected, known as false negatives (can overlook up to 90% of known interactions) (121, 122, 123). Also, bait and prey proteins are expressed as fusions to

modules of transcriptional machinery, which embodies a potential risk. Fusion proteins may alter the activity of the protein by modifying the accessibility of binding site(s) on the protein-protein interface or hinder the proper folding orientation of the protein binding domain. Furthermore, some protein-protein interactions depend on post-translational modifications that may not occur appropriately or may be absent in yeast, including formation of disulfide bridges, glycosylation and most commonly phosphorylation (120). Another disadvantage arises from the high presence of miss-frame proteins or short peptide sequences in the cDNA library; providing a source of non-specific interactions that often pollute the results of the Y2H screens by enriching the number of non-biologically relevant interactions (121). Furthermore, expression of certain bait or prey proteins may affect general colony viability under certain selective conditions and lead to reporter gene activation (without reconstitution of the transcription factor) (120). Mutations or other random recombination events may invoke such activity; however, like all other false positives these uncommon cases are rarely mechanistically explained (120, 121).

Variants of the Yeast Two-Hybrid Systems

A wide range of Y2H systems are currently available, each employing various combinations of DB, AD, and reporter genes. A popular alternative to the classical Gal4 system is the LexA based Y2H system, which is also based on the modular reconstitution of transcription factors (124). The LexA based Y2H system utilizes a DNA binding domain from the *E. coli* LexA protein (amino acids 88-202), and an activation domain from VP16, a genes with promoters containing LexA operator sequences. Interaction mating is a highly efficient technique that operates on the same principle as conventional yeast two hybrid

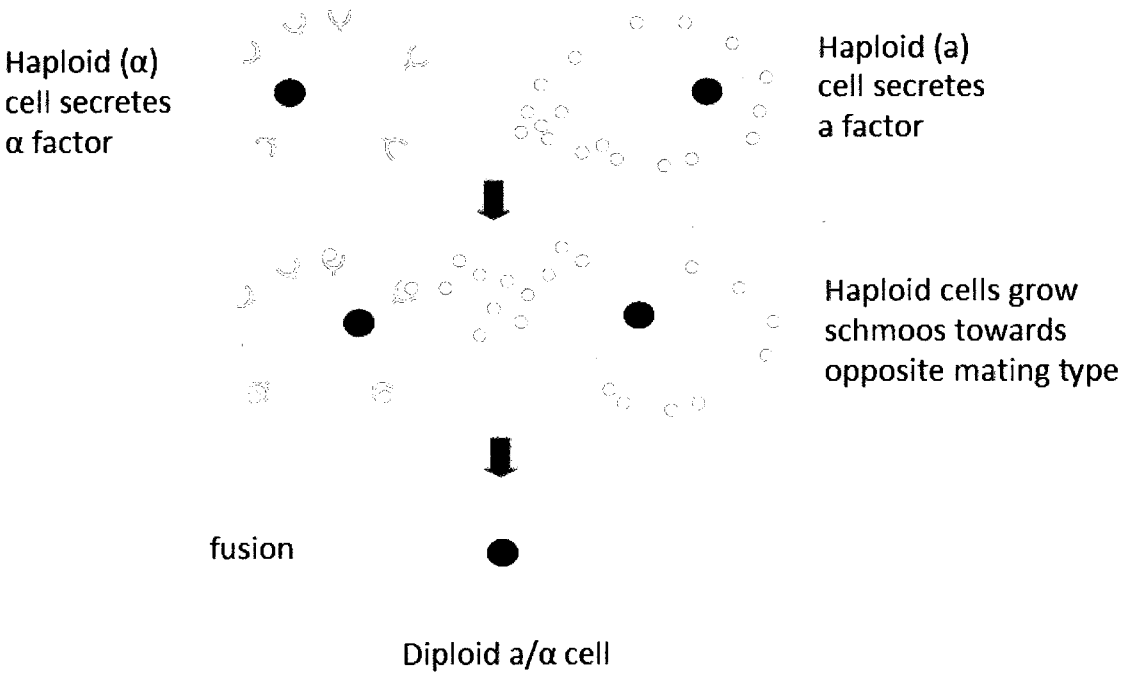

Haploid (α)
cell secretes
α factor
Haploid (a)
cell secretes
a factor
Haploid cells grow
schmoos towards
opposite mating type
fusion
Diploid a/α cell

Figure 10. Simplified schematic of yeast mating. Two haploid cells of opposite mating types secrete mating pheromones (specific to cell surface receptors on opposite mating type), grow projections (schmoos) towards source of pheromone and mate, forming a diploid cell [Adapted from (125)].

systems, but eliminates the need for large scale yeast transformations (**Figure 10**). The cDNA library in the interaction mating system is screened using a yeast mating procedure; may be more suitable for certain bait proteins as it is less intrusive, eliminates variability, and may lead to more reproducible results (126). *S. cerevisiae* can stably exist as either a diploid or a haploid ('a' or 'α'). Haploid cells are capable of mating with other haploid cells of the opposite mating type (by releasing mating pheromone) to produce a stable diploid cell. In interaction mating, the AD and DBD fusion proteins begin in haploid cells of opposite mating type and are brought together via mating (127).

Objective and Rationale

My objective was to identify binding partners for Artemis and PARP-1 in attempt to identify novel functions or to clarify and characterize their previously identified roles in the various branches of the DNA damage response. I started with the theory that proteins belonging to the same functional class have a high probability of working together, and thus potentially have a high number of interactions between one another. I hypothesized that I could identify binding potential substrates or binding partners (likely involved in the DNA damage response network) for Artemis and PARP-1 using a Y2H screen. PARP-1 is an extensively studied protein and has numerous binding partners that have been previously identified; however, little is known about the binding partners of Artemis. I rationalized that a Y2H screening approach would be the most appropriate method for identifying Artemis interacting proteins for several reasons: (i) Artemis is endogenously expressed at very low levels, and previous studies have shown that over-expression of other members of the Snm1 family have led to cellular toxicity (69). (ii) Previous studies have also found that ectopic expression of Artemis is difficult to maintain and sensitive to truncations/amino acid changes

(125). (iii) I speculated that most Artemis interactions would be transient in nature due to the enzymatic function of its catalytic core. The Y2H system eliminates the need for maintaining high levels or stable expression of the protein of interest. Furthermore a wide variety of interactions are readily screened using this technique, while bait and prey proteins are expressed at comparable levels. Also, this system operates in the yeast nucleus, and given that Artemis is a nuclear protein it will likely be in a more physiological state and likely adopt its intrinsic phosphorylation/folding state. I hypothesized that I will identify Artemis and PARP-1 interacting proteins via Y2H screening, which will either shed new light on Artemis functions in NHEJ, or reveal a novel function for Artemis beyond its role in NHEJ. I attempted to employ the C-terminus of Artemis as bait to identify binding partners specific to this region, to gain insight about its largely uncharacterized function/regulation. The regulation and activity of PARP-1 is of great interest in our laboratory. Several studies have been conducted that identified PARP-1 binding proteins using Y2H systems successfully; screening with PARP-1 served as a positive control in this study, in addition to identifying novel binding partners/functions.

MATERIALS AND METHODS

There were two systems used for yeast two-hybrid screening: Hybrid Hunter™ Two-Hybrid LexA based System (Invitrogen™, Burlington, ON) and the Matchmaker™ Yeast Two-Hybrid Gal-4 based system using interaction mating (Clontech Laboratories, Inc.; Mountain View, CA).

Hybrid Hunter™ Two-Hybrid LexA based System

The Hybrid Hunter™ system was used to screen Artemis; using the Hybrid Hunter™ human lung cDNA library (Invitrogen™, Burlington, ON), and DupLEX-A™ Human Fetal Brain cDNA library (OriGene Technologies, Inc.; Rockville, MD).

Yeast Host Genotype and Phenotype

The L40 yeast strain contains two integrated reporters: the yeast *HIS3* gene and the bacterial *lacZ* gene. The first reporter gene is under the control of four LexA operators and the second reporter places the *lacZ* under the control of eight LexA operators. Association of bait and prey protein activates transcription of two reporter genes and as a result yeast grow in the absence of histidine and turn blue when β-galactosidase activity is assayed using the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal). Please refer to **Table 1** for an overview of L40 characteristics and Appendix I for appropriate growth media.

Table 1. The genotype and phenotype of yeast host strain L40. The reporter constucts (*HIS3* and *lacZ*) are integrated into the genome of L40 and contain LexA operator sites as upstream activation sequences (128).

Strain	Genotype	Phenotype	Reporter genes
L40	*MATa his3Δ200 trp1-901 leu2-3112 ade2 LYS2::(4lex*Aop-*HIS3)URA3*::(8*lex*Aop-*lacZ*) *GAL4*	His-, Trp-, Leu-, Ade-	*HIS3* *lacZ*

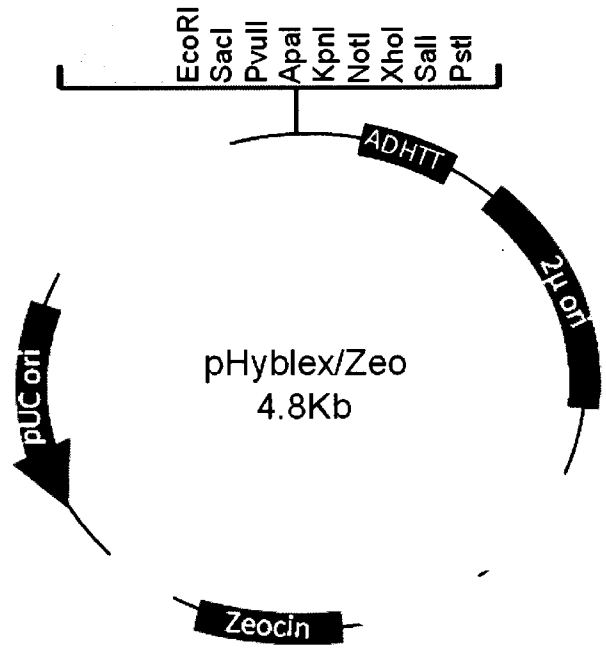
Construction of LexA DNA-BD+ bait Fusion Vectors

The bait plasmid employed in this system was the pHybLex/Zeo vector (**Figure 11**). Three potential bait constructs were prepared by cloning the N-terminus, C-terminus, and full length Artemis in frame with the LexA DBD into pHybLex/Zeo (**Figure 12**).

PCR Reactions

N terminus, C terminus, and Full length Artemis

PCR reactions were carried out with primers designed to generate blunt ended inserts and thus allow for the in-frame insertion of C-terminus, N-terminus, and full fragment regions of Artemis into pHyblexZeo. Primer sequences are listed in Appendix II. The N-terminus region of Artemis (37-1206 bp) was amplified using primers I and II. The C-terminus region of Artemis (1167-2115bp) was amplified using primers III and IV. Full fragment Artemis (37-2115bp) was amplified using primers I and IV. All primers were pre-phosphorylated prior to PCR reactions, consisting of the following cocktail: 400 ng of primer DNA, 10 units of T4 Polynucleotide Kinase (New England BioLabs, Inc), 2 μl of 10X T4 DNA Ligase buffer (New England BioLabs,Inc) in a 20 μl reaction. The reaction was incubated for 2 hours at 37°C , followed by heat inactivation at 65°C for 2 hours. PCR




```
930   AAT AGC GAG TTT AAA CCA ATT GTC GTA GAT CTT CGT CAG CAG AGC TTC ACC
      Asn Ser Glu Phe Lys Pro Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr
                  pHybLex/Zeo Forward priming site                  EcoR I
981   ATT GAA GGG CTG GCG GTT GGG GTT ATT CGC AAC GGC GAC TGG CTG GAA TTC
      Ile Glu Gly Leu Ala Val Gly Val Ile Arg Asn Gly Asp Trp Leu Glu Phe
                Sac I                 Pvu II       Apa I       Kpn I Not I       Xho I      Sal I
1032  AAG CTT GAG CTC AGA TCT CAG CTG GGC CCG GTA CCG CGG CCG CTC GAG TCG
      Lys Leu Glu Leu Arg Ser Gln Leu Gly Pro Val Pro Arg Pro Leu Glu Ser
            Pst I
1083  ACC TGC AGC CAAGCTAATTC CGGGCGAATT TCTTATGATT TATGATTTTT ATTATTAAAT
      Thr Cys Ser
```

Figure 11. Diagram of the pHybLex/Zeo vector. In the Hybrid Hunter yeast two-hybrid system, the pHybLex/Zeo vector was used to express the bait protein (N-terminus, C-terminus and full length Artemis) fused to the LexA DNA binding domain (LexA DBD). Each bait construct was inserted into the multiple cloning site (MCS) of the pHybLex/Zeo vector. In yeast, expression from the constitutive ADH1 promoter (PADH1) results in a protein consisting of an N-terminal LexA DNA-BD followed by the bait. The pHybLex/Zeo vector contains the Zeocin™ resistance gene (expressed from *Streptoalloteichus hindustanus ble* gene) that allows for selection in both *E. coli* and yeast [Adapted from (128)].

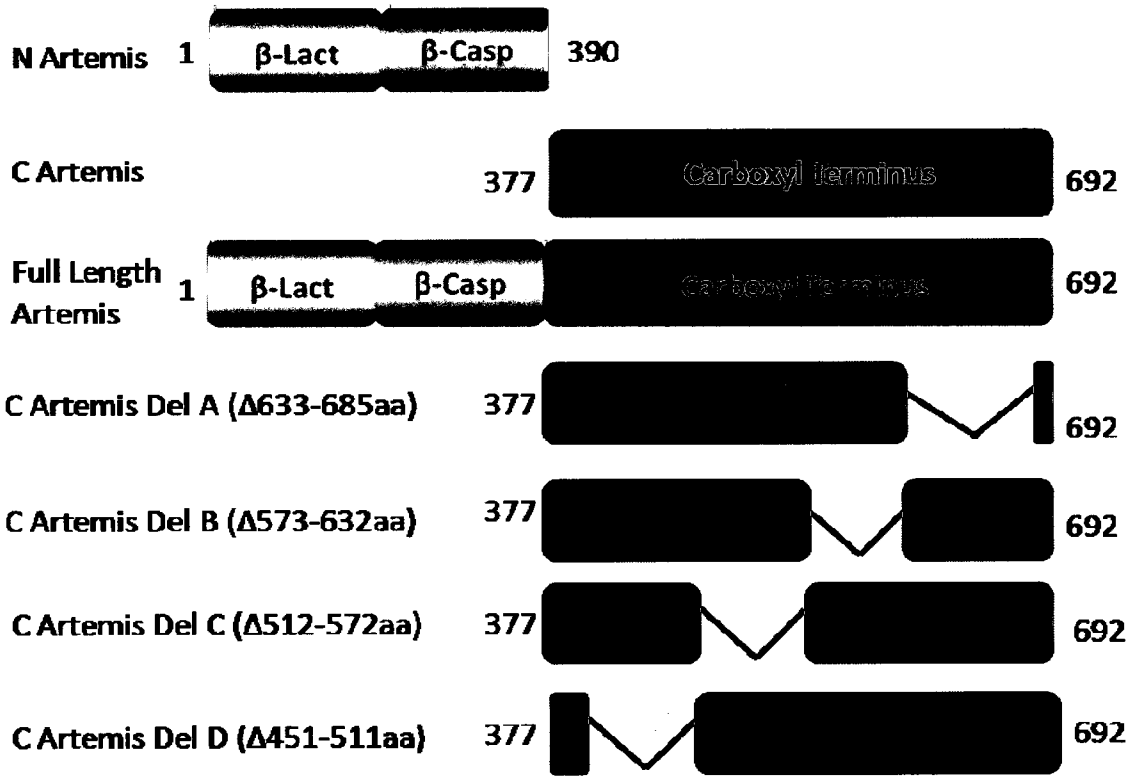

N Artemis
1
β-Lact
β-Casp
390
C Artemis
377
Carboxyl terminus
692
Full Length Artemis
1
β-Lact
β-Casp
692
C Artemis Del A (Δ633-685aa)
377
692
C Artemis Del B (Δ573-632aa)
377
692
C Artemis Del C (Δ512-572aa)
377
692
C Artemis Del D (Δ451-511aa)
377
692

Figure 12. Various Regions of Artemis that were generated via PCR and used to create bait-LexA DBD fusion in pHybLex/Zeo. The same NArtemis, CArtemis and Full Length Artemis regions were used to generate bait constructs in pGBKT7 for interaction mating Y2H screens.

reactions were prepared in 40 µl reaction consisting of the following cocktail: 20 mM Tris-HCl (pH 8.8), 2mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1 mg/ml BSA (bovine serum albumin), 100 ng of each primer, 30 ng template DNA, 400 µM dNTPs (deoxynucleotides), and 2 units of PFU Turbo polymerase. The following program was used in a PTC-200 thermocycler: 30 cycles of denaturing for 30 seconds at 95°C, annealing for 60 seconds at 55°C, and extension for 4 minutes (full length Artemis) or 2 minutes (N and C terminus Artemis) at 68°C.

pHyblexZeo-CArtemis Deletion Constructs

Inverse PCR was used to generate four deletions within the C terminus region of Artemis (1167-2115bp). Inverse PCR reactions were carried out to amplify C terminus regions of Artemis containing the following deletions: pHyblexZeo-CArtemis DelA (Δ 1936-2092bp), pHyblexZeo-CArtemis DelB (Δ1756-1936), pHyblexZeo-CArtemis DelC (Δ1576-1756), pHyblexZeo-CArtemis DelD (Δ1396-1576) (Figure 12). Primers were designed to generate a KpnI site following deletion of appropriate regions from pHyblexZeo-CArtemis for subsequent ligation. Primer sequences are listed in Appendix II. The pHyblexZeo-CArtemis DelA construct was made using primers I and II, pHyblexZeo-CArtemis DelB was made using primers III and IV, pHyblexZeo-CArtemis DelC was made using primers V and VI, and pHyblexZeo-CArtemis DelD was made using primers VII and VIII. The primers were pre-phosphorylated prior to the PCR reactions, as previously described. PCR reactions were prepared in a 50 µl reaction consisting of the following cocktail: 20 mM Tris-HCl (pH 8.8), 2 mM $MgSO_4$, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 0.1% Triton X-100, 0.1 mg/ml BSA (bovine serum albumin), 50 ng of each primer, 15 ng template DNA, 500 µM dNTPs (deoxynucleotides), and 2 units of PFU Turbo polymerase). The

following program was used in a PTC-200 thermocycler: 30 cycles of denaturing for 30 seconds at 95°C, annealing for 60 seconds at 55°C, and extension for 11 minutes at 68°C.

Cloning and Plasmid Preparations

pHybLexZeo-NArtemis, pHyblexZeo-CArtemis, pHyblexZeo-FArtemis

Following PCR amplification, the products were quantified and purified using QIAquick Gel Extraction Kit (QIAGEN, Valenci, CA) from a 1.0% agarose gel: NArtemis, CArtemis, and FArtemis generated 1169 bp, 948 bp, and 2078 bp inserts, respectively. The pHybLex/Zeo plasmid was linearized at the PvuII site of its multiple cloning site generating blunt ends. Vector digestion was set up in a 50 µl reaction containing 1 µl of pHybLex/Zeo (1µg/µl), 1 µl of PvuII (10 units), 5 µl of BSA 10X, 5 µl NEBuffer 2, and 38 µl of water. The reaction was incubated at 37°C for 2 hours. Next, the reaction was incubated with 0.5 µl of calf intestinal alkaline phosphatase (CIP) (New England BioLabs; 1000 units/ml) for 30 minutes at 37 °C. The pHyblexZeo plasmid was then quantified and purified from a 1% agarose gel into 30 µl of EB buffer. A 20 µl blunt-blunt ligation reaction was set up using 3:1 insert to vector ratio using 100 ng of pHybLex/Zeo and 300 ng of insert (C-terminus, N-terminus, and full length Artemis separately) , 0.4 µl of T4 DNA Ligase (2000 units/µl, New England BioLabs, Inc), 2 µl of T4 DNA Ligase reaction Buffer. A 20 µl control reaction was also set up lacking the insert; containing 100 ng of pHybLex/Zeo, 2 µl of T4 DNA Ligase reaction buffer, 0.4 µl of T4 DNA Ligase (2000 units/µl). The reactions were incubated in a 15 °C water bath overnight.

pHyblexZeo-Cterminus Delection(A,B,C,D) constructs

The PCR products were gel purified and re-circularized by setting up a ligation reaction that would regenerate a KpnI site and the appropriate deletions in the construct. A

20 µl blunt-blunt ligation reaction was set up using 200 ng of PCR product, 0.4 µl of T4 DNA Ligase (2000 units/µl), 2 µl of T4 DNA Ligase reaction Buffer in a 20 µl reaction. Reactions were incubated in a 15 °C water bath overnight.

E. coli Transformation

The ligation products were transformed into competent *E. coli* DH5α cells by electroporation, using an *E. coli* pulser (Bio Rad, Mississauga, ON, Canada) and plated on low salt LB plates (5% NaCl) containing 25 µg/ml of Zeocin™ and incubated at 37°C overnight. Several single colonies were selected, and each was used to inoculate 5 ml of liquid low salt LB media containing 25 µg/ml Zeocin™ . The cultures were incubated overnight at 37 °C with shaking at 250 rpm. The plasmid DNA from each culture was then isolated using the small scale plasmid preparations using QIAprep Spin Miniprep Kit (QIAgen, Cat. 27104), following the manufacturer's instruction. The plasmid DNA was then characterized by restriction digest using enzymes that were specific for each insert to identify the clones containing insert in the proper orientation. The clones that contained the appropriate size bands were sent for sequencing using the following primer: 5'-AGGGCTGGCGGTTGGGGTTATTCGC-3'. The sequencing was performed by StemCore Laboratories (OHRI; Ottawa, ON).

Yeast Bait Strain: Yeast Transformation with Bait Construct

The pHyblexZeo-NArtemis, pHyblexZeo-CArtemis, pHyblexZeo-CArtemis Del A,B,C,D and pHyblexZeo-FArtemis bait constructs were then transformed into the yeast L40 strain using the small scale yeast transformation protocol (Appendix III). The transformations were plated on YPAD Z300 plates and incubated for 2-3 days at 30 °C until

large (2-3 mm in diameter) yeast colonies appeared. Yeast cells were stored for up to two months on plates sealed with parafilm at 4 °C. For long term storage, yeast cells were grown in appropriate liquid medium (rich or YC minimal media) at 30 °C overnight. The cells were pelletted via centrifugation (table top microcentrifuge) at 2500 rpm , and an equal volume of 50% sterile glycerol was added the the yeast pellet and stored at -70°C.

Test for Expression of Bait protein

Several fresh colonies of each of the Zeocin™ resistant transformants were grown to characterize the expression of the baits. A single fresh colony from each of the C-Art L40, N-Art L40 and F-Art L40 transformants were each used to inoculate 10 ml overnight culture of YPAD Z300 in 50 ml Falcon tubes. L40 was grown in YPAD as a negative control for expression. The yeast were grown for approximately 16 hours at 30 °C with constant shaking. The yeast culture was centrifuged at 1500 g for 10 minutes. The pellets were flash frozen at -80 °C for 10 minutes. The cells were thawed in 200 µl of cracking buffer (8 M Urea, 5% SDS, 40 mM Tris-HCl pH 6.8, 0.1 mM EDTA, 1% β-mercaptoethanol, 0.4 mg/ml bromophenol blue) pre-warmed (60 °C). The cell suspension was transferred to a 1.5 ml microcentrifuge tube containing 200 µl of acid washed glass beads (Sigma G-8772, 425-600 microns). The solution was incubated at 70 °C for 10 minutes and vortexed for 5-10 minutes. The solution was centrifuged at 14,000 rpm (table top microcentrifuge) for 5 minutes. A 10 µl aliquot of the resulting supernatant was used for immunoblot analysis to detect expression of bait-LexA fusion with Anti-LexA Antibody.

SDS-Polyacrylamide gel electrophoresis

SDS-PAGE was performed essentially as previously described (129). Protein samples were dissolved in sample buffer (62.5 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.125% (w/v) bromophenol blue) and denatured by boiling for 5 min at 95 °C. Samples were loaded on SDS-PAGE gels which consisted of stacking gel (4% (w/v) 36.5:1 acrylamide:N,N'-methylenebisacrylamide, 62.5 mM Tris-HCl pH 6.8, 0.1% (w/v) SDS, 0.1 % (w/v) ammonium persulfate, 0.1% (v/v) TEMED) and a resolving gel of (10% (w/v) 36.5:1 acrylamide:N,N'-methylenebisacrylamide, 375 mM Tris-HCl pH 8.8, 0.1 % (w/v) SDS, 0.05% (w/v) ammonium persulfate, 0.05% (v/v) TEMED). Gel apparatus (Bio Rad) with a gel thickness of 1.5 mm was used. Gels were run at 140 volts for 50-60 minutes in electrophoresis buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1% (w/v) SDS). Gels were examined by western blotting.

Western Blotting

Western blotting was performed as previously described (130). After separation by SDS-PAGE, proteins were electrophoretically transferred from gel to a PVDF (polyvinylidene fluoride) membrane (Immobilon P, Millipore Corp., Bedford, MA) in transfer buffer containing 25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% (w/v) SDS, 10% methanol. The PVDF membrane was prepared by soaking in methanol for 30 seconds. Transfers were carried out for 60 minutes with cooling at 100 V using Bio Rad transblot apparatus. After transfer, the membrane was washed with PBS-T (PBS with 0.1% Tween-20) for 10 minutes at room temperature. The membrane was then blocked with 5% (w/v) skim milk in PBS-T for 20 minutes. The membrane was rinsed with PBS-T and incubated with primary antibody (anti-LexA antibody (LexA D-19 sc-1726 goat polycolonal IgG; Santa Cruz Biotechnology)) at a dilution of 1:1000 in PBS-T overnight at 4 °C. The

membrane was washed with PBS-T for 10 minutes three times at room temperature with gentle shaking. The membrane was then incubated with HRP-conjugated secondary antibody (donkey anti-goat IgG-HR sc2020; Santa Cruz Biotechnology), diluted to 1:50000 in PBS-T. Washing was then repeated as for the primary antibody. The antibody-labelled proteins were detected by enhanced chemiluminescence using Amersham ECL kit (Amersham Life Science Inc., Oakville, ON, Canada) as per manufactures instructions.

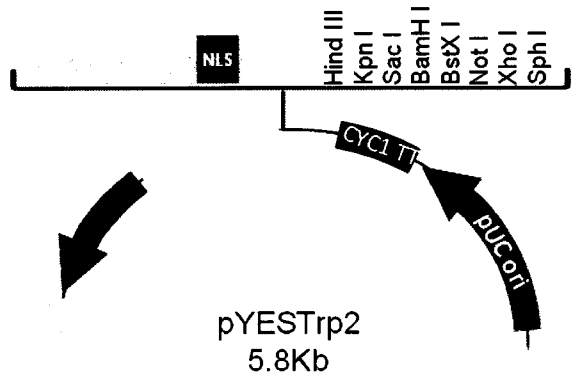
Testing Bait plasmid for non-specific Activation

Each bait strain was transformed with the empty library vector pYESTrp2 (**Figure 13**) using the small scale transformation protocol (Appendix III). The transformation was plated onto YC-W Z300 plates (minimal defined medium for yeast, SD medium is used for selective growth of yeast auxotrophs (Appendix I)) and incubated for 2-3 days at 30 °C. To determine if the bait strain could trans-activate the *HIS3* reporter gene, growth of both the bait strain and bait strain + pYESTrp2 were assayed in the absence of histidine. Individual transformants of bait construct and bait construct + pYESTrp2 were streaked for isolated colonies on YC-W Z300 and YC- WHUK Z300, respectively. Transformants containing pHybLexZeo-Lamin was used as negative control for trans-activation of *HIS3* gene.

Use of 3-AT to Reduce Non-specific Trans-activation

The effect of background trans-activation of *HIS3* by LexA-bait fusion was lowered by inclusion of 3-amino-1,2,4-triazole (3-AT) in the growth medium. 3-AT is a competitive inhibitor of imidazole-glycerolphosphate dehydrate, the product of the *HIS3* gene (131). Non-specific (weak) interactions will produce some dehydratase which is subsequently inhibited by 3-AT, resulting in no growth on histidine-deficient medium. Strong interactions

Figure 13. Diagram of the pYESTrp2 vector. During the yeast two hybrid screen, the pYESTrp2 vector was used to express the cDNA library as a fusion with the B42 activation domain (B42 AD). The B42 AD + cDNA fusions expressed from the pYESTrp2 vector posses a V5 epitope tag. The pYESTrp2 vector contains an antibiotic resistance marker (Amp_r) for selection in E. coli and a nutritional marker (Trp2) for selection in yeast [Adapted from (128)].



V5 epitope

```
541  AATATTAAGC TCACC ATG GGT AAG CCT ATC CCT AAC CCT CTC CTC GGT CTC
                      Met Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu

                                                    SV40 NLS
592  GAT TCT ACA CAA GCT ATG GGT GCT CCT CCA AAA AAG AAG AGA AAG GTA GCT
     Asp Ser Thr Gln Ala Met Gly Ala Pro Pro Lys Lys Lys Arg Lys Val Ala

643  GGT ATC AAT AAA GAT ATC GAG GAG TGC AAT GCC ATC ATT GAG CAG TTT ATC
     Gly Ile Asn Lys Asp Ile Glu Glu Cys Asn Ala Ile Ile Glu Gln Phe Ile

694  GAC TAC CTG CGC ACC GGA CAG GAG ATG CCG ATG GAA ATG GCG GAT CAG GCG
     Asp Tyr Leu Arg Thr Gly Gln Glu Met Pro Met Glu Met Ala Asp Gln Ala

                    B42 activation domain
745  ATT AAC GTG GTG CCG GGC ATG ACG CCG AAA ACC ATT CTT CAC GCC GGG CCG
     Ile Asn Val Val Pro Gly Met Thr Pro Lys Thr Ile Leu His Ala Gly Pro

796  CCG ATC CAG CCT GAC TGG CTG AAA TCG AAT GGT TTT CAT GAA ATT GAA GCG
     Pro Ile Gln Pro Asp Trp Leu Lys Ser Asn Gly Phe His Glu Ile Glu Ala

                                                          Hind III
847  GAT GTT AAC GAT ACC AGC CTC TTG CTG AGT GGA GAT GCC TCC AAG CTT GGT
     Asp Val Asn Asp Thr Ser Leu Leu Leu Ser Gly Asp Ala Ser Lys Leu Gly

     Kpn I      Sac I  BamH I                          BstX I*  EcoR I
898  ACC GAG CTC GGA TCC ACT AGT AAC GGC CGC CAG TGT GCT GGA ATT CTG CAG
     Thr Glu Leu Gly Ser Thr Ser Asn Gly Arg Gln Cys Ala Gly Ile Leu Gln

                    BstX I*  Not I     Xho I          Sph I
949  ATA TCC ATC ACA CTG GCG GCC GCT CGA GGC ATG CAT CTA GAG GGC CGC ATC
     Ile Ser Ile Thr Leu Ala Ala Ala Arg Gly Met His Leu Glu Gly Arg Ile

1000 ATG TAA TTAGTTA TGTCACGCTT ACATTCACGC CCTCCCCCCA
     Met ***
```

will produce more dehydratase resulting in histidine prototrophy (131). The lowest possible concentration of 3-AT that prevented growth of L40 containing bait construct was determined by plating transformants on YPAD Z300 containing various concentrations of 3-AT including 0.5, 1, 2, 5, and 10 mM.

Hybrid Hunter™ human lung cDNA library

Vector system and the number of independent clones

The cDNA library was cloned into pYESTrp2 vector and was pre-transformed into *E. coli* strain Top10f' (Figure 13). The number of independent clones was determined by Invitrogen™ prior to amplification and was guaranteed to be at least $1x\ 10^6$ cfu/ml (colony forming units/ml).

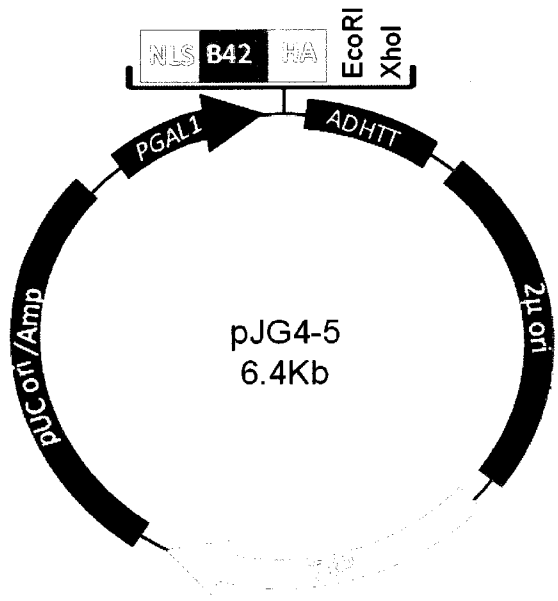
DupLEX-A™ Human Fetal Brain cDNA library

Vector system and the number of independent clones

The human fetal brain cDNA library was cloned into pJG4-5 vector and was pre-transformed into *E. coli* strain Top10f' (**Figure 14**). The number of independent clones was determined by OriGene™ prior to amplification and was guaranteed to be at least 3.5×10^6 cfu/ml.

Determining Library Titer

A small aliquot of Top10f' containing library was thawed on ice. Serial dilutions of the library (10^{-4}, 10^{-6}, 10^{-8}, and 10^{-10}) were prepared using LB medium. An aliquot of 10 µl from each of the 10^{-6}, 10^{-8}, and 10^{-10} dilutions were plated on to LB plates containing 100 µg/ml ampicillin. The plates were incubated over nights and the number of colonies was counted to determine titer of colony forming units/ml (cfu/ml).



The B42 domain and included elements in pJG4-5:

ATG GGT GCT CCT CCA AAA AAG AAG AGA AAG GTA GCT GGT ATC AAT AAA GAT ATC 54
M G A P P K K K R K V A G I N K D I 18
NLS

GAG GAG TGC AAT GCC ATC ATT GAG CAG TTT ATC GAC TAC CTG CGC ACC GGA CAG 108
E E C N A I I E Q F I D Y L R T G Q 36

GAG ATG CCG ATG GAA ATG GCG GAT CAG GCG ATT AAC GTG GTG CCG GGC ATG ACG 162
E M P M E M A D Q A I N V V P G M T 54

CCG AAA ACC ATT CTT CAC GCC GGG CCG CCG ATC CAG CCT GAC TGG CTG AAA TCG 216
P K T I L H A G P P I Q P D W L K S 72

AAT GGT TTT CAT GAA ATT GAA GCG GAT GTT AAC GAT ACC AGC CTC TTG CTG AGT 270
N G F H E I E A D V N D T S L L L S 90

GGA GAT GCC TCC TAC CCT TAT GAT GTG CCA GAT TAT GCC TCT CCC GAA TTC GGC 324
G D A S Y P Y D V P D Y A S P E F G 108
HA tag EcoR I

CGA CTC GAG AAG CTT TGG ACT TCT TCG CCA GAG GTT TGG TCA AGT CTC CAA --
R L E
Xho I

Figure 14. Diagram of the pJG4-5 vector. During the yeast two hybrid screens, the pJG4-5 vector was used to express the cDNA library as a fusion of the B42 activation domain (B42 AD). The B42 AD + cDNA fusions expressed from the pJG4-5 vector possesses an HA epitope tag. The pJG4-5 vector contains an antibiotic resistance marker (Ampr) for selection in *E. coli* and a nutritional marker (Trp2) for selection in yeast [Adapted from (128)].

Amplication of cDNA Library

The library was amplified to ensure representation of every clone in the cDNA library. To ensure even amplification of cDNA library, the Top10f' containing library was grown and amplified on LB (100 μg/ml ampicillin) plates instead of in liquid media. The library was plated directly at an appropriate density so that the resulting colonies would be nearly confluent (~30,000 cfu per 150 mm plate). 30 μl of Top10f' library culture was plated onto 100 x 150 mm plates to obtain 3 times the number of independent clones in the library. The plates were incubated at 37 °C until near confluent growth was achieved (24 hours). Two ml of LB (100 μg/ml ampicillin) was added to each plate and bacteria colonies were scraped into suspension using sterile spreader. Cell suspensions from all plates were pooled into a sterile centrifuge bottle.

Mega preparation of cDNA library

The cells were harvested via centrifugation at 6000 g for 10 min at room temperature. The pellet was resuspended in 250 ml of ice cold P1 buffer (50 mM Tris-Cl, 10 mM EDTA, 0.1 mg/ml RNAse A pH 8.0). Next, 250 ml of P2 buffer was added to the suspension (100 mM NaOH, 1% w/v SDS). The suspension was inverted 10 times and left on ice for 10 minutes. Next, 300 ml of P3 buffer (3.0 M Potassium acetate, pH 5.5 adjusted with glacial acetic acid) was added to the suspension and inverted 10 times. The suspension was left on ice for 20 minutes and coagulated chunks were broken apart by glass pipette. The suspension was centrifuged at 9000 g for 20 minutes. The supernatant was removed and filtered through several layers of cheese cloth to filter out floating pellet particles. 0.7 volumes of isopropanol was added to the supernatant and centrifuged at 9000 g for 20

minutes. The pellet was rinsed with 70% ethanol and re-centrifuged at 9000 rpm for 20 minutes. The remaining supernatant was discarded and the DNA pellet was dried overnight.

CsCl Gradient DNA ultracentrifugation

The DNA pellet was dissolved in 22.4 ml of 1 x TE (10 mM Tris,1 mM EDTA, pH 7.5). The liquid mixture was then transferred to 50 ml Falcon tube to which 28.8 g of CsCl and 800 µl of 2 mg/ml of EtBr were added. The mixture was vortexed and inverted until all particles were completely dissolved. The mixture was then poured to fill 8 Quick Seal ultra-centrifuge tubes (3.9 ml capacity). The tubes were then balanced within +/- 0.05 g and heat sealed. The balanced tubes were placed into TLA 100.3 rotor and locked to spindle. The rotor was then ultra-centrifuged at 80k rpm, 22°C, for at least 16 hours (Beckman TL-100 Ultracentrifuge). The result was a distinct EtBr band (containing plasmid DNA), which was extracted from each tube with an 18-guage needle and syringe. All EtBr extracts were pooled together and plasmid DNA was extracted from EtBr with 5-6 changes of NaCl saturated isopropanol wash (mixed together, spun down and recovered in supernatant). The volume was then dialyzed using dialysis tubing against 1 x TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) 3 times, 1 hour each at 4 °C. The plasmid DNA was then verified by digesting an aliquot of library DNA with Hind III and XhoI, and quantified using gel electrophoresis.

Preparation of sheared salmon sperm DNA

Sheared salmon sperm DNA (SSDNA) is used as a carrier in the transformation process; therefore it is necessary that it is prepared in high quality to ensure high transformation efficiency. 100 mg of SSDNA was added to 10 ml of TE buffer (10 mM

Tris-HCl pH 8.0, 0.5 mM EDTA pH 8.0), and was boiled for 20 minutes with occasionally shaking/vortex. After the SSDNA was completely dissolved, it was sheared by passing through a 21-gauge needle 12 times using a 25 ml syringe. The solution was then boiled again for 10 minutes. The sheared SSDNA was dispensed to 0.5 ml aliquots in 1.5 ml tubes and stored at -20°C.

Interactor Hunt – Large Scale Yeast Transformation of Bait with cDNA Library

Preparing Bait Strain for Transformation

Several fresh colonies of L40 bait strain (Artemis-pHyblexZeo or PARP-1-pHyblexZeo) was used to inoculate 10 ml YPADZ300. The culture was grown overnight at 30°C, shaking at 230-240 rpm. Another overnight culture was set up in 100 ml of YPADZ300 (in a sterile 1 L flask), using 0.5 ml of 10 ml culture to create an initial bait strain culture with an OD_{600} of 0.03. This culture was incubated for 16-17 hours at 30°C, shaking at 230-240 rpm. This 100 ml culture was then used to inoculate a 1L culture of YPAD Z300 (in sterile 2L flask) to an OD_{600} of approximately 0.3. The culture was grown at 30°C, shaking at 230-240 rpm for 3 hours. The cells were pelleted via centrifugation at 6000 g for 15 minutes at room temperature. The pellet was then washed with 500 ml of sterile 1X TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5). The cells were resuspended in 20 ml of 1X lithium acetate/0.5X TE (100 mM lithium acetate, 5 mM Tris-HCl, pH 7.5) and transferred to a sterile 1 L flask.

Transforming the Bait Strain with Library DNA

1 ml of 10 mg/ml denatured SSDNA and 500 µg library DNA were mixed together and added to mixture of cell suspension in the 1 L flask, along with 140 ml of sterilized 1X

lithium acetate/40% PEG-3350/TE (100 mM lithium acetate, 40% PEG-3350, 10 mM Tris-HCl, pH 7.5). The suspension was gently swirled while incubating in a 30°C water bath for 30 minutes. 17.6 ml of DMSO was added to the suspension and swirled to mix, the cells were then heat shocked in a 42°C water bath for 6 minutes with occasional swirling. The cell suspension as then immediately diluted with 400 ml of YPAD (room temperature) and swirled to cool. The cells were pelleted at 6000 g for 10 minutes (room temperature), washed with 500 ml of YPAD and pelleted again at 6000 g for 10 minutes (room temperature). The cells were resuspended in 1L of YPAD (in a sterile 2 L flask) and incubated at 30°C shaking at 230-240 rmp for 1 hour to allow the cells to recover and allow the Zeocin gene to be expressed.

Determining Primary Transformation Efficiency

1% of the final yeast culture after 1 hour recovery (1 ml from 1 L yeast culture) was plated to determine the primary transformation efficiency. The 1 ml culture sample was pelleted and resuspended in YC-WU Z300 to determine the number of double transformants. 10 µl and 1 µl aliquots were plated ($1/10^5$ and $1/10^6$ diluted with YC-WUZ300) on YC-WU Z300 plates.

Plating Large Scale Screen

The remainder of the 1 L yeast culture was pelleted via centrifugation at 6000 g for 15 minutes at room temperature. The cells were then washed with 500 ml of YC-WU media, the cells were then pelleted once again at 6000 g for 15 minutes at room temperature and resuspended in 1 L of YC-WU Z300 (pre-warmed to 30°C). The culture was then inoculated at 30°C for 16 hours overnight shaking at 230-240 rpm. The cells were pelleted at 6000 g

for 15 minutes (at room temperature) and washed with 500 ml of YC-WHUK media and resuspended in a final volume of 10 ml YC-WHUK Z300. Dilutions of $1/10^5$ and $1/10^6$ (diluted with YC-WU Z300) were plated on YC-WU Z300 plates to determine the number of doublings compared to the primary transformation efficiency and to determine the number of His^+ colonies that should be plated to cover the number of primary transformants. One hundred YC-WHUK Z300 (150 mm) plates were plated using 25 µl and 50 µl of the final 10 ml yeast suspension (50 plates each). Plates were incubated for 15-20 days at 30° C.

β-galactosidase Screen Assay

The His^+ colonies that grew on the YC-WHUK Z300 plates were patched onto two YC-WHUKZ300 plates in a grid like manner (Appendix IV): one as a master plate and one for testing β-galactosidase activity. The plates were grown for 2 days at 30°C. A filter lift assay protocol was used to detect β-galactosidase activity (132). Dry nitrocellulose paper was placed on top of plate containing His^+ colonies; pressure was applied on top to ensure each colony adhered to the membrane. The nitrocellulose membrane was then carefully peeled using two forceps and submerged into liquid nitrogen (contained in styrofoam container) colony side up. After 30 seconds of floating, the fragile membrane was carefully removed and placed in the lid of a 150 mm Petri dish containing one #1 Whatman filter circle (sized to fit inside 150 mm plates). The membrane was dried on top of the filter paper at room temperature for 2 minutes. Another Petri dish was prepared for the reaction: one #1 Whatman filter paper was placed inside the bottom half of a 150 mm Petri dish, to which 5 ml of Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$-$7H_2O$, pH 7.0) plus 100 µl of 50 mg/ml of X-gal was added slowly with even distribution. The dry nitrocellulose membrane containing the lysed colonies was then transferred on to Petri dish

contained the Whatman paper moistened with Z buffer/X-gal solution with colonies facing up. The dish was covered with aluminum foil and placed at 30°C. Strong interactions yielded blue color within 2 hours, other colonies were incubated for 6-8 hours until a blue color was observed.

Retrieving Putative Interactors

Isolating Plasmid DNA from Yeast

Yeast plasmid DNA was extracted from double positives (His^+LacZ^+) and transformed into *E. coli* DH5α. 10 ml of YC-W media (50 ml Falcon tube with holes punctured in lid for aeration) was inoculated with a single double positive colony and incubated over night at 30 °C shaking at 230-240 rpm. The cells were pelleted via centrifugation at 2500 rpm (table top microcentrifuge) for 10 minutes at room temperature and frozen at -80°C for 10 minutes. The pellets were then thawed in 300 µl of yeast lysis buffer. The yeast lysate suspension was then transferred to a fresh 1.5 ml round bottom microtube, to which was added 150 µl of phenol/chloroform and 150 µl of acid washed glass beads (Sigma G-8772, 425-600 microns). The microtube lid was snapped shut securely (ensuring no glass beads were adhered to edges) and was vortexed vigorously for 3-5 minutes to ensure efficient lysis. The suspension was then centrifuged at 13 000 rpm (table top microcentrifuge) for 15 minutes. The aqueous phase was transferred to a fresh microtube and clarified by centrifuging at 13 000 rpm (table top microcentrigue) for an additional 5 minutes. The supernatant was then transferred to a fresh microtube.

Ethanol Precipitation of DNA

DNA was precipitated from the clarified supernatant using ethanol precipitation. The supernatant (~300 µl) was transferred to a fresh 1.5 ml microtube. One tenth volume of 3M sodium acetate buffer (3 M sodium acetate buffer, pH 5.0) was added to equalize ion concentrations. Two and half volumes of ice cold 100% ethanol was added to the sample and was left to sit in the -20°C freezer for one hour. The sample was then centrifuged for 15 minutes at 13 000 rpm (table top microcentrifuge) at 4°C. The supernatant was carefully removed and discarded. One volume of ice cold 70% ethanol was then added and used to gently wash the small colorless DNA pellet. The sample was then centrifuged for an additional 10 minutes at 13 000 rpm (table top microcenrifuge) at 4°C. The supernatant was removed and the pellet was dried by leaving the microtube overnight at room temperature with lid open. The pellet was then re-suspended in the desired volume of water.

Characterization of Plasmid containing Putative Interactors

To isolate the proper plasmid cDNA from the precipitated genomic DNA and cellular debris, 5 µl of the DNA suspension was transformed into competent *E. coli* (DH5α) as previously described. The entire transformation was plated on LB plates containing 100 µl ampicillin. Plasmid DNA was isolated from ampicillin-resistant transformants and analyzed by restriction analysis and agarose gel electrophoresis. Empty pYESTrp2/pJG4-5 vector was used as negative control.

Testing Putative Interactors

The retrieved plasmid DNA from double positive colonies (His^+LacZ^+) were then re-transformed into L40 strain alone and with empty bait vector to ensure that the activation of reporter construct is not a result of prey auto-activation or interaction of prey with bait

vector. The putative interactors were transformed into L40 (with and without pHybLex/Zeo) according to the small scale transformation protocol (Appendix III). L40 transformed with prey plasmid alone was plated on YC-WHUK plates and YC-W (positive control for transformation), and L40 transformed with prey and pHybLex/Zeo was plated on YC-WHUK Z300 plates and YC-WU Z300 (positive control for transformation). The plates were incubated for 2-3 days at 30 °C.

Sequencing Putative Interactors

A "well behaved" prey should not activate reporter gene constructs alone or in the presence of pHyblex/Zeo. Unique prey clones that were only capable of activating reporter genes in presence of pHyblex/Zeo-Artemis were sent for sequencing analysis. The primers used for sequencing pYESTrp2 and pJG4-5 cDNAs are listed in Apendix II. The sequencing was performed by StemCore Laboratories (OHRI; Ottawa, ON). The sequence was analyzed using NCBI BLAST to identify the nature of putative interacting proteins (133).

Matchmaker™ Yeast Two-Hybrid Gal-4 based system – Interaction Mating

The Matchmaker™ two-hybrid Gal-4 based system was used to screen PARP-1 and Artemis using Matchmaker™ Pre-transformed Human Bone Marrow cDNA library (Clontech Laboratories, Inc.; Mountain View, CA). There are two yeast strains used in this system, each an opposite mating types (Mat-*a* and Mat-α). The bait plasmid is expressed by Mat-*a* (AH109), while library plasmid is expressed by Mat-α (Y187). The bait and prey hybrids are brought together via yeast mating, not by transformation.

Yeast Host strain genotype

The following table summarizes the genotypic and phenotypic characteristics of each yeast strain used in the interaction mating protocol. In response to a positive interaction between bait and prey hybrid the yeast Strain AH109 expresses four integrated reporter genes under the control of distinct Gal4-responsive upstream activating sequences (UASs) and TATA boxes (**Figure 15**).

Table 2. Complete genotype and phenotype of yeast host strains AH109 and Y187. Yeast strains AH109 and Y187 are opposite mating types and contain compatible reporter genes and transformation markers (134).

Strain	Genotype	Phenotype	Reporter Genes	Transformation markers
AH109	MAT*a*, *trp*1-901, *leu*2-3, 112, *ura*3-52, *his*3-200, gal4Δ, gal80Δ, LYS2 : : $GAL1_{UAS}$-$GAL1_{TATA}$-*HIS3*, $GAL2_{UAS}$-$GAL2_{TATA}$-*ADE2* *URA*3 : : $MEL1_{UAS}$-$MEL1_{TATA}$-*LacZMEL1*	- Ade - His - Leu - Trp + Ura	*HIS3* *ADE2* *lacZ* *MEL1*	*Leu2* *Trp1*
Y187	MATα, *ura*3-52, *his*3-200, *ade*2-101, *trp*1-901, *leu*2-3, 112, gal4Δ, gal80Δ, met–, *URA3* : : $GAL1_{UAS}$-$GAL1_{TATA}$-*LacZ MEL1*	- Ade - His - Leu - Trp + Ura	*lacZ* *MEL1*	*Leu2* *Trp1*

Construction of GAL4 DNA-BD + bait Fusion Vectors

The bait vector is this system is the pGBKT7 vector (**Figure 16**). Artemis and PARP-1 constructs were inserted into the multiple cloning site (MCS) of the pGBKT7 vector. In yeast, expression from the constitutive ADH1 promoter (PADH1) results in a

protein consisting of an N-terminal GAL4 DNA-BD followed by a c-Myc epitope tag (Δ) and the MADAM protein. The pGBKT7 vector contains an antibiotic resistance marker (Kan^r) for selection in *E. coli* and a nutritional marker (TRP1) for selection in yeast. The T7 promotor is used for *in vitro* transcription and translation of the epitope tagged fusion protein.

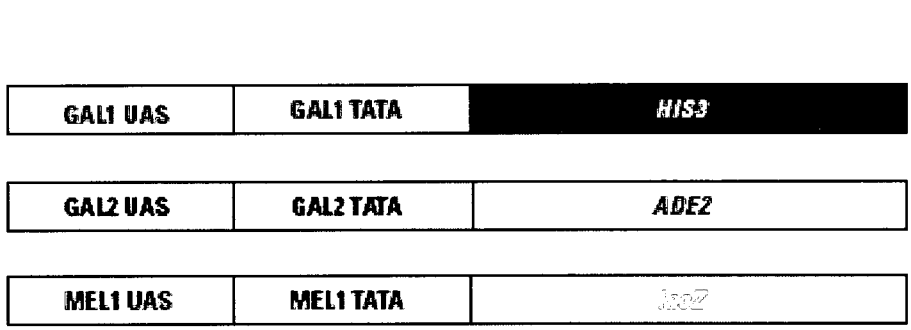
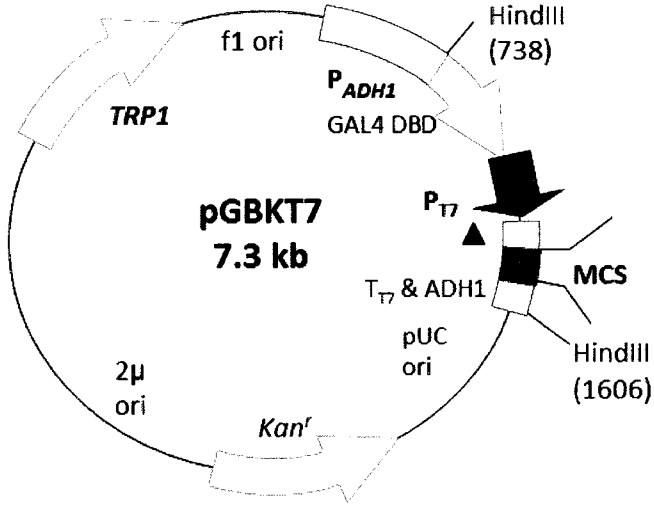

GAL1 UAS
GAL1 TATA
HIS3
GAL2 UAS
GAL2 TATA
ADE2
MEL1 UAS
MEL1 TATA
lacZ
MEL1 UAS
MEL1 TATA
MEL1

Figure 15. Reporter gene constructs in yeast strains MAT-a AH109. In AH109, the *HIS3*, *ADE2*, and *MEL1/LacZ* reporter genes are under the control of three completely heterologous Gal4-responsive UAS and promoter elements: *GAL1*, *GAL2*, and *MEL1*, respectively. The protein-binding sites within the UASs are different, although each is related to the 17-mer consensus sequence recognized by Gal4 (134)



▲ c-Myc epitope tag

MATCHMAKER 5' DNA-BD Vector
Insert Screening Amplimer

1155 GAL4 DNA-Binding Domain

TCA TCG GAA GAG AGT AGT AAC AAA GGT CAA AGA CAG TTG ACT GTA TCG CCG GAA TTT

a.a. 147

1212 T7 Sequencing Primer

T7 Promoter

c-Myc Epitope Tag

GTA ATA CGA CTC ACT ATA GGG CGA GCC GCC ATC ATG GAG GAG CAG AAG CTG ATC TCA GAG GAG GAC CTG

START in vitro

1281

CAT ATG GCC ATG GAG GCC GAA TTC CCG GGG ATC CGT CGA CCT GCA GCG GCC GCA TAACTAGCATAACCCC

Nde I *Nco* I *Sfi* I *Eco*R I *Sma* I/ *Xma* I *Bam*H I *Sal* I *Pst* I STOP (orf 1) STOP (orf 2)

1342 T7 Terminator

TTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCGCGCTTGCAGCCAAGCTAATTCCGGGCGAATTTCTTATGATTT

STOP (orf 3)

1430

ATGATTTTTATTATTAAATAAGTTATAAAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAA

3' DNA-BD Sequencing Primer

MATCHMAKER 3' DNA-BD Vector
Insert Screening Amplimer

Figure 16. Diagram of the pGBKT7 vector. During the yeast two-hybrid screen, the pGBKT7 vector was used to express the bait protein (N term of Artemis and C terminus of PARP-1) fused to amino acids 1-147 of the GAL4 DNA binding domain (GAL4 DNA-BD) [adapted from (134)].

PCR reactions

N terminus, C terminus, and Full length Artemis

PCR reactions were carried out with primers designed to generate NdeI and PstI ended inserts to allow in frame insertion of C-terminus, N-terminus and full fragment regions of Artemis into pGBKT7 (**Figure 17**). Primer sequences are listed in Appendix II. The N-terminus region of Artemis (37-1206 bp) was amplified using primers A and B. The C-terminus region of Artemis (1167-2115 bp) was amplified using primers C and D. Full fragment Artemis (37-2115 bp) was amplified using primers A and D.

C terminus PARP-1

The first 250 amino acids of the N terminus of PARP-1 (encoding the DBD) were omitted from the bait construct to abolish auto-activation of reporter genes. PCR reactions were carried out with primers designed to generate ends containing NdeI and BamHI restriction sites to allow in frame insertion of C-terminus PARP-1 into pGBKT7. The C-terminus of PARP-1 (909-3196 bp) was amplified using the forward primer I and the reverse primer II. PCR reactions were prepared as previously described.

Cloning and Plasmid Preparations

pGBKT7-NArtemis, pGBKT7-CArtemis, pGBKT7-FArtemis

Following PCR amplification, the products were quantified and purified from a 1.0% agarose gel: *pGBKT7*-NArtemis, *pGBKT7*-CArtemis, and *pGBKT7*-FArtemis generated an 1169 bp, 948 *pGBKT7* bp, and 2078 bp insert, respectively. The *pGBKT7* plasmid and PCR inserts were linearized with NdeI and PstI to generate compatible ends. Vector and insert

C-terminus PARP-1

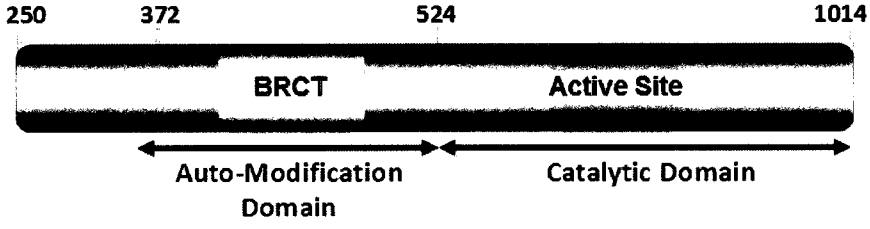


Figure 17. The C terminus region of PARP-1, that was amplified via PCR to create bait-GAL-4 DBD fusion in pGBKT7. The first 250 amino acids of the N terminus of PARP-1 (encoding the DBD) were omitted from the bait construct to abolish auto-activation of reporter genes.

digestion was set up individually; each in a 50 µl reaction containing 1 µg of pGBKT7 or 500 ng of Artemis PCR insert, 1 µl of PstI (10 units), 1 µl of NdeI (10 units), 5 µl of 10X BSA, and 5 µl NEBuffer 2. The reaction was incubated at 37°C for 2 hours. The pGBKT7 plasmid and PCR inserts were then quantified and purified from a 1% agarose gel into 30 µl of EB buffer. A 50 µl blunt-blunt ligation reaction was set up using 3:1 insert to vector ratio using 100 ng pGBKT7 and 300 ng of insert (C-terminus, N-terminus, and full length Artemis separately), 0.4 µl of T4 DNA Ligase (2000 units/µl), 2 µl of T4 DNA Ligase reaction Buffer into a 20 µl reaction. A 20 µl control reaction was also set up lacking the insert; containing 100 ng of pGBKT7, 2 µl of T4 DNA Ligase reaction buffer, 0.4 µl of T4 DNA Ligase (2000 units/µl). The reactions were incubated in a 15 °C water bath overnight.

pGBKT7-cPARP-1

Following PCR amplification, the cPARP-1 product was quantified and purified from a 1.0% agarose gel; generating a 2287 bp insert. The pGBKT7 plasmid was linearized with NdeI and BamHI to generate compatible ends with Artemis inserts. PCR generated cPARP-1 and pGBKT7 were digested with NdeI and BamHI sequentially. A 50 µl blunt-blunt ligation reaction was set up using 3:1 insert to vector ratio using 100 ng of pGBKT7 and 300 ng of insert, 0.4 µl of T4 DNA Ligase (2000 units/µl), 2 µl of T4 DNA Ligase reaction Buffer into a 20 µl reaction. A 20 µl control reaction was also set up lacking the insert; containing 100 ng of pGBKT7, 2 µl of T4 DNA Ligase reaction buffer, 0.4 µl of T4 DNA Ligase (2000 units/µl). The reactions were incubated in a 15 °C water bath overnight.

E. coli Transformation

The ligation products were transformed into competent *E. coli* DH5α cells as previously described, and plated on LB plates containing 50 μg/ml of KAN and incubated at 37°C overnight. Several single colonies were selected, and each was used to inoculate 5 ml of liquid LB media containing 50 μg/ml of ampicillin. The cultures were incubated overnight at 37 °C with shaking. Plasmid DNA was isolated and characterized by restriction digest using enzymes that were specific for each insert to identify the clones containing proper insert. The clone that contained appropriate size bands was sent off for sequencing using forward primer (5'-TAATACGACTCACTATAGGG-3').

Yeast Transformation with Bait Construct (Creating Bait Strain)

The pGBKT7-N, -C, -FArtemis and pGBKT7-cPARP-1 bait constructs were transformed into the yeast AH109 strain using the small scale yeast transformation protocol (Appendix III). The transformations were plated on YC-W plates and incubated for 2-3 days at 30 °C until large (2-3 mm diameter) yeast colonies appeared.

Test for Expression of Bait protein

A single fresh colony from each of AH109 [pGBKT7-CArtemis], AH109 [pGBKT7-NArtemis] , AH109 [pGBKT7-FArtemis] and AH109 [pGBKT7-cPARP-1] were each used to inoculate 10 ml overnight culture of YC-W in 50 ml Falcon tubes. AH109 was grown in YPAD as a negative control for expression. Lysates were prepared from transformants capable of growing in TRP deficient medium for immunoblot analysis to check for expression of the bait as previously described.

Testing Bait plasmid for non-specific Activation

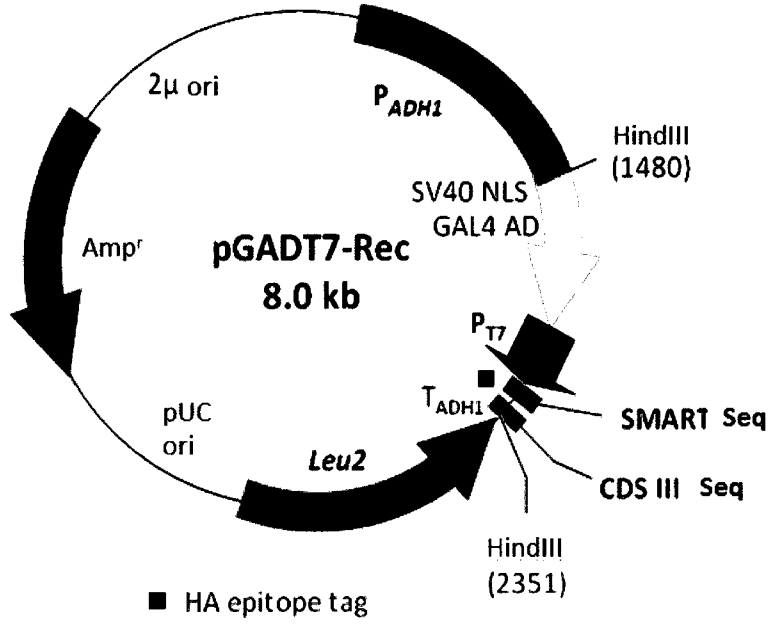
Each bait strain was transformed with the empty library vector pGADT7 (Appendix V) using the small scale transformation protocol (Appendix III). The transformation was plated onto YC-LW plates and incubated for 2-3 days at 30°C. To determine if the bait strain could trans-activate the *HIS3* reporter gene, growth of both the bait strain and bait strain + pGADT7 were assayed in the absence of histidine and adenine. Individual transformants of bait construct and bait construct + pGADT7 were streaked for isolated colonies on YC-LW and YC- WHAL, respectively. The lowest possible concentration of 3-AT that prevented growth of AH109 containing bait construct was determined.

Matchmaker™ Pre-transformed Human Bone Marrow cDNA Library

The matchmaker™ Pre-transformed human bone marrow cDNA library is a high-complexity cDNA library cloned into a yeast GAL-4 AD pGADT7-Rec vector and is pre-transformed by the manufacturer into host strain *MAT*αY187. The pGADT7-Rec vector is designed for recombination-mediated cloning; it is used to construct cDNA libraries *in vivo* in yeast (**Figure 18**).

In vivo cloning of Pre-transformed cDNA library using SMART™ Technology

The *in vivo* cloning of Matchmaker cDNA library involves two main steps: cDNA synthesis and yeast transformation. The following is an outline of the manufacturer's protocol (134). Messenger RNA transcripts were copied into cDNA using SMART™ (Switching Mechanism at 5' end of RNA Transcript) technology. In the first-strand cDNA synthesis step, MMLV (Moloney Murine Leukemia Virus) Reverse Transcriptase (RT) is used to transcribe RNA into DNA. To prime RNA for cDNA synthesis, the modified



SMART III^{TM} terminus

1858
MATCHMAKER 5' AD LD-Insert Screening Amplimer
GAL4 Activation Domain
a.a. 881
T7 Sequencing Primer
T7 Promoter
CTA TTC GAT GAT GAA GAT ACC CCA CCA AAC CCA AAA AAA GAG ATC TTT AAT ACG ACT

1915
START *in vitro*
HA Epitope Tag
CAC TAT AGG GCG AGC GCC GCC ATG GAG TAC CCA TAC GAC GTA CCA GAT TAC GCT

1969
SMART III Primer
CAT ATG GCC ATG GAG GCC AGT GAA TTC CAC CCA AGC AGT GGT ATC AAC GCA GAG TGG

2026
CCA TTA TGG CCC

CDS III terminus

2038
CDS III Primer
GGG AAA AAA CAT GTC GGC CGC CTC GGC CTC TAG AGG GTG GGC ATC GAT ACG GGA TCC

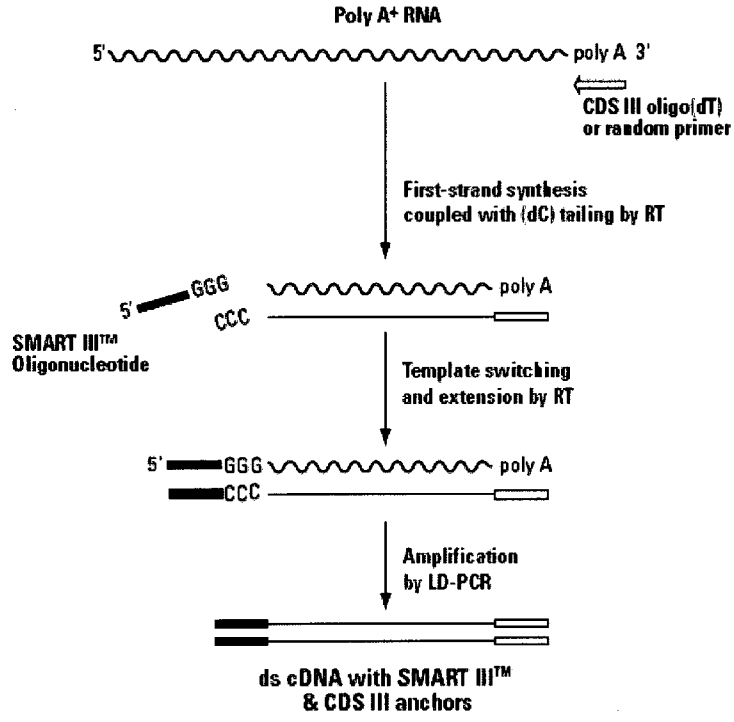
2095
STOP
ATC GAG CTC GAG CTG CAG ATG AAT CGT AGA TAC TGA AAAACCCCGCAAGTTCACTTC
MATCHMAKER 3' AD LD-Insert Screening Amplimer

2152
AACTGTGCATCGTGCACCATCT
3' AD Sequencing Primer

Figure 18. Diagram of the pGADT7-Rec vector. pGADT7-Rec was used to express proteins in the human bone marrow cDNA library as a fusion with the GAL4 activation domain (GAL4 AD). This vector is engineered for HRR-mediated cloning. The GAL4 AD+ cDNA fusions expressed from the pGADT7-Rec vector possess a hemagglutinin (HA) epitope tag (■). The pGADT7-Rec vector contains an antibiotic resistance marker (Amp^r) for selection in *E. coli* and a nutritional marker (LEU2) for selection in yeast [Adapted from (134)].

oligo(dT) (CDS III primer) is used to prime the RNA for cDNA synthesis. When the RT encounters the 5'-end of the RNA, the terminal transferase activity of the enzyme will add a deoxycytidine (dC) stretch to the 3'-end of the single-stranded cDNA (ss cDNA). The SMART III™ oligonucleotide contains an oligo (G) sequence at its 3'-end that will base-pair with the dC stretch on the ss cDNA. RT will switch templates and continue replicating to the end of the SMART III™ oligonucleotide. Only ss cDNA with the SMART III™ oligonucleotide sequence at the 5'-end can serve as a template during long-distance PCR (LD-PCR). LD-PCR of the ss cDNA results in the generation of double-stranded cDNA (ds cDNA) with complete SMART III™ and CDS III anchors (**Figure 19A**). To generate fusion proteins between ds cDNA and the GAL4 AD, the SMART ds cDNA and SmaI linearized pGADT7-Rec vector are co-transformed into the yeast, the ds cDNA will recombine with the pGADT7-Rec vector to yield the complete GAL4 AD+cDNA library plasmid (**Figure 19B**). The GAL4 AD+cDNA library plasmid will express the cDNA insert as a fusion with the GAL4 AD. Recombination is possible because the SMART III and CDS III sequences that were incorporated into the cDNA are also engineered into the pGADT7-Rec vector. Yeast repair enzymes will restore the SmaI linearized pGADT7-Rec vector to its circular form by recombining the SMART™ III and CDS III homologous ends present in the cDNA and vector. Successful constitution of the pGADT7-Rec+cDNA plasmid can be determined by growth of the yeast on SD/-Leu medium. The yeast transformation procedure was performed to introduce the SmaI linearized pGADT7-Rec vector and human bone marrow ds cDNA into AH109 yeast cells that already contained the pGBKT7+MADAM (ΔSS/ΔTM) plasmid. Transformations were performed using a lithium acetate mediated method according to the Matchmaker Library Construction & Screening Kit (Clontech Laboratories, Inc.; Mountain View, CA) (134). Yeast transformants were evaluated for growth on the following synthetic

A


Poly A+ RNA
5'
poly A 3'
CDS III oligo(dT)
or random primer
First-strand synthesis
coupled with (dC) tailing by RT
5'
GGG
CCC
poly A
SMART III™
Oligonucleotide
Template switching
and extension by RT
5'
GGG
CCC
poly A
Amplification
by LD-PCR
ds cDNA with SMART III™
& CDS III anchors

B

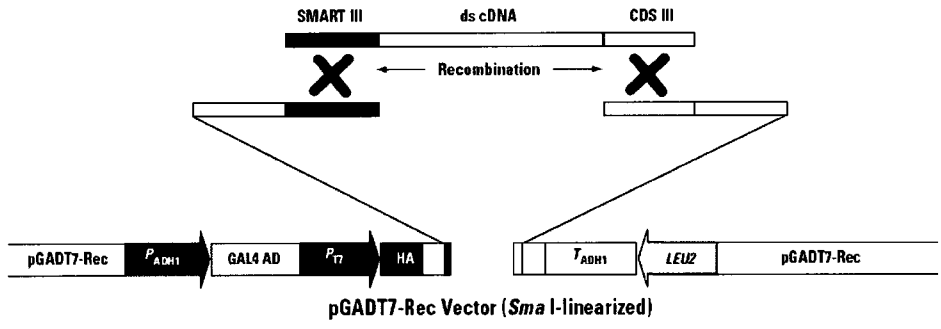

SMART III
ds cDNA
CDS III
Recombination
pGADT7-Rec
P_{ADH1}
GAL4 AD
P_{T7}
HA
T_{ADH1}
LEU2
pGADT7-Rec
pGADT7-Rec Vector (Sma I-linearized)

Figure 19. Construction of cDNA library using the SMART™ technology. **(A)** Double stranded cDNA synthesis from RNA using SMART and CDS III anchors. During first strand synthesis, RT is used to transcribe mRNA into cDNA. The modified oligo (dT) (CDS III primer) is used to prep the RNA for cDNA synthesis. **(B)** HRR of ds cDNA and the pGADT7-Rec vector. To introduce the ds cDNA into the pGADT7-Rec vector, first SMART III and CDS III sequences must be generated on the 5'- and 3'-terminal of the ds cDNA. When co-transformed into yeast, the ds cDNA will recombine with the pGADT7-Rec linearized vector, thereby restoring the circular topology of the vector [Adapted from (134)].

dropout (SD) minimal medium: SD/-Leu/-Trp (double dropout or DDO) and SD/-Leu/-Trp/-His (triple dropout or TDO).

Number of Independent Clones and Library Titer

The number of independent clones of the human bone marrow pre-transformed library was determined by Matchmaker™, and was guaranteed to be at least $8.8x10^6$. Prior to library screening, a 1 ml aliquot of library culture was thawed on ice. An aliquot of 10 µl of the culture was removed to prepare serial dilutions of the library (10^{-4}, 10^{-6}, 10^{-8}, and 10^{-10}) using YC-L medium. An aliquot of 100 µl from the 10^{-6}, 10^{-8}, and 10^{-10} dilutions were plated on to YC-L plates. The plates were incubated for 2-3 days and the number of colonies was counted to determine titer of colony forming units/ml (cfu/ml).

Control Experiments

Mating efficiency of Y187 [pGADT7-T] and AH109 [pGBKT7-p53]

Y187 and AH109 pre-transformed with plasmids encoding known interacting proteins (SV40 large T-antigen and murine p53, respectively) were included with the pre-transformed library as positive controls. Control mating procedure was carried out using the following small-scale mating protocol prior to, and in parallel with library screening. One large colony of each type (ie. AH109 [pGBKT7-53] and Y187 [pGADT7-T]) was selected from fresh stock plates (<2 weeks old). Both colonies were placed in one 1.5 ml microtube, containing 0.5 ml of 2X YPAD medium. The tube was gently vortexed and inverted until colonies were uniformly resuspended in the medium. Tubes of single, non-mating cultures were similarly prepared to determine the mating efficiency of the positive control. The cultures were incubated overnight at 30°C overnight (~24 hours) with shaking at 200 rpm. 100 µl aliquots of 10^{-2} and10^{-3} dilutions of the mating culture was plated on YC-WHAL

medium, YC-WL, YC-L, and YC-W plates. 100 µl aliquots of 10^{-2} and10^{-3} dilutions of non-mating cultures were plated on YC-W (AH109 [pGBKT7-53]) and YC-L (Y187 [pGADT7-T]). The plates were incubated at 30°C for 3-5 days to allow diploid cells to form. The number of colonies the plates was determined by counting to calculate the mating efficiency.

Testing effect of AH109 [pGBKT7-bait] on Mating Efficiency

The AH109[pGBKT7-NArtemis] and AH109[pGBKT7-cPARP-1] were mated with Y187[pGADT7-T] using the small-scale mating protocol to determine the effect of the bait on mating efficiency in comparison with control experiment prior to library screening. 100 µl aliquots of 1:100 and 1:1000 dilutions of each mating culture was plated on YC-LW, YC-L, and YC-W medium plates. 100 µl aliquots of 10^{-2} and10^{-3}dilutions of non-mating cultures were plated on YC-W (AH109 [pGBKT7-NArtemis] and AH109 [pGBKT7-CPARP-1]) and on YC-L (Y187 [pGADT7-T]). The plates were incubated at 30°C for 3-5 days to allow diploid cells to form. The number of colonies the plates was determined by counting to calculate the mating efficiency.

Two Hybrid Library Screening Using Yeast Mating

Large Scale Yeast Mating Protocol

An overnight culture of bait strain was prepared by inoculating several fresh colonies of bait strain into 50 ml of YC-W in a 500 ml flask. The culture was incubated overnight shaking at 30°C at 230-240 rpm to ensure the culture had grown to OD_{600} >0.8. The cells were pelleted via centrifugation at 1000 rpm for 10 minutes. The cells were resuspended in the residual liquid (~5 ml) and left on ice. A 1 ml aliquot of library culture (in Y187) was thawed on ice prior to use. The entire library culture was gently mixed with 5 ml bait cell culture and transferred to sterile 2 L flask. 45 ml of 2X YPAD/KAN (4% peptone, 2% yeast

extract, 4% D-glucose, 0.02% adenine, 15 μg/ml kanamycin) was added the cell suspension along with two 1 ml aliquots of 2X YPAD/KAN to rinse cells from the library tube. The culture was incubated overnight (24 hours) with gentle swirling (40 rpm). The cells from the mating culture were then pelleted via centrifugation at 1000 rpm for 10 minutes. The flask of the culture was rinsed with 50 ml of 0.5 X YPAD/KAN (1% peptone, 0.5% yeast extract, 1% D-glucose, 0.005% adenine, 15μg/ml kanamycin) and used to resuspend the culture pellet. The cells were pelleted at 1000 rpm for 10 min. The cell pellet was resuspended in 0.5 X YPAD/KAN to a total volume of 10 ml. To determine the mating efficiency of the screen, serial dilutions of 10^{-2}, 10^{-3}, and 10^{-4} were prepared and 100 μl of each dilution was plated on YC-W, YC-L, and YC-LW plates. The remainder of the screen was plated on YC-WHAL plates containing 5 mM 3-AT and incubated at 30°C for 7-10 days. The colonies that grew on YC-WHAL plates were selected on patched on two master plates as previously described. One master plate was used to detect β-galactosidase activity using the filter lift assay protocol as previously described.

Optimizing Mating Conditions to Increase Mating Efficiency

Optimal mating efficiency is essential in interaction-mating since the yeast two hybrid system requires a sufficient number of doubly-transformed cells expressing the representative fusion proteins. The number of doubly-transformed cells required for the system is dependent on the number of independent clones in the library. In general, it is necessary to obtain at least 5-6X double transformants as the number of independent clones in the cDNA library. Mating conditions were optimized by applying an approach reported by a study that developed a protocol for optimizing mating conditions prior to library screening (135). It was reported that mating efficiencies were strongly dependent on the

MATa:MATα ratio, and that an optimum ratio of 2.5:1 lead to a maximum mating efficiency. Additionally that pre-incubation on low pH media before mating was shown to increase the mating competence of yeast cells. The bait strain AH109 [pGBKT7-bait] was pooled in 2.5 fold excess of library strain Y187 [pGADT7-cDNA library]. The two haploid strains were pre-incubated in 2X YPAD medium pH 4.5 for 4.5 hours at 30°C prior to mating protocol.

Decreasing Plating Volume

Dense plating can result in growth by cells obtaining nutrients from surrounding dead cells rather than from the medium. His^+ colonies in particular grow poorly in high density plates. Dense plating can also give the illusion of growth; a lawn of growth after 1-2 days appears making it difficult to distinguish between background and a positive transformant. The volume of screen to plate is dependent on both the number of independent clones in a cDNA library and also the number of double positive transformants. Enough doubly-positive yeast should be plated to screen all the representative clones in the library. To decrease the plating volume required for the interaction-mating screen, the yeast were grown in YC-LW medium for an additional 16 hours to increase the number of diploid cells. The OD_{600} of the culture was measured before and after 16 hour incubation period to determine the approximate growth of the yeast. Calculations were made to determine the appropriate volume of yeast to plate.

Retrieving Putative Interactors

Isolating Plasmid DNA from Yeast

Yeast plasmid DNA was extracted from triple positives ($His^+Ade^+LacZ^+$) using the Y-DER® Yeast DNA Extraction Reagent Kit (Thermo Fisher Scientific Inc.; Rockford, IL),

using manufacturer's extraction protocol (136) and transformed to *E. coli* DH5α. 10 ml of YC-L media (50 ml Falcon tube with holes punctured in lid for aeration) was inoculated with a single colony of triple positive colony and incubated overnight at 30°C shaking at 230-240 rpm. The cells were pelleted via centrifugation at 10000 rpm (table top microcentrifuge) for 5 minutes at room temperature and transferred to a 1.5 ml microtube. This result was a 70-100 mg pellet. The cells were then resuspended into 800 µl of Y-PER® Reagent by inverting the microtube and gently vortexing until a homogenous mixture was established. The mixture was then incubated at 65°C for 10 minutes. 200 µl of Protein Removal Reagent was added to the to mixture and the microtube was inverted several times. The mixture was centrifuged at 13 000 rpm (table top microcentrigue) for 10 minutes. The supernatant was transferred to a new 1.5 ml microtube, and 600 µl of isopropanol was added to fill tube. The mixture was inverted several times and DNA was precipitated by centrifuging the mixture at 13 000 rpm (table top microcentrigue) for 20 min. The supernatant was removed carefully and the pellet was rinsed with 1 ml of 70% ice cold ethanol to wash off cellular debris and/or residual salts. The DNA was centrifuged once again at 13 000 rpm (table top microcentrifuge) for 1 min. The tube was inverted to dry the pellet overnight at room temperature. The DNA pellet was then resuspended in 20 µl of sterile water.

Characterization of Plasmid containing Putative Interactors

To isolate the proper plasmid cDNA from the precipitated genomic DNA and cellular debris, 5 µl of the DNA suspension was transformed into competent *E. coli* (DH5α) as previously described. The entire transformation was plated on LB plates containing 100 µl ampicillin. Plasmid DNA was isolated from ampicillin-resistant transformants and analyzed

by restriction analysis and agarose gel electrophoresis. The empty pGADT7 vector was used as negative control.

Testing Putative Interactors

The retrieved plasmid DNA from triple positive colonies ($His^+Ade^+LacZ^+$) were then re-transformed into AH109 strain alone and with empty bait vector to ensure that the activation of reporter construct is not a result of prey auto-activation or interaction of prey with bait vector. The putative interactors were transformed into AH109 (with and without pHybLexZeo) according to small scale transformation protocol (Appendix III). AH109 transformed with prey plasmid alone was plated on YC-WHAL plates and YC-L (positive control for transformation), and AH109 transformed with prey and pGBKT7 was plated on YC-WHAL plates and YC-WL (positive control for transformation). The plates were incubated for 2-3 days at 30 °C.

Sequencing Putative Interactors

A "well behaved" prey should not activate reporter gene constructs alone or with presence of pGBKT7. Unique prey clones that were only capable of activating reporter genes in presence of pGBKT7-bait were sent for sequencing analysis. The pGADT7-Rec forward and pGADT7-Rec reverse oligomers were used for sequencing (Appendix II). The sequencing was performed by StemCore Laboratories (OHRI; Ottawa,ON). The sequence was analyzed using NCBI BLAST to identify the nature of putative interacting proteins (133).

RESULTS

A series of yeast two hybrid screens were carried out to identify Artemis and PARP-1 interacting factors using two different systems. The Hybrid Hunter™ Two-Hybrid LexA based System was used to screen for Artemis interacting factors using a Hybrid Hunter™ human lung cDNA library and a DupLEX-A™ Human Fetal Brain cDNA library. The Matchmaker™ Yeast Two-Hybrid Gal-4 based system was used to screen for Artemis and PARP-1 interacting factors using a pre-transformed Matchmaker™ human bone marrow cDNA library via interactive mating. Below, the bait constructs, the characterization of bait-DBD fusions, as well as the results of the Y2H screens are presented.

Hybrid Hunter™ Two-Hybrid LexA Based System

Characterization of pHybLex/Zeo-Artemis Bait constructs

Bait Expression: pHyblexZeo-NArtemis, C-Artemis, and FArtemis

A direct assay of bait expression was determined via western blot analysis. The pHybLex/Zeo-NArtemis and -CArtemis Bait constructs were expressed in L40 at detectable levels at the appropriate size (**Figure 20**). Expression of pHybLex/Zeo-FArtemis bait construct in L40, was not confirmed (**Figure 21**). The expression of various pHybLex/Zeo-CArtemis deletion constructs in L40 were detected via western blot analysis (**Figure 22**).

Testing Bait plasmids for non-specific Transcriptional Activation

The lowest possible concentration of 3-AT included in the YPAD Z300 medium that reduced the growth of background microsatellite L40 colonies expressing bait constructs was determined to be 5 mM. The pHybLex/Zeo-NArtemis bait construct did not trans-activate any of the reporter genes in L40 (**Figure 23A**): L40 [pHybLex/Zeo-NArtemis] was not able

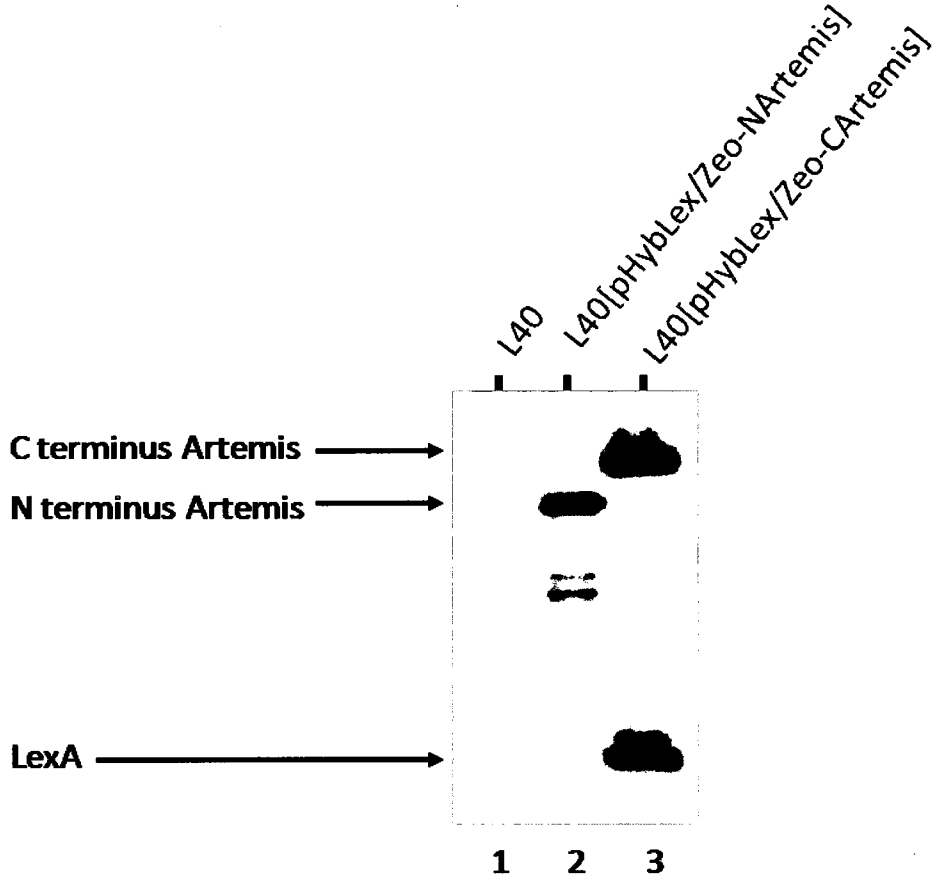

L40
L40[pHybLex/Zeo-NArtemis]
L40[pHybLex/Zeo-CArtemis]
C terminus Artemis
N terminus Artemis
LexA
1
2
3

Figure 20. Western blot analysis of L40 transformed with pHybLex/Zeo-Artemis constructs. The expression of Artemis-LexA DBD was detected with the anti-LexA antibody on a 10% acrylamide gel. Lane 1 contains the lysate of L40 untransformed (negative control) without detectable Artemis expression. Lane 2 contains the lysate of L40 transformed with pHybLex/Zeo-NArtemis construct; showing expression of N-terminus Artemis. The third lane contains the lysate of L40 transformed with pHybLex/Zeo-CArtemis showing expression of C-terminus Artemis and un-fused LexA.

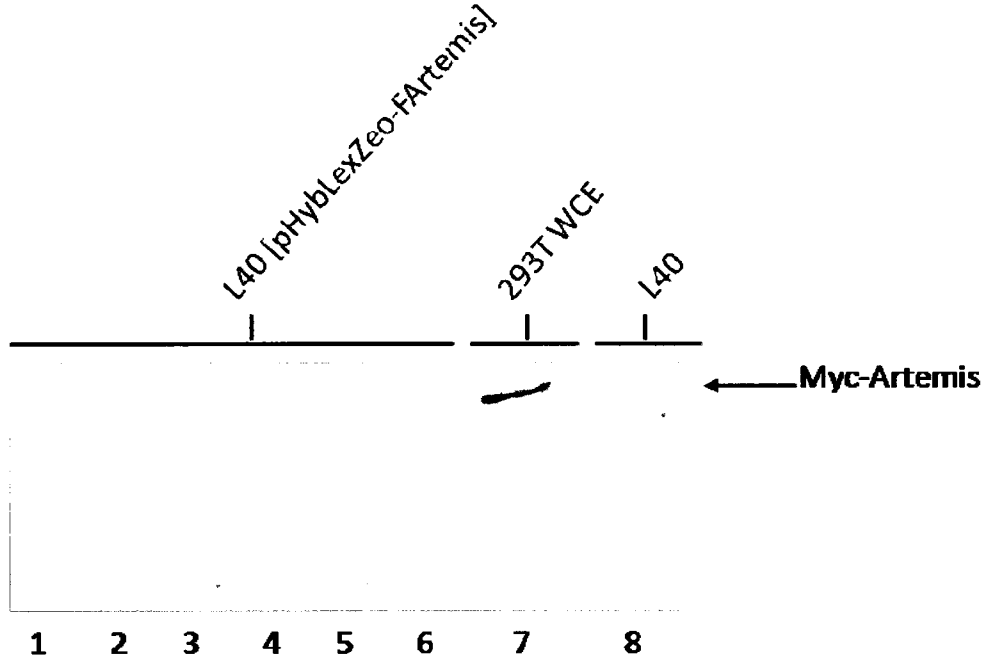

L40 [pHybLexZeo-FArtemis]
293T WCE
L40
Myc-Artemis
1
2
3
4
5
6
7
8

Figure 21. Western blot analysis of L40 transformed with pHybLex/Zeo-Full length Artemis. The expression of Artemis-LexA DBD was not detected with anti-LexA antibody. Lanes 1-6 contain lysates of L40 transformed with the pHybLexZeo-FArtemis. Lane 7 contains WCE (whole cell extract) of 293T cells showing expression of Myc-Artemis. Lane 8 contains lysate of L40 untransformed (negative control). Detection with Anti-LexA antibody results in no detectable difference between L40 transformed with pHybLexZeo-FArtemis and untransformed L40.

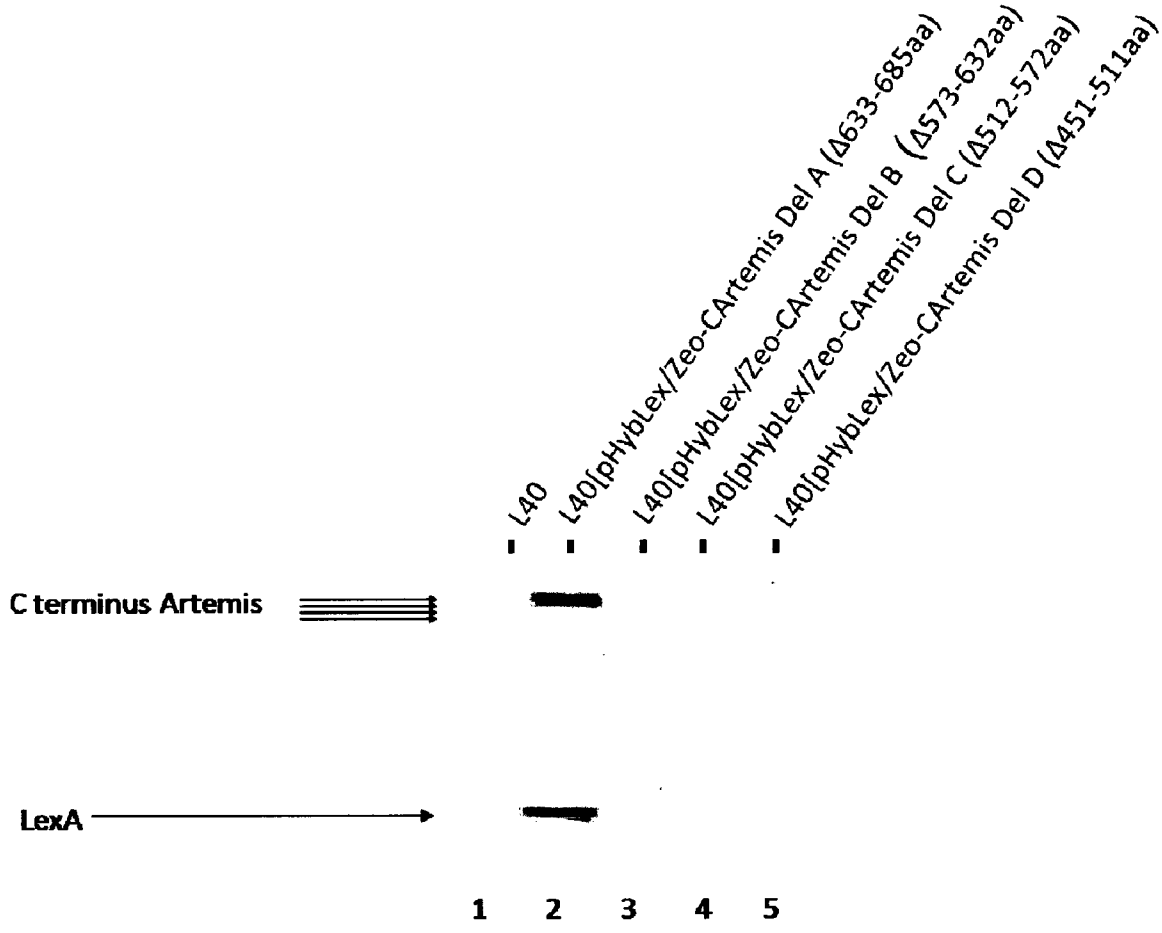

L40
L40[pHybLex/Zeo-CArtemis Del A (Δ633-685aa)
L40[pHybLex/Zeo-CArtemis Del B (Δ573-632aa)
L40[pHybLex/Zeo-CArtemis Del C (Δ512-572aa)
L40[pHybLex/Zeo-CArtemis Del D (Δ451-511aa)
C terminus Artemis
LexA
1
2
3
4
5

Figure 22. Western blot analysis of L40 transformed with various pHybLex/Zeo-CArtemis deletion constructs. Artemis-LexA DBD expression was detected using anti-LexA antibody on a 10% acrylamide gel. Lane 1 contains lysate of untransformed L40 showing no Artemis expression. Lanes 2-5 contains the lysate of L40 transformed with representative pHybLex/Zeo-CArtemis deletion constructs showing expression of truncated C-Artemis and un-fused LexA DBD.

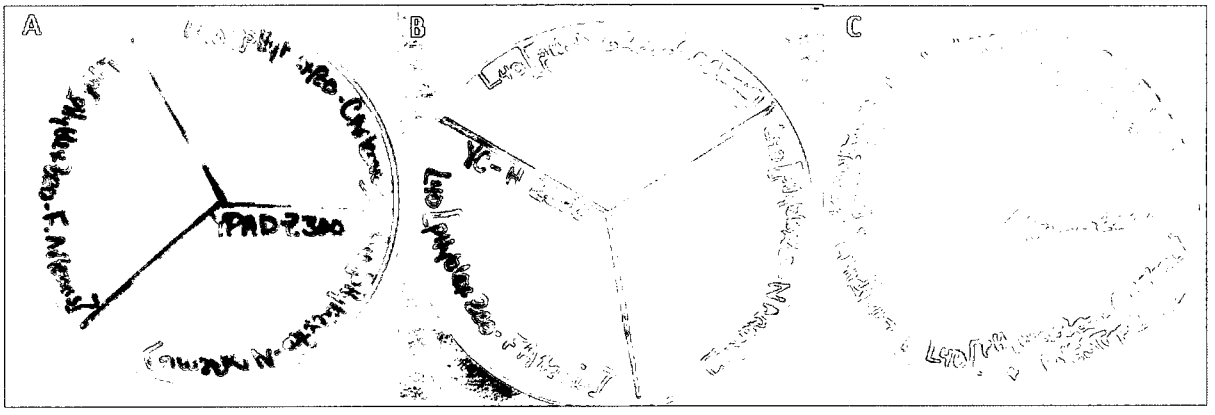

A
B
C

Figure 23. Bait construct pHybLexZeo-NArtemis and pHybLexZeo-FArtemis do not exhibit auto-activation in L40. **(A)** L40 [pHybLexZeo-NArtemis], L40 [pHybLexZeo-FArtemis], and L40 [pHybLexZeo-CArtemis] all grow on YPAD Z300 with similar size and growth kinetics. **(B)** L40 [pHybLexZeo-NArtemis] and L40 [pHybLexZeo-FArtemis] are unable to grow on YC-H Z300, while L40 [pHybLexZeo-CArtemis] is able to grow on YC-HZ300; the pHybLexZeo-CArtemis auto-activates *His* reporter gene alone. **(C)** L40 [pHybLexZeo-FArtemis + pYESTrp2] and L40 [pHyblexZeo-NArtemis + pYESTrp2] do not grow on YC-WH Z300; pHybLexZeo-NArtemis and pHybLexZeo-FArtemis do not auto-activate *His* reporter gene, when transformed with empty pYESTrp2 vector.

to grow on YC-HUK Z300 medium, L40 [pHybLex/Zeo-NArtemis + pYESTrp2] was not able to grow on YC-WHUK Z300. The pHybLex/Zeo-CArtemis bait construct did auto-activate the reporter genes in L40 (Figure 23B): L40[pHybLex/Zeo-CArtemis] was able to grow as efficiently on YC-HUK Z300 as on YPAD Z300 medium in parallel. The C-terminus Artemis was then truncated into four deletion constructs to obliterate the auto-activity of the C-Artemis-LexA fusion. Each of the four deletion constructs was able to trans-activate all of the reporter genes just as the pHybLex/Zeo-CArtemis construct. Only pHybLex/Zeo-NArtemis construct was "well-behaved" and appropriate to use in the screen.

Hybrid HunterTM human lung cDNA library

Determining library Titer

The number of colonies growing on each library serial dilution was counted and used to determine library titer.

Library titer (cfu/ml) = No. colonies / (plating volume (ml) x dilution factor)

Cfu/ml = number of colonies / (dilution x 10^2)
= 200 colonies (10^4 x 10^{-2} ml)
= 2.0 x 10^8 cfu/ml

Amplification of pYESTrp2-Human Lung cDNA library

Number of clones = Number of independent clones
= 5.96 x 10^6

Number of plates required = Number of clones to screen/ ~ colonies /150 mm plate
= 6 x 10^6 / 60 000
= 100 plates

Volume of library to plate = Number of clones/library titer
= 6 x 10^6 clones / 2.0 x 10^8cfu/ml
= 3.0 x 10^{-3} ml
= 30 µl

Result of Screen

Determining Primary Transformation Efficiency and Doubling factor

Primary transformation efficiency was determined by plating appropriate dilutions of library screen after initial transformation step on YC-WU Z300 plates to number of total doubly transformed cells. Dilution of $1:10^5$ and $1:10^6$ resulted in 105 and 9 colonies on YC-WUZ300 plates, respectively.

Transformation efficiency = number of colonies/ [(plating volume (ml)) x dilution factor]

$= 105 / (0.1 \text{ ml} \times 10^{-6})$

$\approx 10^7$ transformants

The doubling factor was determined by plating appropriate dilutions of library screen after growing library screen in YC-WUZ300 media for 16 hours selecting for doubly transformed cells. Dilution of $1:10^6$ and $1:10^7$ resulted in 161 and 12 colonies on YC-WUZ300 plates respectively.

Doubling factor = # doubly transformed cells / primary transformation efficiency

$= 160 \times 10^6 / 10^6$ transformants

$\approx$ 16 fold increase in doubly transformed cells

Doubling factor $\approx 2^4$

Determining Volume of Screen to Plate

The volume of library screen to be plated was determined based on primary transformation efficiency and doubling factor to ensure that a sufficient number of doubly transformed cells were plated to cover approximately 6 times the number of independent clones in the cDNA library.

Final volume of screen suspension = 10 ml

Volume of Screen to plate = 10 ml/ (primary transformation efficiency x Doubling factor)

(6 x Number of independent clones)

$= 10 \text{ ml} / [10^7 \times 2^4 / (6 \times 5.96 \times 10^6)]$

$= 2.3 \text{ ml}$

Following 7-10 days of incubation, 15 colonies grew on QDO (YC-WHUK Z300) medium, 12 of which exhibited β-galactosidase activity (**Table 3**). Five of these clones were identified as putative interactors, as they did not auto-activate the reporter genes alone or when transformed with pHybLex/Zeo. Furthermore, transformation of the putative interactors along with pHybLex/Zeo-NArtemis successfully reconstituted transcriptional activation of the reporter genes, allowing diploid L40 to grow on YC-WHUK Z300. **Figure 24** shows the restriction profiles of these clones. Sequence analysis of the putative interactors confirmed the presence of library insert into pYESTrp2 vector. Basic Local Alignment Search Tool (BLAST) of cDNA inserts revealed that none of the sequences contained a proper protein ORF in any of three possible frames (both 5' and 3' orientation), all sequences translated into short stretches of amino acids in all frames. A nucleotide BLAST search (searching a nucleotide database using a nucleotide query) revealed the presence of genomic sequences in the cDNA clones retrieved from all five colonies (**Table 4**).

Table 3. The putative interactors of pHybLex/Zeo-NArtemis that were retrieved from yeast colonies displaying the His+LacZ+ phenotype. Five putative interactors demonstrated a specific interaction with NArtemis, as they did not auto-activate the reporter genes alone or when co-transformed with pHybLex/Zeo. These putative interactors displayed β-galactosidase activity (√), did not auto-activate reporter genes alone or when co-transformed with pHybLex/Zeo (√), and were able to grow on YC-WHUK Z300 when co-transformed with pHybLex/Zeo-NArtemis. The prey-cDNAs that did not exhibit β-galactosidase activity (×), or auto-activated reporter genes alone or when co-transformed with pHybLex/Zeo (×), or were not able to grow on YC-WHUK Z300 when co-transformed with pHybLex/Zeo-NArtemis (×) were eliminated from further characterization.

Colony #	β-galactosidase activity	Prey cDNA does not Auto-activiate	L40 [Bait + Putative Interactor] grows on YC-WHUK Z300
1	×	-	-
2	√	√	√
3	√	√	√
4	√	×	-
5	√	×	-
6	√	×	-
7	×	-	-
8	√	√	√
9	√	×	-
10	√	×	-
11	×	-	-
12	√	√	√
13	√	√	√
14	√	×	-
15	√	×	-

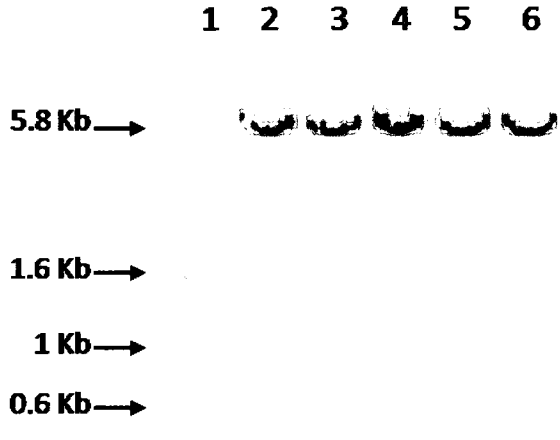

1 2 3 4 5 6
5.8 Kb
1.6 Kb
1 Kb
0.6 Kb
0.3 Kb

Figure 24. Restriction digest of five putative interactors which represented a specific interaction with NArtemis electrophoresed on a 1% agarose gel. Each cDNA clone was digested with *Hind*III and *Xho*I releasing the corresponding library inserts out of the 5.8 Kb pYESTrp2 vector. A 1-kb plus DNA ladder (lane 1) was used a size marker. Restriction digest revealed that cDNA clones 2, 3, 4, 5 and 6 contain 450 bp, 900 bp, 350 bp, 1 kb and 1.5 kb inserts respectively.

Table 4. BLAST identification of cDNA inserts from yeast colonies demonstrating potential interactions with N-terminus Artemis. BLAST analysis was conducted using "Translated query vs protein database" tool, comparing translated products of cDNA nucleotide sequences (in all 6 frames) to a protein database. BLAST analysis resulted in "No Significant Similarity Found" revealing that none of the sequences contained a protein ORF in any six possible frames. All nucleotide sequences were located in non-protein encoding regions and translated into short stretches of amino acids containing hydrophobic (in bold) and acidic amino acids (underlined). Indicated below are the short polypeptide translated in frame with LexA DBD of each clone.

Colony #	Organism	NCBI Map element	Protein ORF	Translated polypeptide of Prey cDNA
2	*Homo sapiens* chromosome 20	NT_011362	None	V5 AD L Q T G P K R **E** R **E** P R G **D** STOP
3	*H. Sapiens* chromosome 7	NT_007933.14	None	V5 AD R T R F C T **E** K K **E** K F **E** I STOP
8	*H. Sapiens* chromosome 6	NT_025741.14	None	V5 AD **E** Q A V L **E E E** Q V C G L T F **E** A P P C **D** L **E** G A L Q G I P R A A STOP
12	*H. Sapiens* chromosome 1	NT_004350	None	V5 AD G **E** G R **D** Q Q G **E** G R A G W V G G P A V G R **D** G Q STOP
13	*H. Sapiens* chromosome 9	NT_008470.18	None	V5 AD **D** A S K L G T **E** L G S T STOP

Six random clones from the Hybrid Hunter™ human lung cDNA library were sent for sequencing to examine the quality of the cDNA inserts in the library. Sequence analysis revealed that 5/6 clones contained no protein ORF in any possible frame and 4/6 clones encoded genomic sequences; thus, the quality of this cDNA library was believed to be inadequate for the purpose of Y2H experiment. A second cDNA library (DupLEX-A™

Human Fetal Brain cDNA library), compatible with the Hybrid hunter vector system, was used to conduct the following screens with the pHybLex/Zeo-NArtemis.

DupLEX-A™ Human Fetal Brain cDNA library

Determining Library Titer

► Cfu/ml = number of colonies / (dilution x 10^2)
= 75 colonies x (10^5 x 10^{-2})
= 7.5 x 10^7 cfu/ml

Amplification of pJG-45 Human Fetal Brain cDNA library

► Number of clones = Number of independent clones
= 3.5 x 10^6

►Number of plates required = Number of clones to screen/ ~ colonies per 150 mm plate
= 3.5 x 10^6 clones/ 60 000
≈ 60 plates

► Volume of library to plate = Number of clones/library titer
= 3.5 x 10^6 clones / 7.5 x 10^7 cfu/ml
= 47 µl

Result of Screen

Determining Primary Transformation Efficiency and Doubling Factor

Dilution of 1:10^5 and 1:10^6 resulted in 196 and 23 colonies on YC-WUZ300 plates respectively.

► Transformation efficiency = number of colonies/ [(plating volume (ml)) x dilution factor]

= 196/ (0.1 ml x 10^{-5})

≈ 2 x 10^7 transformants

Dilutions of 1:10^6 and 1:10^7 of screen culture following 16 hours of growth resulted in 112 and 8 colonies on YC-WUZ300 plates, respectively.

► Doubling factor = # of doubly transformed cells / primary transformation efficiency

$= 110 \times 10^6 / 10^6$ transformants

▶ Doubling factor $\approx 2^{3.46}$

Determining Volume of Screen to Plate

The volume of screen to be plated was determined based on primary transformation efficiency and doubling factor to ensure that a sufficient number of doubly transformed cells were plated to cover approximately 6 times the number of independent clones in the cDNA library.

Final volume of screen suspension = 11 ml

▶ Volume of Screen to plate $= 11 \text{ ml} / \left(\frac{\text{transformation efficiency} \times \text{doubling factor}}{6 \times \text{number of independent clones}} \right)$

$= 11 \text{ ml} / [2 \times 10^7 \times 2^{3.46} / (6 \times 3.5 \times 10^6)]$

$= \sim 1 \text{ ml}$

Following 7-10 days of incubation, 23 colonies grew on QDO (YC-WHUK Z300) medium, 17 of which exhibited β-galactosidase activity. Eleven of these clones were identified as putative interactors as they did not auto-activate the reporter genes alone or when transformed with pHybLex/Zeo. Furthermore, transformation of the putative interactors along with pHybLex/Zeo-NArtemis successfully reconstituted transcriptional activation of the reporter genes, allowing diploid L40 to grow on YC-WHUK Z300 (**Table 5**). The restriction digest profile of these clones are shown in **Figure 25**.

Sequence analysis of these putative interactors revealed the presence of library insert into pJG4-5 vector. Basic Local Alignment Search Tool (BLAST) of cDNA inserts revealed protein identification and regions of homology and frame of library inserts to putative interactors (Table 6).

Table 5. The putative interactors of pHybLex/Zeo-NArtemis that were retrieved from yeast colonies displaying the His+LacZ+ phenotype. Seven putative interactors demonstrated a specific interaction with NArtemis. These putative interactors displayed β-galactosidase activity (√), did not auto-activate reporter genes alone or when co-transformed with pHybLex/Zeo (√), and were able to grow on YC-WHUK Z300 when co-transformed with pHybLex/Zeo-NArtemis. The prey-cDNAs that did not exhibit β-galactosidase activity (×), or auto-activated reporter genes alone or when co-transformed with pHybLex/Zeo (×), or were not able to grow on YC-WHUK Z300 when co-transformed with pHybLex/Zeo-NArtemis (×) were eliminated from further characterization.

Colony #	β-galactosidase activity	Prey-cDNA does not Auto-activate	L40 [Bait + Putative Interactor] grows on YC-WHUK Z300
1	√	√	×
2	√	√	√
3	√	×	-
4	√	√	√
5	√	√	-
6	×	-	-
7	√	√	√
8	√	×	-
9	√	√	√
10	√	×	-
11	×	-	-
12	×	-	-
13	√	√	×
14	√	√	√
15	×	-	-
16	√	×	-
17	×	-	-
18	√	√	√
19	√	√	×
20	√	×	-
21	√	√	×
22	√	√	√
23	×	-	-

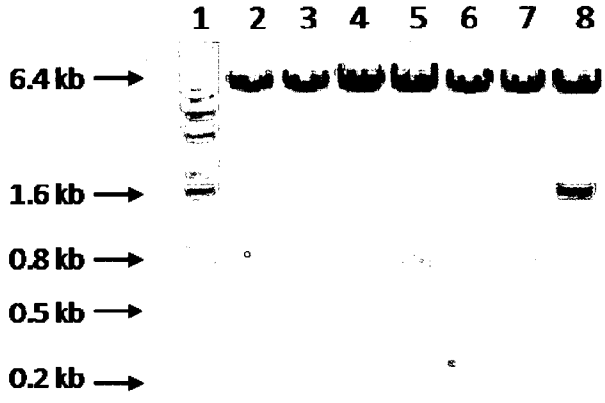

1 2 3 4 5 6 7 8
6.4 kb
1.6 kb
0.8 kb
0.5 kb
0.2 kb

Figure 25. Restriction digest of putative interactors electrophoresed on a 1% agarose gel. Each cDNA clone was digested with *EcoR*I and *Xho*I releasing the corresponding library inserts out of the 6.4 Kb pJG4-5 vector. A 1-kb plus DNA ladder (lane 1) was used a size marker. Restriction digest revealed that cDNA clones 2, 3, 4, 5, 6, 7, and 8 contain 750 bp, 500 bp, 700 bp, 800 bp, 950 bp, 850 bp, and 1.9 kb inserts respectively.

Table 6. BLAST identification of cDNA inserts from yeast colonies demonstrating potential interactions with N-terminus Artemis. BLAST analysis was conducted using "Translated query vs protein database" tool, comparing translated products of cDNA nucleotide sequences (in all 6 frames) to a protein database. All proteins identified were from the *H. Sapiens* genome.

Colony	BLAST Identification of cDNA	Protein ORF in frame with V5 AD	Sub-cellular localization	NCBI element
1	NADH dehydrogenase subunit 1	Yes	Mitochondrion (Cristae)	ACA21981
2	TIMP3 metallopeptidase inhibitor	Yes	ECM (retina)	CAG30479
4	Lysosomal-associated multispanning membrane protein 5	Yes	Lysosomal Membrane	NM_006762
7, 9	Cytochrome c oxidase subunit I	Yes	Mitochondiron (Cristae)	ABY86862
13	Heat shock protein 90 (Hsp90)	Yes	Cytosol/Endoplasmic Reticulum/ Mitochondrion	NP_005339
14	ATRX protein	No (+2 of V5 AD)	Nucleus	dbj|BAC81110.1|

ATRX (alpha thalassemia/mental retardation syndrome X-linked) protein

Due to its physiological activity as a DNA-interacting biomolecule, clone 14 represents a particularly interesting potential interactor (**Figure 26**). The ATRX mRNA was inserted into Frame +2 of the V5 AD. BLAST alignment of the retreived cDNA insert in the +2 frame of the V5 AD versus the known sequence of ATRX protein, as well as the pepide generated in frame with the V5 AD are shown below:

Protein sequence alignment of translated cDNA from colony 14 versus ATRX protein

Positives = 89/91 (98%)
E-value = $3x10^{-38}$

```
Query          RTTKKRIPNPKDFXSSEDEKHSKKGMDNQGHKNLKTSQEGSSDDAERKQERETFSSAEGT
               RTTKKRIPN KDF SSEDEKHSKKGMDNQGHKNLKTSQEGSSDDAERKQERETFSSAEGT
ATRX    837     RTTKKRIPNTKDFDSSEDEKHSKKGMDNQGHKNLKTSQEGSSDDAERKQERETFSSAEGT

Query          VDKDTTIMELRDRLPKKQQASASTDGVDKLS
               VDKDTTIMELRDRLPKKQQASASTDGVDKLS
ATRX           VDKDTTIMELRDRLPKKQQASASTDGVDKLS   927
```

Peptide generated in frame with V5 AD

V5 AD- ESLSTPSVEALECEFLGSRSLNSMMVVSLXTVPSAEEKVSLSCFLSASSDDP SXEXFKFLCP- **STOP**

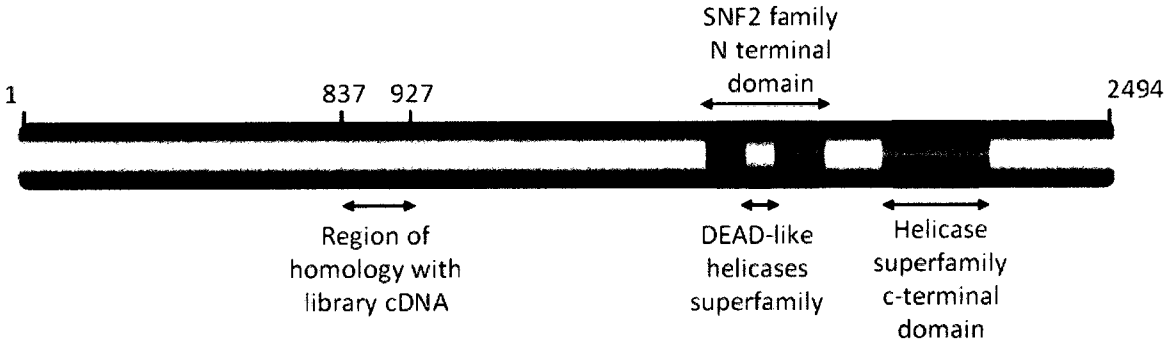

SNF2 family
N terminal
domain
1
837
927
2494
Region of
homology with
library cDNA
DEAD-like
helicases
superfamily
Helicase
superfamily
c-terminal
domain

Figure 26. Structure and domain architecture of the ATRX protein. The ATRX protein contains the SNF2 family N terminal domain (aa1562-1881), the DEAD-like helicase superfamily domain(aa1590-1752), and helicase superfamily c-terminal domain (aa1813-1959); all shared among proteins that are involved in DNA or RNA unwinding [Adapted from (137)].

Matchmaker™ Yeast Two-Hybrid Gal-4 Based System

Bait Expression: pGBKT7-cPARP-1 and -Artemis constructs

A direct assay of bait expression was determined via western blot analysis. The pGBKT7-cPARP-1 (**Figure 27**), pGBKT7-NArtemis (**Figure 28A**), pGBKT7-CArtemis and pGBKT7-FArtemis (**Figure 28B**) constructs were expressed in AH109 at detectable levels at the appropriate size.

Testing Bait plasmids for non-specific Transcriptional Activation

pGBKT7-cPARP-1

The pGBKT7-cPARP-1 bait construct did not trans-activate any of the reporter genes in AH109 (**Figure 29A**). AH109[pGBKT7-cPARP-1], and AH109 [pGBKT7-cPARP-1 + pGADT7] were not able to grow on YC-WHA and YC-WHAL respectively (**Figure 29B**). AH109 transformed with positive control constructs, AH109[pGBKT7-p53 + pGADT7-T] was able to grow on YC-WHAL (supplemented with 2.5 mM of 3-AT) (**Figure 29C**).

pGBKT7-NArtemis, pGBKT7-CArtemis, pGBKT7-FArtemis

The pGBKT7-NArtemis or pGBKT7-FArtemis bait construct did not trans-activate any of the reporter genes in AH109. AH109[pGBKT7-NArtemis or –FArtemis], and AH109 [(pGBKT7-NArtemis or –FArtemis) + pGADT7] were not able to grow on YC-WHA and YC-WHAL respectively. The pGBKT7-CArtemis bait construct did trans activate *His3* and *Ade* reporter genes alone. AH109 transformed with positive control constructs, AH109[pGBKT7-p53 + pGADT7-T] was able to grow on YC-WHAL.

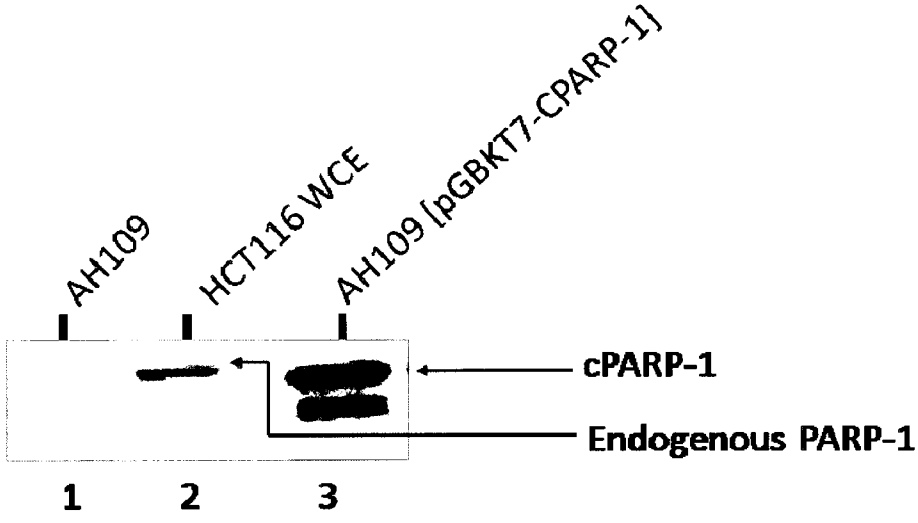
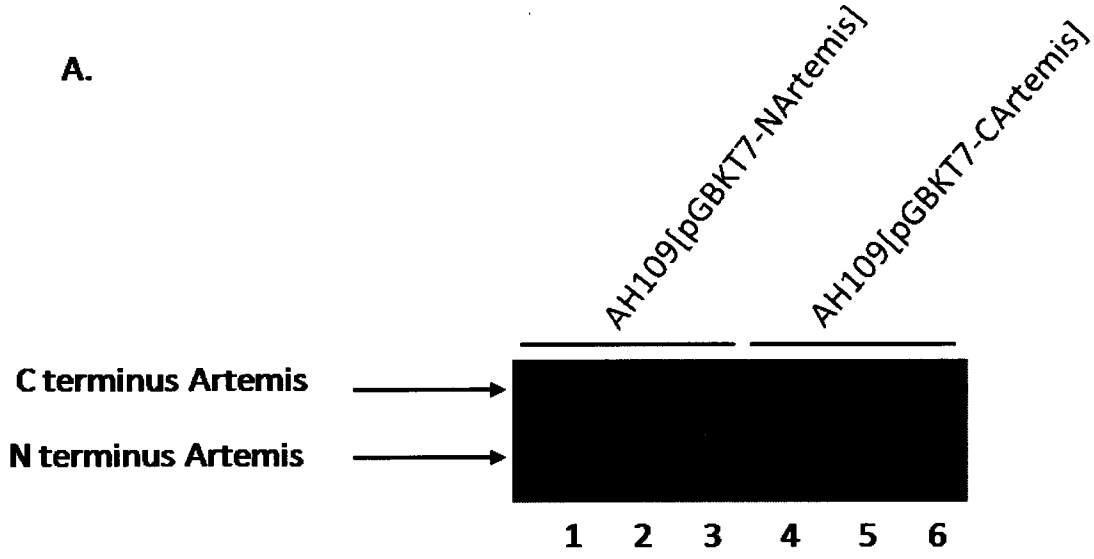

AH109
HCT116 WCE
AH109 [pGBKT7-CPARP-1]
cPARP-1
Endogenous PARP-1
1
2
3

Figure 27. Western blot analysis of AH109 transformed with pGBKT7-CPARP-1. PARP-1 expression was detected with anti-PARP-1 Ab (6D640 sc-71849; Santa Cruz) on a 10% acrylamide gel. Lane 1 contains the lysate of untransformed AH109 showing no PARP-1 expression. Lane 2 contains the lysate of HCT116 cells (positive control) showing expression of endogenous PARP-1. Lane 3 contains lysate of AH109 transformed with pGBKT7-CPARP-1, showing expression of C-terminus of PARP-1.

A.


AH109[pGBKT7-NArtemis]
AH109[pGBKT7-CArtemis]
C terminus Artemis
N terminus Artemis
1
2
3
4
5
6

B.

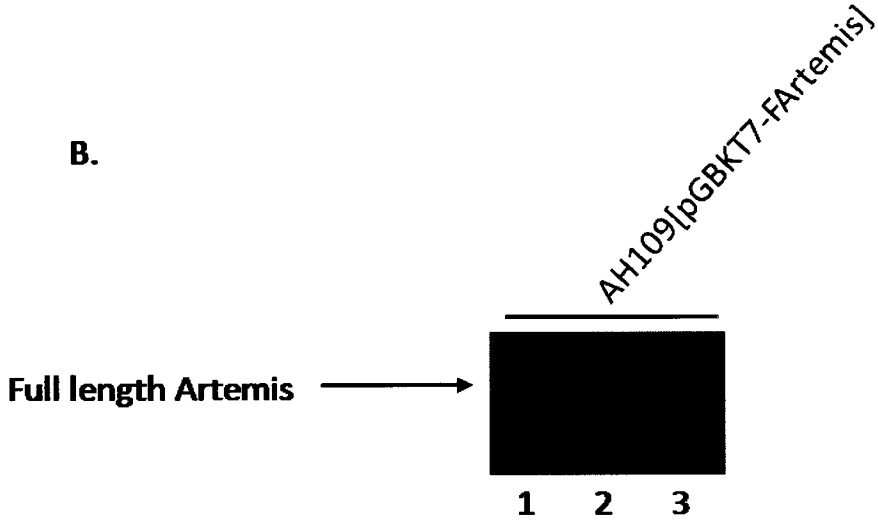

AH109[pGBKT7-FArtemis]
Full length Artemis
1
2
3

Figure 28. Western blot analysis of AH109 transformed with pGBKT7-Artemis constructs. (A) Lanes 1-3 contain lysate of AH109 transformed with pGBKT7-NArtemis, expression of N-terminus Artemis-Gal4 DBD fusion is detected on a 10% acrylamide gel. Lanes 4-6 contain lysate of AH109 transformed with pGBKT7-CArtemis, expression of C-terminus Artemis-Gal4 DBD fusion is detected. **(B)** Lanes 1-3 contain lysate of AH109 transformed with pGBKT7-FArtemis, expression of full length Artemis-Gal4 DBD fusion is detected on a 10% acrylamide gel. Artemis-Gal4 DBD expression was detected with anti-GAL4 DBD Ab [Rabbit GAL4 (N-19), Polyclonal; Santa Cruz Biotechnology].

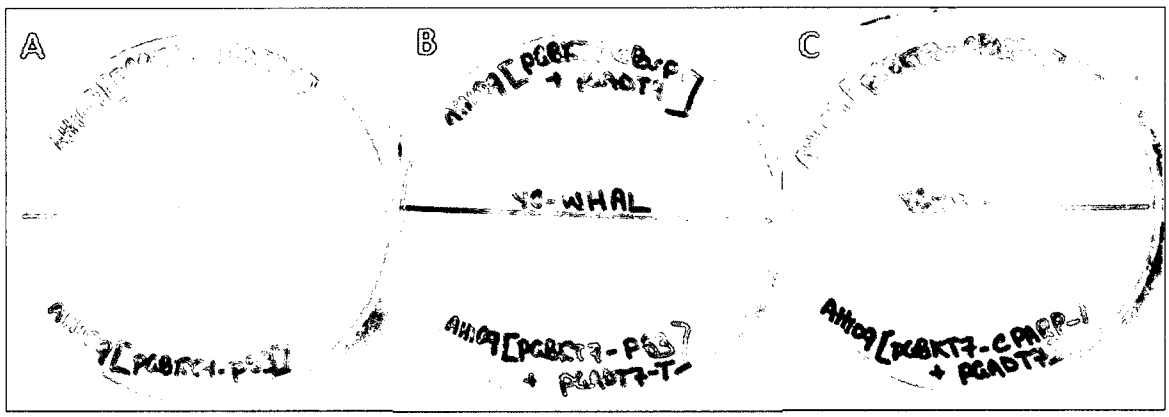

A
B
C

Figure 29. Bait construct pGBKT7-cPARP-1 does not exhibit auto-activation in AH109. **(A)** AH109 [pGBKT7-cPARP-1] and AH109 [pGBKT7-p53] (positive control construct) are unable to grow on YC-WHA; the bait constructs do not auto-activate *His* and *Ade* reporter genes alone. **(B)** AH109 [pGBKT7-cPARP-1 + pGADT7] does not grow on YC-WHAL; bait construct does not auto-activate *His* and *Ade* reporter genes when transformed with empty prey vector. Three AH109 [pGBKT7-p53 + pGADT7-T] colonies are able to grow on YC-WHAL; positive control vector express p53 and large T proteins which interact allowing yeast to express *Ade* and *His* reporter genes. **(C)** AH109 [pGBKT7-cPARP-1 does not grow on YC-WL, while three AH109 [pGBKT7-cPARP-1 + pGADT7-T] colonies grow on YC-WL; confirming correct phenotype of bait transformed AH109, bait strain transformed with pGADT7 grows on YC-WL, while alone bait strain does not grow on YC-WL.

Measuring viability of mating partners

The viability of each mating partner and diploid yeast was calculated to determine the mating efficiency of the screen. The strain with the lower viability is the "limiting partner".

Viability of Y187 partner = # of cfu/ml on YC- L plate

Viability of AH109 partner = # of cfu/ml on YC-W plate

Viability of Diploids = # of cfu/ml on YC-LW plate

▶ Viability (cfu/ml) $= \frac{\text{cfu}}{\text{(volume plated (µl) x dilution factor)}} \times 1000\ \mu l/ml$

Calculate Mating Efficiency (percentage of diploids)

▶ % Diploids $= \frac{\text{Viability of Mating culture}}{\text{Viability of limiting partner}} \times 100\ \%$

Mating Efficiency of Positive Control: AH109[pGBKT7-p53] and Y187[pGADT7-T]

Characteristics of the host yeast strains and including the various transformants are described in **Table 7**.

Viability of AH109[pGBKT7-p53] $= \frac{1500 \times 1000\ \mu l/ml}{100\mu l \times 10^{-3}}$

$= 1.5 \times 10^7$ cfu/ml

Viability of Y187[pGADT7-T] $= \frac{3500 \times 1000\ \mu l/ml}{100\mu l \times 10^{-3}}$

$= 3.5 \times 10^7$ cfu/ml

Viability of mating Culture $= \frac{45 \text{ colonies} \times 1000\ \mu l/ml}{100\mu l \times 10^{-3}}$

$= 4.5 \times 10^5$ cfu/ml

% diploids $= 4.5 \times 10^5$ cfu/ml / 1.5×10^7 cfu/ml x 100%

Therefore, mating efficiency = 3 %

Table 7. The viability of haploid partners and the mating efficiencies of diploid yeast after yeast mating using the small scale yeast mating protocol. The haploid partner with the lower viability is the limiting partner (√) in each mating reaction.

Host Strain	Selective Medium	Dilution Factor	# of Colonies	Viable cfu/ml	Limiting Partner	Mating Efficiency
AH109 [pGBKT7-p53]	YC-W	10^{-3}	1500	1.5×10^{7}	√	-
Y187 [pGADT7-T]	YC-L	10^{-3}	3500	3.5×10^{7}	-	-
Mating culture (Diploids) (AH09+Y187[pGBKT7-p53+ pGADT7-T])	YC-WL	10^{-3}	45	4.5×10^{5}	-	3%
AH109 [pGBKT7-FArtemis]	YC-W	10^{-3}	1150	1.2×10^{7}	√	-
AH109 [pGBKT7-cPARP-1]	YC-W	10^{-3}	950	9.5×10^{6}	√	-
Y187 [pGADT7-T]	YC-L	10^{-3}	3000	3.0×10^{7}	-	-
Mating culture (Diploids) (AH109+Y187[pGBKT7-FArtemis + pGADT7-T])	YC-WL	10^{-1}	30	1.5×10^{3}	-	0.03 %
Mating culture (Diploids) (AH109+Y187[pGBKT7-cPARP-1 + pGADT7-T])	YC-WL	10^{-1}	45	4.5×10^{3}	-	0.05%

Results of Screen

Mating efficiencies of screens with pGBKT7-cPARP-1 and pGBKT7-FArtemis were significantly lower than the control strains (0.03-0.06%). Optimizing mating conditions prior to screening as previously described (135) increased the mating efficiency approximately 10 fold for both bait constructs.

Determining Mating Efficiency and Doubling Factor: Bait: pGBKT7-cPARP-1

Final volume of screen suspension = 11.5 ml

Viability of AH109 [pGBKT7-cPARP-1] = 1.0×10^7 cfu/ml

Viability of Y187 [pGADT7-library] = 9.5×10^7 cfu/ml

Viability of diploid mating culture = 7.0×10^4 cfu/ml

▶ Therefore, mating efficiency = 7.0×10^4 cfu/ml / 1.0×10^7 cfu/ml x 100%

= 0.7 %

Dilutions of $1:10^4$ and $1:10^5$ of screen culture following 16 hours of growth resulted in 110 and 9 colonies on YC-WL plates, respectively.

▶ Doubling factor = # of diploids following 16 hr growth / # of diploids of mating culture

= 1.0×10^6 cfu/ml / 7.0×10^4 cfu/ml

$\approx 2^{3.4}$ fold increase in diploid cells

Determining Volume of Screen to Plate

Viability of mating culture following 16 hour growth = 1.0×10^6 cfu/ml

Therefore # diploids in final screen = 1.0×10^6 cfu/ml x 11.5 ml

= 1.2×10^7 diploids

The number of independent clones in the library was 8.8 x 10^6 , while 1.2 x 10^7 diploids were screened; therefore, plating the entire volume of yeast suspension would be necessary to screen every independent clone in the cDNA library. However, because His^+ colonies grow poorly if the plating density is too high, the entire suspension could not be plated. Optimal plating volume was determined to be 25-30 µl of suspension / 150 mm plate. Therefore, the library could not be screened until saturation. A 25 µl aliquot of the final screen was plated on each of 200 large 150 mm plates to screen. Following 7-10 days of incubation, 230 colonies grew on QDO (YC-WHAL) medium, 129 of which exhibited β-galactosidase activity. Restriction analysis of the retrieved cDNA from a single yeast colony revealed the presence of an additional re-occurring "false positive", that did was not lineralized with Hind III digestion and appeared to have a lower copy number (~10 fold less) (**Figure 30**). This "false positive" was retrieved repeatedly from each one of the 129 yeast colonies that were analyzed for the presence of a putative interactor. Although 14 of the preparations revealed the presence of the false positive as a sub-population that accompanied the presence of the correct pGADT7-Rec library plasmid, 115 of the preparations revealed presence of only the false positive without any sign of pGADT7-Rec (**Figure 31**). The 14 prey cDNAs retrieved were identified as putative interactors as they did not auto-active the reporer genes alone, or when transformed with pGBKT7 (Table 8). Transformation of the putative interactors along with pGBKT7-cPARP-1 successfully reconstituted transcriptional activation of the reporter genes, allowing the diploid AH109 to grow on YC-WHAL medium. Sequence analysis of these putative interactors revealed the presence of library insert into pGADT7- Rec vector (Table 9). Basic Local Alignment Search Tool (BLAST) of cDNA inserts revealed protein identification and regions of homology and frame of library inserts to putative interactors (133).

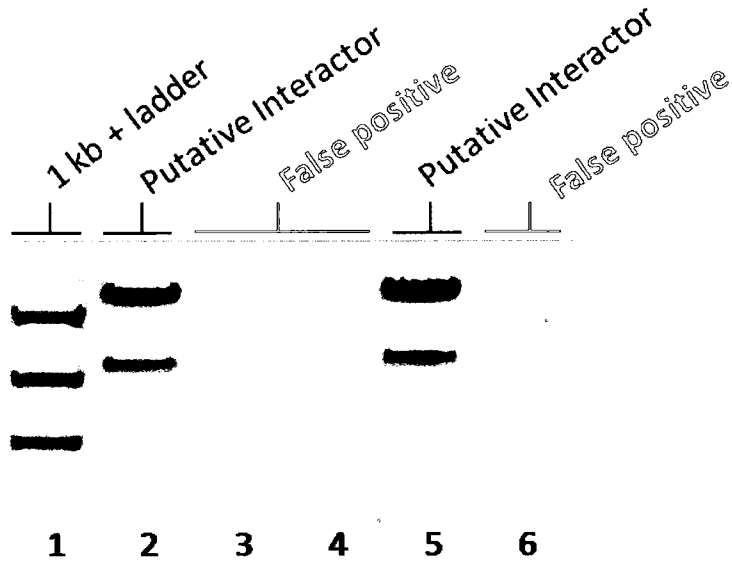

1 kb + ladder
Putative Interactor
False positive
Putative Interactor
False positive
1
2
3
4
5
6

Figure 30. Characterization of retrieved cDNA from a single diploid Ade+His+LacZ+ yeast colony # 5 by restriction analysis. Lanes 1 and 4 contain the product of Hind III digestion of the correct pGADT7-Rec vector and library insert . Lanes 2,3 and 5 (outlined in red) contain a "false positive" that remained undigested after HindIII digestion, and was retrieved from the lysate of the same yeast colony.

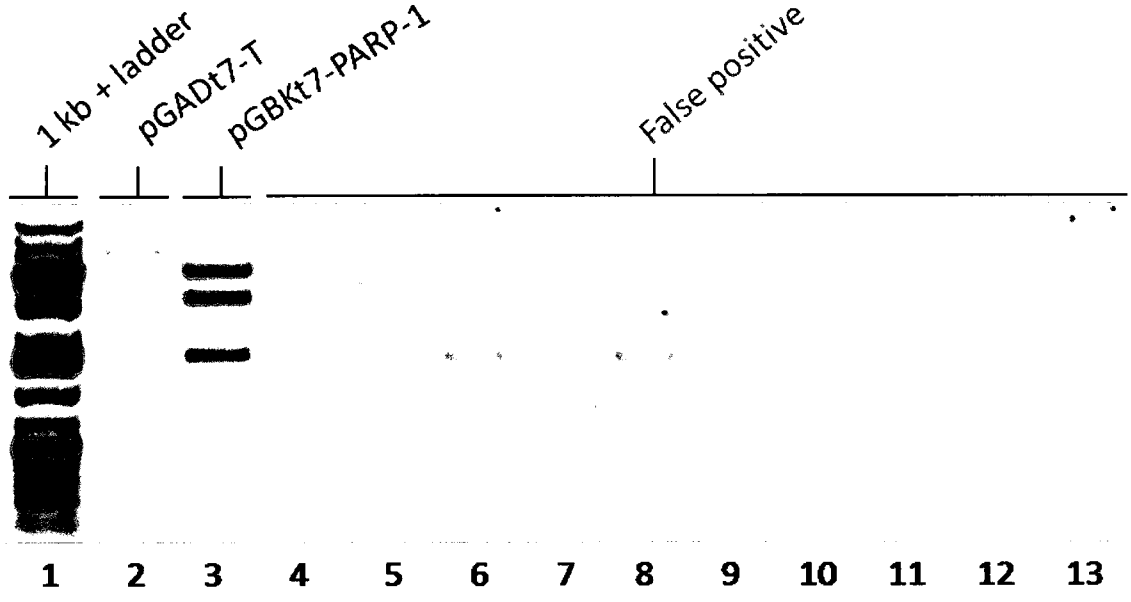

1 kb + ladder
pGADt7-T
pGBKt7-PARP-1
False positive
1
2
3
4
5
6
7
8
9
10
11
12
13

Figure 31. Characterization and comparison of retrieved bait and prey plasmids versus the re-occurring "false positive". Lane 1 contains the HindIII digest of pGADT7-T that was retrieved from diploid AH109 expressing positive control plasmids. Lane 2 contains the HindIII digest of pGBKT7-cPARP-1 retrieved from diploid *Ade+His+LacZ+* colony #1. Lanes 3-13 contain the reoccurring undigested false positive that was retrieved from10 separate preparations from DNA that was recovered from the same diploid *Ade+His+LacZ+* colony #1.

Table 8. The putative interactors that were retrieved from yeast colonies displaying Ade+His+LacZ+ phenotype. Seven putative interactors demonstrated a specific interaction with cPAPR-1. These putative interactors displayed β-galactosidase activity (√), did not auto-activate reporter genes alone or when co-transformed with pGBKT7 (√), and were able to grow on YC-WHAL when co-transformed with pGBKT7-cPARP-1. The prey-cDNAs that did not exhibit β-galactosidase activity (×), or auto-activated reporter genes alone or when co-transformed with pGBKT7 (×), or were not able to grow on YC-WHAL when co-transformed with pGBKT7-cPARP-1 (×) were eliminated from further characterization.

Colony	β-galactosidase activity	Prey cDNA does not auto-activate	AH109[Bait + Putative Interactor] grows on YC-WHAL
5	√	√	√
13	√	√	√
19	×	-	-
25	×	-	-
28	√	√	√
40	√	√	√
41	√	√	√
48	√	×	-
55	×	-	-
70	√	×	-
77	×	-	-
80	√	×	-
103	√	√	√
138	√	√	√

Table 9. BLAST identification of cDNA inserts from yeast colonies demonstrating potential interactions with C-terminus of PARP-1. BLAST analysis was conducted using "Translated query vs protein database" tool, comparing translated products of cDNA nucleotide sequences (in all 6 frames) to a protein database. All proteins identified were from the *H. Sapiens* genome.

Colony #	BLAST identification of cDNA	Protein ORF in frame with GAL4 AD	Sub-cellular localization	NCBI Map element
13	Caspase-7	No (+1 of GAL4 AD)	Nucleus/ Cytosol	emb|CAI12641.1|
28	Unknown Protein	No (+2 of GAL4 AD)	Nucleus	gb|AAA88038.1|
5,41	Hemoglobin α-1 globin chain	Yes	Cytosol	AAK37554
48,103	Hemoglobin β chain variant	Yes	Cytosol	P02024
138	Immunoglobulin lambda heavy chain	Yes	Cytosol	emb|CAA75032.1|

Caspase-7

Due to its physiological role in the apopotic pathway, the cDNA encoding the C-terminus of caspase-7 (isolated from clone 13) represents a particularly interesting potential interactor for PARP-1 (**Figure 32**). The Caspase-7 mRNA was inserted into Frame +1 of the GAL4 AD. BLAST alignment of the retreived cDNA insert in the +1 frame of the GAL 4 AD versus the known sequence Caspase-7, as well as the pepide generated in frame with the GAL4 AD are shown below:

Sequence alignment of putative interactor (colony 13) versus C terminus of Caspase-7

Identities = 67/85 (78%), Positives = 70/85 (82%), Gaps = 3/85 (3%)

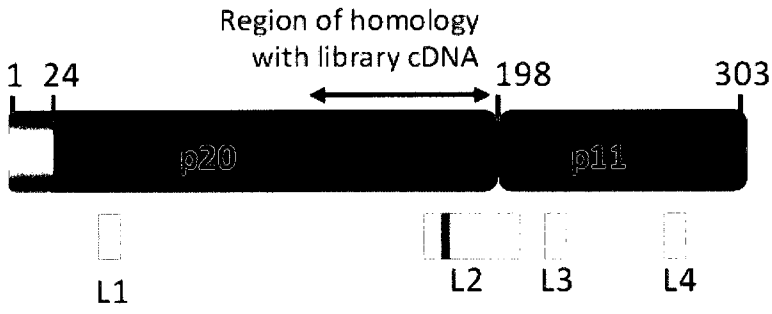
```
Query  1      SVTQAGVQRRHLHRLRPPPRFFKQRSSLPSILSSWDYRSVSPRLANFCIFRRYDERHVGQ
              SVTQAGVQRR L RL+PPP  FKQ S L S+LSSWDYR   PRLANFCIFRR   RHVGQ
Casp-7 115    SVTQAGVQRRDLGRLQPPPPRFKQYSCL-SLLSSWDYRHAPPRLANFCIFRREGFRHVGQ

Query  61     SGLKLLASSDPPT--SQSAGITGVS  83
              +GLKLLASSDPPT  SQSAGITGVS
Casp-7 174    AGLKLLASSDPPTSGSQSAGITGVS  198
```

Peptide generated in frame with GAL4 AD

GAL4 AD –NRSHSVTQAGVQWHDHGPPQPRPSRLKRSSHPSILSGWDYRDVSPR PANF-**STOP**

(A)


Region of homology
with library cDNA
1
24
198
303
p20
p11
Caspase-7
L1
L2
L3
L4

(B)

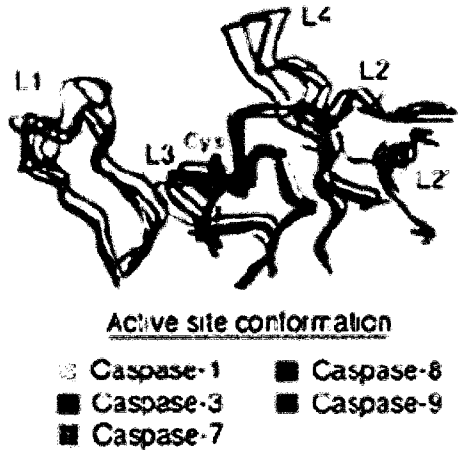

L1
L3
L2
L2
Active site conformation
Caspase-1
Caspase-3
Caspase-7
Caspase-8
Caspase-9

Figure 32. Structure and domain organization of human Caspase-7. (A) Caspase-7 encodes a 35-kDa protein. The active caspase-7 is a homodimer, each monomer consisting of a large (p20) and a small (p10) catalytic subunit. The active site of caspases are formed from four protrouding loops (L1-4) depicted in grey blocks. Loop L1 and a portion of L2, which contains the catalytic Cys residue (depicted in red line), constitute a portion of the p20 subunit, whereas L3 and L4 constitute a portion of the p10 subunit. **(B)** Representation of active site conformations of the caspases with known structures. Loops L1 and L3 are highly conserved among caspases, while L2 and L4 represent regions that ensure differences in substrate binding specificity [Adapted from (137, 138)].

"Unknown Protein"

Due to its physiological role as a chromatin associated protein with functions in DNA dynamics, the unknown protein (isolated from colony 28) represents a particularly interesting putative interactor for PARP-1 (Figure 33). The "unknown protein" mRNA was inserted into Frame +2 of the GAL4 AD. BLAST alignment of the retreived cDNA insert in the +2 frame of the GAL4 AD versus the known sequence the "unknown protein", as well as the pepide generated in frame with the GAL4 AD are shown below:

Sequence alignment of putative interactor versus C terminus of "Unknown Protein"

Identities = 92/120 (76%), Positives = 98/120 (81%), Gaps = 1/120 (0%)

```
Query  1          SVRLLLRDLTEEPLFYDAILLGSIYP-EKKSFPYKDTCTFMFIAALFTIAKTWNQPKCPS  59
                  SV   LRDL  E  F  AI L  IYP E KS  YKDTCT MFIAALFTIAKTWNQPKCP+
Unknown 1073      SVWRFLRDLELEIPFDPAIPLLGIYPNEYKSCCYKDTCTRMFIAALFTIAKTWNQPKCPT 1132

Query  60         MIDWITKMWHIYTIEYYASIKKDEFLSFSGTSMKLETIILSKLTQEQKTKHHMFSLIGGN  119
                  MIDWI KMWHIYT+EYYA+IK DEF+SF GT MKLETIILSKL+QEQKTKH +FSLIGGN
Unknown 1133      MIDWIKKMWHIYTMEYYAAIKNDEFISFVGTWMKLETIILSKLSQEQKTKHRIFSLIGGN 1192
```

Sequence of peptide generated in frame with GAL4 AD

GAL4 AD–SNPKEKKSFPPKDTCTFMCITALFHNSKDLEPTQKSINDRLDNKNVAHI HHRILCIHKKG-**STOP**

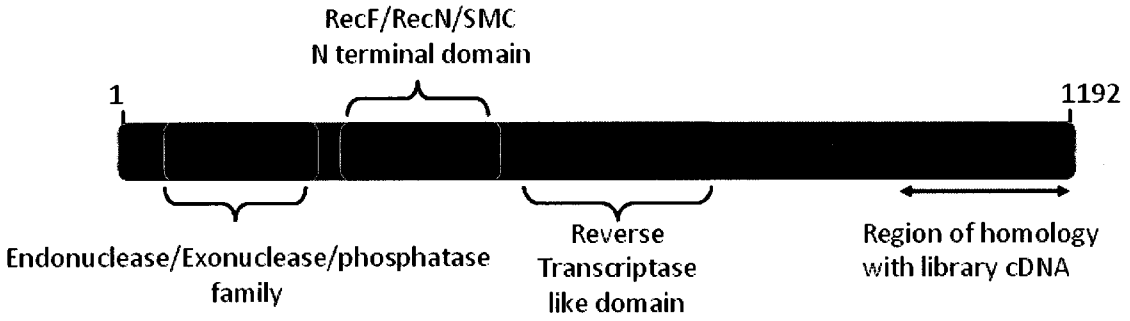

RecF/RecN/SMC
N terminal domain
1
1192
Endonuclease/Exonuclease/phosphatase
family
Reverse
Transcriptase
like domain
Region of homology
with library cDNA

Figure 33. Structure and conserved domain organization of unknown protein. The most N terminal region of this protein classifies this protein as a member of the endonuclease/exonuclease/phosphatase family. The RecF/RecN/SMC N terminal domain classifies this unknown protein as one of the SMC (structural maintenance of chromosomes) superfamily of proteins; involved in chromatin and DNA dynamics, and the RecF and RecN family of proteins; involved in DNA metabolism and recombination. The Reverse Transcriptase like domain [Adapted from (137)].

Calculating Mating Efficiency of Screen and Doubling Factor. Bait: pGBKT7-FArtemis

Final volume of screen suspension = 12 ml

Viability of AH109 [pGBKT7-FArtemis] = 8.8×10^6 cfu/ml

Viability of Y187 [pGADT7-library] = 1.0×10^8 cfu/ml

Viability of diploid mating culture = 5.1×10^4 cfu/ml

Therefore, mating efficiency = 5.1×10^4 cfu/ml / 8.8×10^6 cfu/ml x 100%

= 0.6 %

Dilution of $1:10^4$ and $1:10^5$ of screen culture following 16 hours of growth resulted in 72 and 6 colonies on YC-WL plates respectively.

Doubling factor = # of diploids following 16 hr growth / # of diploids of mating culture

= 8.0×10^5 cfu/ml / 5.1×10^4 cfu/ml

Doubling factor $\approx 2^4$ fold increase in diploid cells

Determining Volume of Screen to Plate

Viability of mating culture following 16 hour growth = 8.0×10^5 cfu/ml

Therefore # diploids in final screen = 8.0×10^5 cfu/ml x 12 ml

= 9.6×10^6 diploids

There were 8.8×10^6 independent clones in the cDNA library, while 9.6×10^6 diploids were screened; therefore, the final number of diploids was roughly equivalent to what was required to screen the number of independent clones in the library. Optimal plating volume was 25-30 µl of suspension / 150 mm plate. 25 µl of library screen was plated on each of 200 large 150 mm plates that screened approximately 5 ml of the screen

suspension (~42% of the number of independent clones). Following 7-10 days of incubation, 197 colonies grew on QDO (YC-WHAL) medium, 99 of which exhibited β-galactosidase activity. Restriction analysis of the retrieved cDNA from each single yeast colony revealed the presence of the re-occurring "false positive", that was encountered with previous screen using bait pGBKT7-cPARP-1. Analysis of DNA preps retrieved from 96 colonies revealed the presence of this false positive, only 3 DNA preps contained the proper plasmid and cDNA insert (**Table 10**).

Table 10. BLAST identification of cDNA inserts from yeast colonies demonstrating potential interactions with full length Artemis. BLAST analysis was conducted using "Translated query vs protein database" tool, comparing translated products of cDNA nucleotide sequences (in all 6 frames) to a protein database. All proteins identified were from the *H. Sapiens* genome.

Colony #	BLAST Identification of cDNA	Protein ORF in frame with Gal4 AD	NCBI Map element	Sub-cellular localization
26	CAP1	Yes	ref|NP_006358.1|	Cytoplasm
47	Hemoglobin beta chain	Yes	sp|P02024|HBB_GORGO	Red blood cell
55	Alpha 2 globin variant	Yes	dbj|BAD97112.1|	Red blood cell

Characterization of False Positive

In order to characterize the false positive, a restriction enzyme that could linearize the plasmid was identified (Figure 34). The false positive revealed to be a 2.5 kb plasmid which was then cloned into the pGBKT7 plasmid and sent for sequencing analysis. The 2.5 kb plasmid revealed to be the by-product of site specific intra-plasmid recombination event

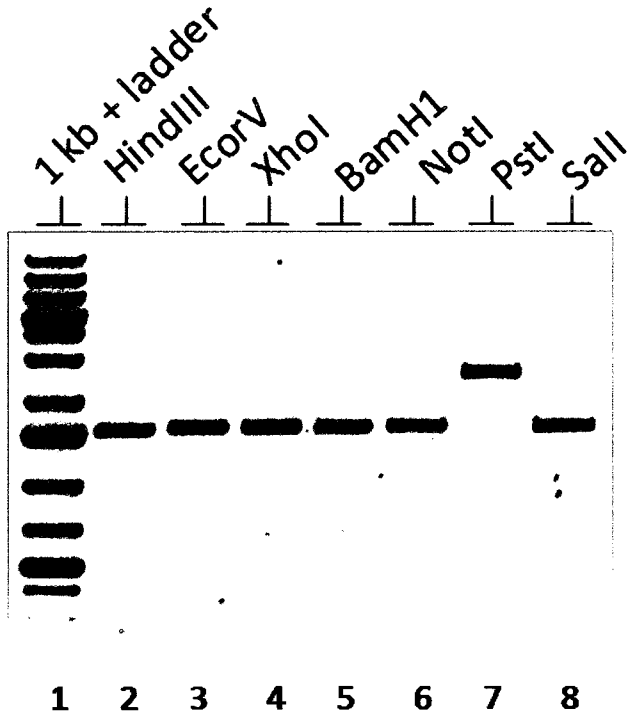

1 kb + ladder
HindIII
EcorV
XhoI
BamH1
NotI
PstI
SalI
1
2
3
4
5
6
7
8

Figure 34. Agarose gel electrophoresis characterizing the "false positive". The false positive retrieved from various diploid colonies obtained from library screen using restriction analysis. A 1-kb plus DNA ladder (lane 1) was used a size marker. The plasmid is linearized with PstI digestion into an approximately 2.5 Kb plasmid (lane 7).

within pGADT7-Rec. The yeast had spliced two major sections of the plasmid including the cDNA insert (**Figure 35**). What remained of the original vector was the 2 μ and pUC origin of replications, the T7 promoter and the Ampicillin resistance gene.

pGADT7-T exhibits instability

The pGADT7-Rec vector is a modified version of pGADT7(Appendix V); it is engineered for synthesis of Matchmaker pre-transformed cDNA libraries to allow the vector to undergo recombination event with SMART-amplified cDNA in yeast cells (134). The prey vector expressing the large T antigen (pGADT7-T) supplied by the manufacturer for control experiments is cloned into the pGADT7 vector, while the cDNA library used for the screen is cloned into pGADT7-Rec. In order to asses the stability (tendency to recombine) of pGADT7-T, ten yeast colonies expressing pGADT7-T were lysed and plasmid DNA was retrieved. The plasmid DNA was isolated and analyzed via restriction analysis (**Figure 36**). Hind III digest of recovered plasmid revealed that 8 out of 10 pGADT7-T vectors generated two DNA bands corresponding to pGADT7 vector (~ 8 kb) and large T antigen insert (~1.2 kb). However, 2 out of the 10 vector digested with HindIII generated DNA bands that were incorrect in size demonstrating potential recombination events in a site-specific manner (Figure 36: Lanes 4 and 8).

pGADT7-Rec exhibits instability: Site-Specific Recombination

In order to assess the instability of pGADT7-Rec containing human bone marrow (HB) cDNA library inserts without stress from selective dropout media or changes in pH, plasmid DNA was recovered from Y187 expressing random clones of the cDNA library.

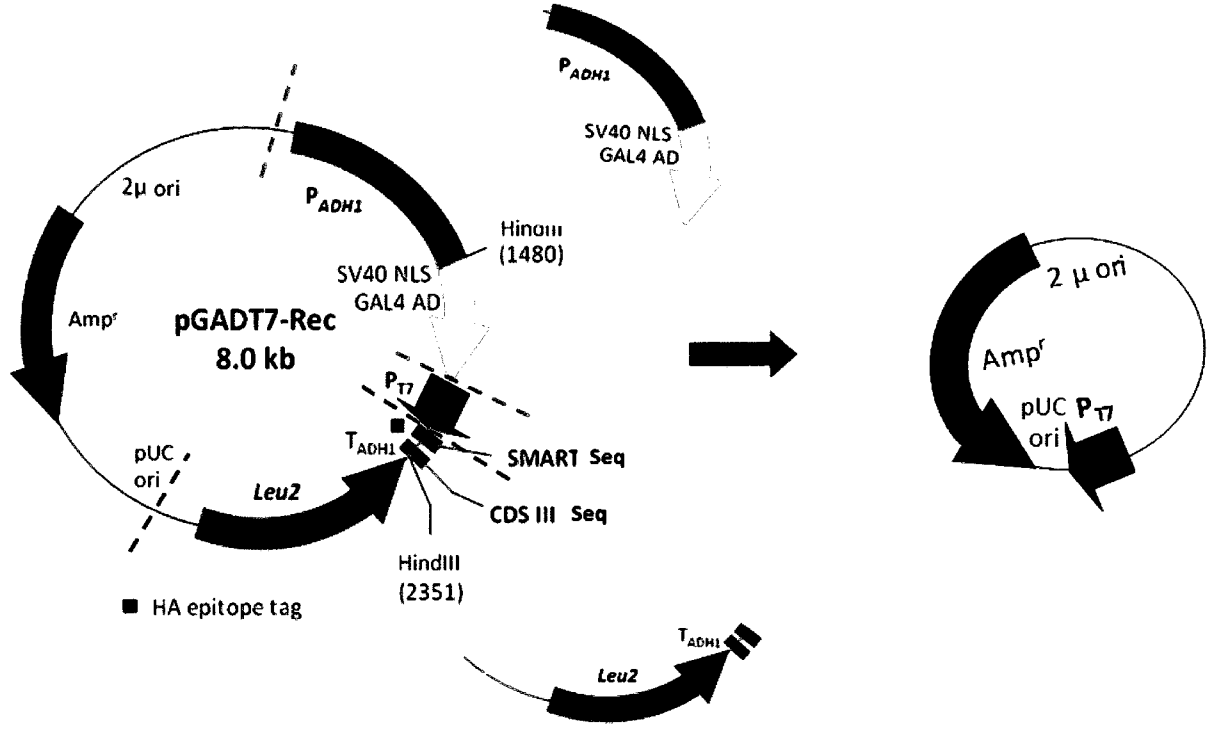

P ADH1
SV40 NLS
GAL4 AD
2μ ori
P ADH1
HindIII
(1480)
SV40 NLS
GAL4 AD
Amp r
pGADT7-Rec
8.0 kb
P T7
T ADH1
SMART Seq
pUC
ori
Leu2
CDS III Seq
HindIII
(2351)
HA epitope tag
T ADH1
Leu2
2 μ ori
Amp r
pUC P T7
ori

Figure 35. Depiction of pGADT7-Rec site-specific recombination. Sequence analysis of 2.5 Kb plasmid revealed that it was the product of site-specific recombination of the pGADT7-Rec plasmid.

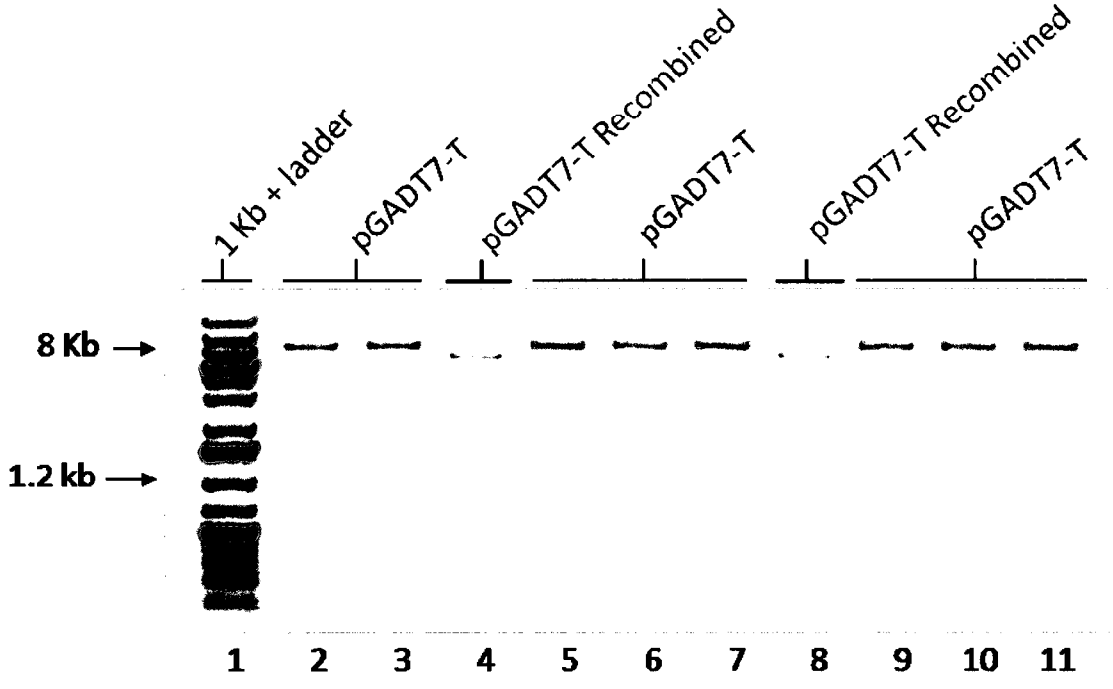

1 Kb + ladder
pGADT7-T
pGADT7-T Recombined
pGADT7-T
pGADT7-T Recombined
pGADT7-T
8 Kb
1.2 kb
1
2
3
4
5
6
7
8
9
10
11

Figure 36. pGADT7-T exhibits instability. HindIII digest of plasmid recovered from 10 unique yeast colonies transformed with pGADT7-T resulted in generation of an approximately 8 kb band (pGADT7 plasmid) and release of large T antigen (~1.2 kb) [lanes 2,3,5,6,7,9,10,11]. However, 2 out of 10 plasmids recovered demonstrated potential recombination events (lanes 4 and 8). Hind III digestion of the recombined plasmid resulted is generation of two DNA fragments that were incorrect in size (6 kb and 3 kb), as they did not correspond to the pGADT7 plasmid or the large T antigen.

Hind III digest of 10 yeast colonies expressing unique cDNA library inserts were analyzed (**Figure 37**). HindIII digest of six out of the ten clones analyzed, resulted in generation of appropriately sized DNA bands corresponding to linearlized pGADT7-Rec vector (~8 kb) and various library inserts (Figure 37: Lanes 2, 3, 5, 6, 8,and 10). However, 4 out of the 10 HindIII digested pGADT7-Rec plasmids analyzed, resulted in the same variant as pGADT7-T recombinants (Figure 37: Lanes 4 and 8), potentially demonstrating the same site-specific recombination event.

Does "false positive" auto-activate reporter genes?

In order to determine whether the "false positive" was able to induce reporter gene expression on its own or if it required the presence of the bait plasmid to produce transcriptional activation, the plasmid was transformed into AH109 alone and with pGBKT7. None of the yeast that were transformed with this false positive were able to grow on YC-W media and the plasmid was not able to auto-activate any of the reporter genes alone or with pGBKT7, pGBKT7-cPARP-1 or pGBKT7-FArtemis.

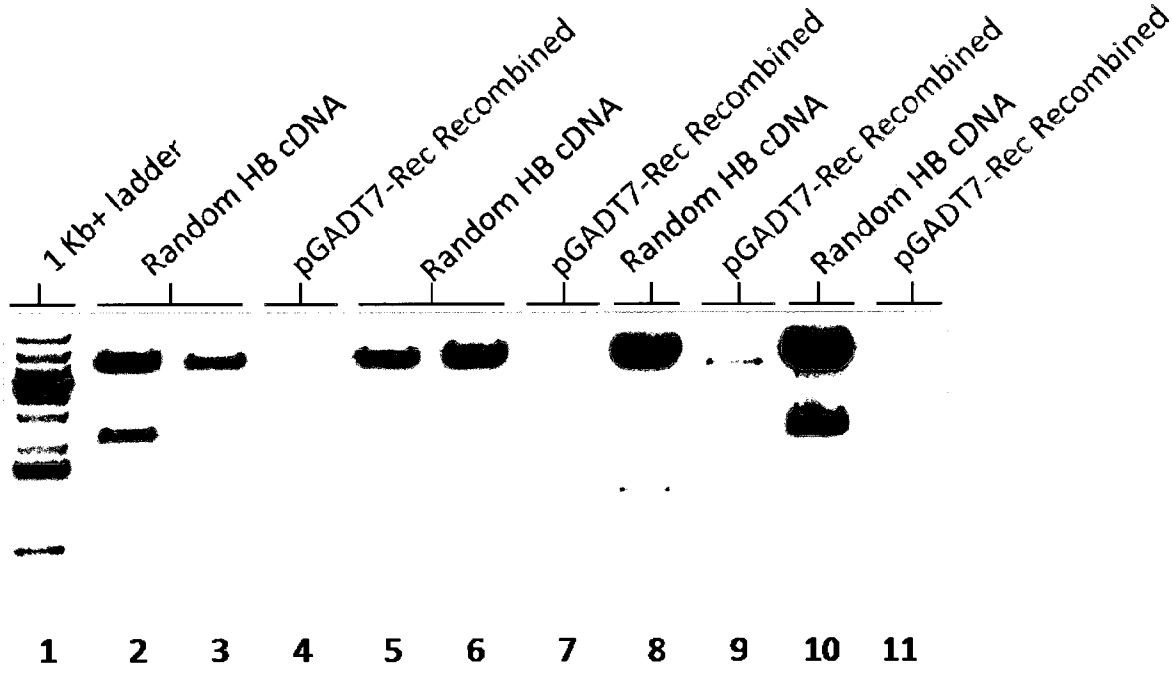

1 Kb+ ladder
Random HB cDNA
pGADT7-Rec Recombined
Random HB cDNA
pGADT7-Rec Recombined
Random HB cDNA
pGADT7-Rec Recombined
Random HB cDNA
pGADT7-Rec Recombined
1
2
3
4
5
6
7
8
9
10
11

Figure 37. pGADT7-Rec-cDNA exhibits instability. HindIII digest of plasmid recovered from 10 unique yeast colonies pre-transformed with pGADT7-Rec-cDNA resulted in generation of an approximately 8 kb band (pGADT7-Rec plasmid) and release of unique cDNA insert (Lanes 2, 3, 5, 6, 8, and10). However, 4 out of 10 plasmids demonstrated potential recombination events (lanes 4,7,9, and 11), HindIII digested of the recombined product resulted is generation of two DNA fragments that were incorrect in size (6 kb and 3 kb) and did not correspond to pGADT7-Rec or a unique cDNA insert.

DISCUSSION

Yeast Two Hybrid Screening Approach

Critical Requirements for a successful Y2H screen

A series of extensive Y2H screenings were conducted to find Artemis and PARP-1 interacting factors using both interactive mating and classical Y2H systems. Consequently, several essential parameters required for successful Y2H screening experiment have been identified and will be discussed in the following text. Primarily, it is essential that the cDNA library used to screen the bait of interest consists of high quality cDNAs and a high number of primary clones, representing the accurate complexity of the particular tissue from which it was derived. Moreover, amplification of classical cDNA libraries must be uniform to ensure the proper representation of each cDNA clone. The disadvantage of classical cDNA libraries (created by cloning pools of DNA fragments to the coding sequence of the activation domain) is that in numerous cases the hybrid protein is expressed in the wrong frame, or from the 5' to 3' un-translated regions of the mRNA (120). In fact, most classical cDNA libraries contain only one out of six fused cDNAs in the correct open reading frame; therefore, it is crucial to screen at least 6 times the number of independent clones in such cDNA libraries (120, 121). Equally important is the number of doubly positive host yeast cells that are transformed with both bait and prey plasmids. The lower the transformation or mating efficiency of the screen, the lower the number of independent cDNAs screened, and therefore, the higher the volume of yeast cells that must be plated. If the number of double positive yeast is equal to or is less than the number of independent clones in the cDNA library, thousands of large 150 mm plates would be required to plate the screen (within optimal plating density), making it practically infeasible to screen the cDNA library until saturation. Additionally, plating yeast at high densities results in growth of microsatellite

colonies that grow on surrounding dead cells and do not represent true interactions and in effect, prevent growth of colonies that express the reporter genes. Furthermore, optimal growth conditions (pH, temperature, concentration of solutes) and selection pressure (nutrient availability) must be determined prior to plating and monitored continuously to prevent induction of yeast stress responses and maintain colony viability.

Regulation of Transcription Factors

A positive signal in a Y2H screen requires the reconstitution of a two-component transcription factor complex; therefore, both fusion proteins must be expressed at sufficient levels relative to their affinity for one another and must reform an adequate amount of transcription factor complexes to activate reporter gene activity at detectable levels. If the fusion proteins are poorly expressed, or if sufficient amounts of transcription complexes are not reconstituted to overcome the activation threshold, reporter gene activity will not be detected (even if the two proteins bind and have a high affinity for each other), leading to a "false negative", where a real interaction is not detected (139). Furthermore, there is often non-specific activity associated with DBD or AD alone. A "well-behaved" bait construct can exhibit non-specific auto-activity that will activate the reporter genes at very low detectable levels, such that when expressed in yeast and patched on to selective medium, only a few colonies will be apparent rather than a solid streak.

Bait Selection and Characterization

The first step in conducting an Y2H screen is to select and characterize the bait of interest. The Y2H system, however, is not suitable for all bait proteins. This system is limited to proteins that can be localized to the nucleus, are able to stably fold in yeast cells

(undergo appropriate posttranslational modifications), and must be able to retain their activity as fusion proteins without trans-activating reporter genes. Several Artemis bait constructs were cloned in pHybLex/Zeo to optimize the probability of obtaining a suitable bait, since several previous studies, including our own laboratory have encountered difficulties expressing Artemis (unpublished data). Expression of pHybLex/Zeo-full length Artemis was not detectable in L40; this may reflect potential toxicity due to over-expression of Artemis over-expression, a previous observation described as a result of the over-expression of SNM1 proteins (69). Expressions of Artemis truncations were detectable at high levels in L40, however only pHybLex/Zeo-NArtemis was a suitable bait and did not exhibit auto-activation of reporter constructs. The C-terminus of Artemis exhibited auto-activity and was able to activate both *HIS3* and *LacZ* reporters alone and was deemed to be unsuitable as bait for the Y2H system. This was undesirable because the C-terminus of Artemis was of particular interest to us, as it is unique to Artemis and non-conserved among other members of the SNM1 family. To determine if the C-terminus of Artemis contained a region that functioned as an activation domain, we truncated the region into four separate deletion constructs in an attempt to abrogate the auto-activation capacity. None of the deletion constructs were able to abrogate the auto-activation of the full C-terminus, and hence a specific region within the polypeptide could not be credited with potentially functioning as an activation domain. Western blot analysis of the bait constructs using anti-LexA antibody revealed that, unlike pHybLex/Zeo-NArtemis, both pHybLex/Zeo-CArtemis (Figure 20) and pHybLex/Zeo-CArtemis deletion constructs (Figure 22), showed expression of LexA DBD both fused and un-fused to Artemis alone (LexA at 32-kDa). It has been reported that the activity of un-fused LexA from pHybLex/Zeo is high enough to weakly activated *HIS3* and *LacZ* reporters, and that protein fusions to LexA generally decrease this

background activation (128). The auto-activity of pHybLex/Zeo-CArtemis and the deletion constructs generated much higher auto-activity than pHybLex/Zeo alone. When L40 expressing pHybLex/Zeo was patched on His^- plates, only a few colonies were visible. Conversely, when L40 expressing PHybLex/Zeo-CArtemis was patched on His^- plates, a solid streak was apparent. It is possible that fusion of C-terminus Artemis to LexA generates very high activities by increasing the intrinsic auto-activity of the LexA DBD, resulting in degradation of the fusion protein, marked by detection of un-fused LexA as well as LexA-C terminus Artemis (Figures 20 and 22). The same Artemis truncations were used to construct bait-DBD fusion in pGBKT7 in interactive mating Y2H system. Intriguingly, expression of pGBKT7-Full length Artemis was detected in AH109 in contrast to L40. Once more, the Artemis C-terminus construct resulted in high auto-activity and was not suitable for screening. Both pGBKT7-NArtemis and pGBKT7-FArtemis were well expressed and exhibited low levels of auto-activity (similar to that of pHybLex/Zeo alone). However, pGBKT7-FArtemis was the bait of choice, since this construct includes the C-terminus of Artemis, a region of great interest in our study.

The PARP-1 bait construct was designed omitting the N-terminal DNA-binding domain (aa1-250) to abrogate auto-activity. pGBKT7-cPARP-1 expression was detected in AH109, and resulted in very low background activation of reporter genes (similar to that detected in AH109 containing pGBKT7 alone) and was therefore deemed an appropriate bait construct for the Y2H analysis.

pHybLex/Zeo-NArtemis screening of Hybrid Hunter^{TM} Human Lung cDNA library

The primary transformation efficiency of this screen was 10^7, and approximately $3.5x10^7$ double transformants were screened, resulting in growth of 15 colonies (positive for

presence of the bait and prey vectors and potential interaction). The transformation protocol used for this screen was optimized to yield approximately 10 to 100 million double positives (128). The primary transformation efficiency was several orders of magnitude higher than the number of independent clones in the cDNA library ($5.96x10^6$); therefore, the cDNA library was screened until saturation. Typically, screens with high transformation efficiencies result in growth of 100-300 colonies. Growth of only 15 colonies was the initial indication that an adequate number of interactions were not assayed to identify biologically significant interactions. Further examination of the putative interactors revealed that only five of the cDNAs demonstrated specific interactions with pHybLex/Zeo-NArtemis. BLAST analysis of the putative interactors revealed that none of the cDNA inserts contained an ORF in any of the possible frames. The pYESTrp2+cDNAs all produced short peptides that demonstrated specific interactions with pHybLex/Zeo-NArtemis. However, since NCBI blast identification revealed that the cDNA sequences were in fact located in non-protein encoding regions, it was concluded that these putative interactors do not represent biologically relevant interactions (these Y2H interactions are considered to be false positives). Analysis of the short stretches of amino acid sequences (translated in the same frame as LexA DBD), revealed that all of these polypeptide stretches contained both hydrophobic and acidic amino acid residues (Table 4). Prevailing evidence has shown that interspersed acidic and hydrophobic amino acids are critical features of activating domains of yeast transcription factors (140, 141, 142). It is feasible that fusing a polypeptide that mimics as an additional activation domain to the existing activation domain (V5 AD) of the prey plasmid increases the weak intrinsic transcriptional auto-activity of pYESTrp2-cDNA, such that it becomes a stronger activation domain. However, each pYESTrp2-cDNA was assessed for auto-activation after retrieval from yeast, and could not trans-activate reporter

gene activity alone in yeast host. Furthermore, co-transformation of the retrieved pYESTrp2-cDNA with pHybLex/Zeo (empty bait vector) could not auto-activate reporter gene activity. Yeast transformed with the pYESTrp2-cDNA could grow on selective medium when and only co-transformed with pHybLex/Zeo-NArtemis specifically. There are two possibilities for this observation. The first is that the short polypeptides produced (although not biologically relevant), are indeed interacting with N-Artemis specifically within the yeast cell. The alternative possibility is that producing additional activating regions coupled to the V5 AD increases the overall auto-activity of the hybrid and stabilizes the transcriptional machinery, such that when present with LexA DBD-NArtemis, more transcription factors are reconstituted, which ultimately overcomes the specific threshold required to drive transcription at detectable levels. Alone however, the auto-activity of pYESTrp2 or pYESTrp2-cDNA with the pHybLex/Zeo would not overcome the threshold required to activate transcription on a level that would result in detectable expression of reporter genes. Various studies have shown that yeast activating regions function by recruiting transcriptional machinery to DNA (143). According to this principle, activating regions function like "glue" to recruit transcriptional machinery to DNA (143). In line with this theory, it is possible that the accumulation of an additional activation domain to LexA DBD-NArtemis and V5-cDNA (from the short polypeptides) would result in higher recruitment of transcriptional machinery to DNA, leading to stronger gene activation and detectable expression of reporter genes, without a true interaction between library DNA and N-terminus Artemis.

Because all five putative interactors that were identified encoded sequences that contained no protein ORF and blast identification revealed homology only to non-protein encoding regions, the quality of the Hybrid Hunter^{TM} Human Lung cDNA library was

assessed. Sequencing analysis from six random cDNA inserts from the library revealed that 5/6 clones contained no ORF in any possible frame. BLAST search using nucleotide query of cDNA sequences against the protein database resulted in "no significant similarity found" for 4/6 cDNA clones. BLAST search using nucleotide query of the sequences against the nucleotide database identified the randomly selected clones as non-protein encoding sequences. A cDNA library is theoretically produced from fully transcribed mRNA found in the nucleus, and therefore should encode for only expressed genes in the tissue from which it was derived. Although most classical cDNA libraries contain only 1 out of 6 clones in the proper frame, the presence of non-protein encoding sequences in a cDNA library is not anticipated. It is likely that this cDNA library was generated from poor quality RNA preparation or poly(A)$^+$ RNA from which double stranded DNA copies were generated. Because a high quality cDNA library is essential to yield high quality data, it is important to assess the quality of a newly generated cDNA library before conducting exhaustive screening experiments.

pHybLex/Zeo-NArtemis screening of DupLEX-A^{TM} Human Fetal Brain cDNA library

The DupLEX-A^{TM} human fetal brain cDNA library was chosen for subsequent screenings with pHybLex/Zeo-NArtemis. Human fetal brain tissue has been shown to be positive for expression of Artemis (69), and this cDNA library was guaranteed to encode 1-2/6 full length cDNA clones, and was compatible with the Hybrid Hunter system. The quality of the cDNA library was assessed by sequencing six random clones prior to library screening. All 6 clone sequences encoded protein sequences; however, only 2/6 cDNA clones contained proteins in frame with LexA DBD. The primary transformation efficiency

of this screen was determined to be $2.0x10^7$, and approximately $2.1x10^7$ double transformants were screened. The primary transformation efficiency was several orders of magnitude higher than the number of independent clones in the cDNA library ($3.5x10^6$); therefore, the cDNA library was screened until saturation resulting in growth of 23 colonies (Table 5). Seven of these clones were identified as putative interactors as they did not auto-active the reporter genes alone or when transformed with pHybLex/Zeo.

BLAST Identification of Putative Interactors

BLAST identification of the cDNA insert retrieved from colony #1 identified partial mRNA encoding the NADH dehydrogenase subunit 1 protein (Table 6). NADH dehydrogenase subunit 1 is a mitochondrial protein imbedded within the cristae of the mitochondrion, and catalyses the transfer of electrons from NADH to coenzyme Q in the electron transport chain (144). Another mitochondrial protein was identified from colony #9 encoding cytochrome c oxidase subunit I (Table 6). This protein is a multi-chained transmembrane protein embedded within the cristae of the mitochondrion, and functions as a terminal enzyme of the mitochondrial respiratory chain (145). Mitochondrial proteins are one of the common re-occurring false positives encountered in the Y2H system (120). Artemis is a nuclear protein, while NADH dehydrogenase and cytochrome c oxidase subunit I are expressed in the mitochondrial membrane. Although Artemis has only been characterized as a nuclear protein, one cannot exclude the possibility that Artemis may also be expressed in other subcellular compartments of the cell that have thus far remains uncharacterized. Therefore, it cannot be concluded with absolute certainty non-nuclar proteins such as mitochondrial proteins or commonly-encountered false positives are in fact

representing biological false positive interactors. However, it can be stated that the resulting interaction between such proteins and Artemis proteins expressed simultaneously within the yeast cell likely represent "biologically" false positive interactions. In the cell these two proteins would not physically interact under physiological conditions within the cell as they are likely never expressed in the same organelle to produce a genuine interaction. Other such "biological" false positives were identified, including the TIMP3 metallopeptidase inhibitor, a protein expressed in the extracellular matrix (ECM) of the retina (Table 6). Other such examples of "biological" false positives include interactions that occur between two proteins that are not physiologically expressed at the same time (temporally) or in the same tissue (spatially). These false positives are one of the major setbacks of the Y2H system and are not feasible to avoid (121).

Another putative binding partner identified from the cDNA clone isolated from colony #13 was the heat shock protein (hsp90) (Table 6). Hsp90 is a molecular chaperone and its subcellular localization includes the cytoplasm, the endoplasmic reticulum, and the mitochondrion (146). Hsp90 is one the most abundantly expressed proteins in a cell, and consequently one of the highest encountered false positives reported in Y2H systems was therefore omitted from further characterization (121, 146). The only putative interactor discovered that was a nuclear protein was identified from colony #14 corresponding to partial mRNAs of the ATRX protein (Table 6). The cDNA retrieved from colony #14 shares 91% sequence homology with the N terminal region (837-927aa) of the ATRX protein (Figure 26). At first glance, this finding was particularly interesting since identifying a nuclear protein with helicase activity is a promising binding partner for Artemis.

Biological significance of the ATRX protein

The ATRX protein was identified in 1995 as a novel X-encoded protein whose mutations gave rise to a characteristic form of a syndromal mental retardation associated with [alpha] thalassemia known as alpha thalassemia/mental retardation syndrome, X-linked (ATR-X syndrome) (147). The ATRX protein was classified as a new member of the SNF2 family of proteins, which encode core enzymatic subunits of chromatin-remodeling protein complexes, including ATPase and helicase domains (147). Based on additional sequence homology, this protein was assigned to a subgroup within the SNF2 family consisting of a group of proteins involved in cellular processes such as transcriptional regulation, recombination, DNA repair and replication (148). Although the physiological function of ATRX has not been fully elucidated, it is known that ATRX localizes at pericentromeric heterochromatin, and that ATRX mutations have been correlated with changes in DNA methylation patterns. Furthermore, it has been found that ATRX forms a complex with the transcription cofactor, Daxx, and displays chromatin-remodeling activities (149). Chromatin-remodeling proteins are involved in the proper regulation of gene expression in eukaryotes via an effect on chromatin structure/function (150). ATRX has also been identified in Y2H screens as a binding partner of the heterochromatin protein 1 (HP1) and the polycomb group protein EZH2 (enhancer of zeste homolog 2) (151, 152). HP1 proteins are chromatin associated proteins involved in the cellular response to DNA damage, they regulate gene repression by promoting heterochromatinization of euchromatin in response to DNA damage which triggers the DNA damage response (153). The polycomb group protein EZH2 is the human homologue of the Drosophila protein Enancer Zeste (E(z)), both contain a SET (Su(var)3-9, Enhancer-of-zeste, Trithorax) domain which is found in

chromatin-associated regulators of gene expression (154). Polycomb group EZH2 has been implicated in the regulation of homeotic gene expression, and dysregulation of EZH2 leads to dysregulation of the transcriptional maintenance system which can lead to malignancy (154).

As previously mentioned, recent studies have speculated that the endonucleolytic activity of Artemis may be required to act upon specific regions of DNA that are heterochromatinized. Identifying a binding partner for Artemis with helicase activity would support this theory; given that unwinding of heterochromatinized DNA would be required prior to the processing activity of Artemis. Several members of the SNF2 family of helicases possess chromatin-remodeling activity in regulation of DNA repair and transcriptional maintenance. HP1 was identified as a binding partner of ATRX, and since it is important for DNA repair, it may link the helicase activity of ATR in chromatin-remodeling in DNA repair and may therefore share functional relevance with Artemis (153). The polycomb group protein EZH2 was also identified as a binding partner for ATRX, a protein that is involved solely in homeotic transcriptional maintenance (154). It is possible that ATRX has chromatin-remodelling functions in the regulation of both DNA repair and transcriptional maintenance; ATRX and Artemis could potentially be involved in the same branch(es) of DNA damage response.

A closer examination of the retrieved cDNA encoding ATRX revealed that the ORF of the putative interactor was out-of-frame with the V5 activation domain of the pGJ4-5 vector, rather it was cloned in the +2 reading frame of V5. Among the most important criteria of the Y2H screen is to ensure that the pGJ4-5 AD vector is selected because it encodes a genuine fusion protein, and not a short irrelevant polypeptide. As previously mentioned, another main source of false-positives routinely encountered in Y2H systems are

out-of-frame activation domain and protein fusions encoding short peptide fragments (121). It is significant to note that previous studies have found and confirmed real interactions with out-of-frame prey cDNA inserts; in rare cases it is possible for yeast to generate fusion proteins in +1 or -1 frames via a "frame shifting" event, potentially due to ribosome "slippage" or mRNA peptidyl detachment and re-pairing (155, 156, 157, 158). Therefore, one cannot automatically exclude out-of-frame fusion proteins in the identification of potentially promising putative interactors from further investigation. However, frame shifting would have to occur downstream of the activation domain, as to generate a fusion of polypeptide mRNA in frame with activation domain in order to reconstitute transcriptional machinery and to activate the reporter gene (s). For example, if a putative interactor encoded a polypeptide within +2 frame of the activation domain, a frame shift event upstream of the activation domain +1 would generate an in frame polypeptide, but an out of frame activation domain. Therefore, to confirm that reporter gene expression is the result of an interaction between bait and a "frame shifted" prey mRNA, one must clone the cDNA into prey plasmid in the proper frame and re-test for interaction. It must also be considered that the cDNA retrieved sharing 91% sequence homology with the ATRX protein, represented a small region (aa 837-927) of the large protein (2494 aa) (Figure 26). Furthermore, this region of homology does not represent a full length cDNA or an independent domain within the protein, but a small portion of the large N-terminus. Therefore, it is less probable that this small region can exist in stable manner independent of the rest of the protein or fold appropriately to adopt its correct 3D structure when expressed as fusion to GAL4 AD. Hence, despite the promising likelihood of an interaction between the ATRX protein and Artemis, due to the fact that they are both nuclear proteins and may share functional relevance, (and despite the possibility of a genuine physiological interaction

between Artemis and ATRX), the interaction identified within this experiment may represents a false positive (further elucidation is required).

pGBKT7-PARP-1 Screening of Pre-transformed Matchmaker™ HB cDNA library

The mating efficiency of this screen was 0.7%, and only 57% of the total number of independent clones in the HB cDNA library were screened. The result was the growth of 230 colonies on YC-WHAL medium, 129 which exhibited β-galactosidase activity. Analysis of library cDNA retrieved from the 129 colonies, revealed the presence of a re-occurring "false positive" (2.5 kb recombinant plasmid), polluting the results of the screen. The correct pGADT7-Rec vector was retrieved from only 14 colonies as a sub-population along with the recombinant plasmid (Figure 30). The recombinant plasmid was retrieved exclusively from the remaining 115 colonies (without any trace of the pGADT7-Rec) and remained un-linearized following HindIII restriction digest (Figure 32), unlike the intact pGADT7-Rec plasmid which contains two HindIII sites (Figure 18). Intriguingly, the correct and intact pGBKT7-cPARP-1 was retrieved from the same colonies using the same plasmid isolation techniques (Figure 33). Furthermore, the correct and intact pGADT7-T plasmid (positive control) was appropriately retrieved from control strain (Figure 33). This observation led us to believe that there was an intra-plasmid site-specific recombination event exclusively obliterating the pGADT7-Rec vector. Sequence analysis of the recombinant vector revealed that it was indeed the product of an intra-plasmid site-specific recombination event of the pGADT7-Rec vector. It appeared that the recombination event had spliced out two major regions of the plasmid, including the site encoding the library

cDNA insert, leaving intact the 2-μm and pUC origins of replications, the T7 promoter and the ampicillin resistance gene (Figure 35).

A study conducted in 1998 by El Housni and colleagues reported the finding of a similar recombination event in their yeast two-hybrid screening experiment, in which site-specific rearranged bait constructs were continually selected because they produced functional GAL4 activity. El Housni and colleagues explained that although it was likely that various other recombination events took place, only reporter gene activating recombinants were continuously selected (159). Our next objective became to determine the cause of this recombination event that we had encountered in our own study. I hypothesized that either the yeast were promoting this recombination event because the resulting plasmid sequence was producing GAL4 transcriptional activity, or that the recombination was a random event induced by the intrinsic instability of pGADT7-Rec itself. To investigate, we transformed the recombinant plasmid into AH109 alone, to determine if it was in fact capable of auto-activating reporter genes alone. Indeed, AH109 transformed with the recombinant plasmid could not grow on YC-L medium (as it lacks the *LEU2* auxotrophic marker) and could not auto-activate any of the reporter genes alone or with presence of pGBKT7-cPARP-1. We rationalized that host yeast containing recombined plasmids were able to grow on nutrient deficient medium because the regions of pGADT7-Rec that were spliced out (auxotrophic markers) were likely integrated into the yeast genome; thus, evading the requirement of maintaining a stable intact plasmid. In order to assess the instability of the pGADT7-Rec itself, we recovered ten random library clones from the host yeast pre-transformed with HB cDNA library. Restriction analysis revealed that six out of ten library plasmids represented the correct pGADT7-Rec vector, while the other four appeared to be recombinants of pGADT7-Rec (different recombination variant than "false positive"

encountered in library screen) (Figure 37). To determine if this intrinsic instability was exclusive to pGADT7-Rec (the engineered version of pGADT7), we also analyzed the stability of pGADT7-T (pGADT7 expressing large T antigen, supplied as prey for positive control). Restriction analysis of pGADT7-T revealed that two out of ten plasmids recovered from positive control strain did harbour the same recombinant variant as those recovered from ten random HB library plasmids (Figure 36). Therefore, the pGADT7 and pGADT7-Rec vectors both exhibited intrinsic instability when selected with YC-L medium, without strong selection pressure or environmental stress (low pH, inclusion of 3-AT, and growth on QDO medium). Both pGADT7-Rec and pGADT7 replicate from the 2-μm origin of replication in yeast and encode the *LEU2* auxotrophic marker for selection in yeast (134). The *LEU2* auxotrophic marker is expressed by the defective leucine marker *leu2*-d, a useful variant of *LEU2* (160). The *leu2*-d lacks the *LEU2* promotor and therefore is expressed at very low levels, and in order to provide sufficient levels of β-isopropylmalate dehydrogenase (the *LEU2* product to facilitate growth), vectors carrying *leu2*-d for selection must be present at very high copy numbers (>200 copies per cell) (160, 161). Studies have shown that strong selection pressure against transformants that harbour plasmid DNA carrying the 2-μm origin of replication and the *LEU2* gene fragment are expressed at insufficiently high copy numbers and result in instability that may lead to plasmid rearrangement events (160, 162). It is significant to note that yeast cells harbour an endogenous 2-μm plasmid that is autonomously replicating at 50-100 copies per diploid cell (163, 164). Recombinant plasmids such as pGADT7 and pGADT7-Rec that replicate from the 2-μm origin of replication use the same replication and propagation systems as the endogenous yeast 2-μm plasmid (160, 161, 163). Transformed cells that contain such a recombinant plasmid in addition to the endogenous plasmid (depending on the type of vector and DNA insert) may

exhibit segregation and intra-plasmid site-specific recombination events varying between 1 and 5% loss of original plasmid per generation (160). Furthermore, it has been shown that plasmid curing (elimination of plasmid DNA from cell) and recombination events were prevalent under transformation conditions selecting for the *LEU2* gene (160, 165, 166). Furthermore, the selection pressure and environmental stress of the QDO (YC-WHAL) medium may have further enhanced this recombination event. Like other single-celled organisms, yeast live freely in nature and encounter large variations in their physical environment. Therefore, in order to survive and proliferate, single-celled organisms have evolved rapid response and powerful adaptation mechanisms to stress and sudden environmental changes (167). It is likely that the growth and plating conditions in the screen imposed environmental stress conditions upon the yeast. Such stress conditions include nutrient availability (stringent QDO medium), presence of toxic chemical agents (inclusion of 3-AT), and changes in pH prior to mating (pre-incubation of yeast in low pH prior to mating). It is possible that the recombination events were induced as a result of yeast stress response to plating and growth conditions. Taking into consideration these previous findings on the nature of plasmids harbouring the 2-μm ori and the *LEU2* gene fragment, and the intrinsic instability of the library vector alone, one can infer that the pGADT7-Rec was susceptible to recombination events. However additional stress responses evoked by yeast host likely further evoked the recombination to the scale that was encountered during the screening. In addition, these recombination events were likely favoured growth conditions, as yeast no longer depended on the reconstitution of the AD and DBD in order to survive; therefore, yeast harbouring this recombination were continuously selected throughout the screen. This hypothesis is favoured by our finding that the recombination observed during the Y2H screen plated on high stringency medium (YC-WHAL medium) was evident at

much higher frequencies than that observed via the segregation of pGADT7-Rec library clones (4/10 plasmids recombined) and pGADT7-T (2/10 plasmids recombined) which were plated on less stringent medium (YC-L medium). There are however various publications using the Pre-transformed Matchmaker^{TM} HB cDNA library, yielding positive results. The main difference in procedure between our screening and manufacturer's protocol was pre-incubation of yeast in low pH medium. This modification led to a ten fold increase in mating efficiency, but may have polluted the overall results of the screen. However, we are likely not the only group to have encountered instability or recombination events with the pGADT7-Rec or the Matchmaker pre-transformed libraries, as both the vector and the system are no longer commercially available.

Mating efficiency

The mating efficiencies of both bait strains (AH109 [pGBKT7-cPARP-1] and AH109 [pGBKT7-FArtemis]) with the library-expressing strain (Y187 [pGADT7-library cDNA]) was significantly lower in comparison to the mating efficiency of the control strains (AH109 [pGBKT7-p53] with Y187 [pGADT7-T]) (Table 7). Mating efficiencies of both bait strains (AH109 [pGBKT7-cPARP-1] and AH109 [pGBKT7-FArtemis]) with Y187 [pGADT7-T] versus Y187 [pGADT7-Rec-librarycDNA], were both comparably low compared to mating efficiency of control strains (Table 7). This is an indication that expression of PARP-1 and Artemis proteins are to some extent toxic to yeast, or may interfere with the mating process. Over-expression of a protein with nuclease activity such as Artemis protein or PARP-1 (a cellular indicator of DNA damage) may impose toxic side effects within the yeast and induce stress response(s).

BLAST Identification of Putative Interactors

The intact pGADT7-Rec vector was retrieved from only 14 of the 115 colonies (Table 9). BLAST identification of the cDNA inserts retrieved from colonies 5 and 41 revealed the presence of a long polypeptide encoding the hemoglobin α-1 globin chain protein, and that cDNA retrieved from colonies 48 and 103 revealed the presence of the hemoglobin β chain variant protein (Table 9). The α-1 and β globin polypeptide chains are subunits that assemble the hemoglobin protein– the iron-containing oxygen-transport metalloprotein in the red blood cells of vertebrates (168). Before assembly into hemoglobin, these globin chains are synthesized by ribosomes in the cytosol. The immunoglobulin

lambda heavy chain polypeptide (retrieved from colony 138) is the large polypeptide subunit of an immunoglobulin, found in the blood and other bodily fluids of vertebrates (169) (Table 9). However, PARP-1 is a nuclear protein, and hence all PARP-1 binding partners will also be expressed in the nucleus. Therefore, PARP-1 interactions with immunoglobulin lambda heavy chain and hemoglobin polypeptides represent "biological" false positive interactions. Two of the putative interactors identified were nuclear proteins, the cDNA inserts encoding CASP-7 and "unknown protein", were retrieved from colonies 13 and 28 respectively (Table 9). Both of these proteins are promising binding partners for PARP-1, the biological significance of CASP-7 and the "unknown protein" and the characterization of their retrieved cDNAs are discussed in the text below.

Biological Significance of Caspase-7

At first glance, CASP-7 is a very promising binding partner for PARP-1, as both proteins have implicated functions in the same apoptotic pathway. Caspase-7 is a member of a family of cysteine protease caspases (14 members in mammals), which coordinate apoptosis, and are activated in response to pro-apoptotic stimuli (170). Caspases involved in apoptosis are divided into two functional sub-groups: upstream or "initiator" caspases (caspase-2, -8, -9 ,-10), are those that are involved in initiating the caspase cascade upon the receipt of stimuli, while downstream, or "effectors", caspases (caspase-3, -6, -7) are responsible for the annihilation of the cell (170, 171). The initiator caspases appear to be more specific proteases that cleave only a few proteins, their own precursors and the effector caspases. Effector proteases are responsible for the majority of proteolysis that is carried out during apoptosis ultimately resulting in the morphological and biochemical characteristics of

apoptosis (ie. DNA fragmentation) (171). Caspase cleavage can inactivate a subgroup of substrates such as PARP, nuclear lamin, the Bcl-2 (B-cell lymphoma 2) family of genes invovled in regulation of mitochondrial membrane permeability, and β-catenin (170). Caspases can also activate a different subgroup of substrates including MEKK1 (MAP/ERK kinase kinase 1), p21-activated kinase, protein kinase Cdk and the ICAD/CAD (Inhibitor of caspase-activated DNase/ caspase-activated DNase) complex in the apoptotic pathway (172). Caspase-3 and caspases-7 are the most closely related caspases in terms of sequence similarity and substrate specificity compared to all other mammalian caspases (171). The proteolytic specificity of caspases-3 and caspases-7 have been shown to be virtually indistinguishable; therefore, it is widely accepted that caspases-3 and -7 are redundant enzymes in term of their proteolytic activity and specificity during apoptosis (171, 173). However, caspase-3 and -7 are differentially regulated in subcellular localization during apoptosis, and the role of caspase-3 has been more clarified in apoptosis execution (173). Caspases are synthesized as inactive precursors and become activated with the receipt of apoptotic stimuli via proteolytic processing at conserved aspartic residues to produce two subunits, large and small, that dimerize to form the active enzyme (170). Effector caspases cannot undergo autocatalytic processing alone and require "initiator" caspases for processing, they must be cleaved between NH2-terminal of the large subunit (p20) and the COOH-terminal of the small subunit (p10) to be activated (Figure 31). The small and large subunits then combine into a catalytically active heterodimer (173). In addition, ectopic expression of the short subunit caspases-3 and -7 precursors in mammalian cells is not sufficient to execute apoptosis, suggesting that the presence of their large subunit is necessary for proper activation. To surmount the inactivity of *in vitro* translated wild type effector caspases, Srinivasual *et al* engineered contiguous caspase-3 and -6 molecules in

which the short subunit was cloned in-frame N terminal to the large subunit (reverese-caspases), and a cleavage site (i.e., DEVDG in the case of caspase-3) was introduced between the two subunits (174). In this manner, and in contrast to their wildtype counterparts, these engineered reverse-caspases were able to undergo auto-catalysis and fold into active conformation, inducing apoptosis when over-expressed in cells (174). This reverse-caspase engineering was employed by Araya *et al* who used reverse-caspase-7 as bait in a Y2H screening to identify caspase-7 interacting proteins (172). It was found that reverse-caspase-7 successfully folded into active conformation and was autoprocessed in an *in vitro* translation system (172). Reverse-caspase-7 was used to screen a mouse adult brain cDNA library to obtain 110 positives clones by screening $2x10^6$ transformants (172). In addition to the previously identified substrates of caspase-7, Araya *et al* identified the proteosome activator PA28γ (172).

Intriguingly, PARP-1 contains a caspase-3 and caspase-7 cleavage site within its NLS site (aa202-231); however this cleavage site was excluded from the bait construct. Furthermore, the caspase-7 cDNA insert is fused in the +1 frame of the GAL4 AD. There are two possibilities explaining the observed association, either CASP-7 is translated in the +1 frame and in fact represents a true interaction with PARP-1, or the interaction occurs with the polypeptide generated in the correct frame. Furthermore the cDNA retrieved is merely a partial of the caspase-7 (PARP-1-caspase-7) protein, bearing 82% homology to the C-terminal portion of the p20 subunit (aa115-198) of the CASP-7 protein (Figure 33). This region of homology represents merely a segment of the CASP-7 p20 domain, constitutes a portion of the catalytic core, and alone is not sufficient to undergo auto-catalysis to adopt the proper active conformation. If the cDNA had encoded an independent domain or the entire catalytic core, it would increase the likelihood of a true interaction. Moreover, the cleavage

site of caspase-7 was excluded from the bait construct; therefore, a proteolytic reaction does not occur with co-transformation of the bait (pGBKT7-cPARP-1) and prey (pGADT7-par-caspase-7). Therefore, in the case that the PARP-1-caspase-7 is translated in the +1 frame, the interaction observed is the result of a specific association between the C-terminus of PARP-1 and partial region of caspase-7, and not as a result of the characterized proteolytic reaction. It is possible that these regions interact because alone they are sufficient to exert the same affinity for one another as their wildtype counterparts under the event of a proteolytic reaction; however, this is unlikely since the PARP-1-caspase-7 would not be physically folded in its proper active three dimensional conformation. It is significant to note that sometimes an interaction between proteins or polypeptides that are over expressed within a cell may occur not because of a physiological interaction, but because it is energetically more favorable and may render stability to the protein structure or assembly (175).

Protein surfaces are irregular; this is what enables proteins to bind to specific ligands or to associate specifically with other proteins. However, the binding affinity between one protein surface and another depends on various parameters. There are energetic contributions exerted by interface amino acids (hydrophobicity and polarity) that change the binding affinity of two proteins by altering the number of hydrogen bonds and van der Waal interactions (176). The stronger the complementarity (between interface residues), the more energetically favorable the interaction; and therefore, the higher the binding affinity (176, 177). Surface complementarity between two proteins ensures that as many as possible van der Waals contacts are formed and that hydrogen bonds are formed at the interface between two molecule pairs instead of forming hydrogen bonds to water (177). Given the three dimensional structures of the proteins, or assembly of the amino acids, various interactions

between disulfide bonds, hydrophobic residues, ionic interactions, hydrogen bonds, aromatic-aromatic interactions, and aromatic-sulphur interactions could produce an interaction at a given threshold detectable by reporter gene activity (175, 177).

Biological Significance of the "Unknown Protein"

The retrieved polypeptide encoding a partial region of the "unknown protein" shared 81% sequence homology (aa1073-1192) to the most C-terminal region of the protein (Figure 33), and was cloned within a +2 frame of the GAL4 AD. The "unknown protein" contains an N-terminal region which classifies the protein as a member of the endonuclease/exonuclease/phosphatase family (137). Members of the endonuclease/exonuclease/phosphatase family share a common structural domain that is also found among a large family or proteins that include Mg^{2+}-dependent endonucleases involved in DNA repair, as well as a diverse array of protein phosphatases involved in signal transduction and cell-cycle regulation (178). The reverse transcriptase-like domain is common to mobile genetic elements (retroelements) including retroviruses, retrotransposons, group II introns and hepadnaviruses (137, 179). These retroelements are mobile DNA species that are integrated into various positions of the eukaryotic genome (in humans they make up a heavy load of approximately 42% of the genome), and are able to translocate to different regions within the genome (180). The majority of these mobile elements have a reverse transcriptase-like domain, as they require reverse transcription for their replicate transposition. Retroelements translocate by being transcribed into single stranded RNA, and then reverse transcribed into double stranded DNA and are integrated back into the host genome, duplicating the element (181). This "unknown" protein also encodes the

RecF/RecN/SMC N terminal domain exclusive to a superfamily of proteins with ATP-binding domains, involved in chromatin and DNA dynamics (137). The SMC proteins are essential for successful chromosome transmission during replication and segregation of the genome in all organisms. This family also includes the RecF and RecN proteins that are involved in DNA metabolism and recombination (137). Based on sequence and domain homology it is likely that the "unknown protein" has implications in structural maintenance of chromosomes (DNA metabolism and recombination). A protein with ATPase and reverse transcriptase activity would be a likely binding partner for PARP-1— a chromatin associated protein. All of the domains that are responsible for the enzymatic function of this protein are encoded by the N-terminus of the "unknown" protein. However, the cDNA insert is encoded in +2 frame of the GAL4 AD and shares homology only with the most C-terminus. The region with homology to the "unknown protein" isolated from the cDNA represents approximately 10% of the entire protein and does not encode an independent domain with a known function or independent folding capability. Ultimately there are two possibilities: A) PARP-1 is interacting with the polypeptide that is translated with the same frame as the GAL 4 AD and represents a biological false positive and not a genuine interaction with PARP-1, B) or that the depicted region of the "unknown" protein (Figure 33) is indeed translated in the +2 frame and is interacting with PARP-1 which represent a potential biological interaction. Further elucidation is required.

pGBKT7-FArtemis Screening of Pre-transformed Matchmaker^{TM} HB cDNA library

The results of the pGBKT7-FArtemis screen were very much similar to that of the pGBKT7-cPARP-1 screen. Both screens had low mating efficiencies (0.6-0.7%) and the results of both screens were polluted with the site-specific recombination of the pGADT7-

Rec vector. From the 197 colonies that were obtained from this screen, the proper pGAD7-Rec vector was only isolated from three colonies. Two of the retrieved cDNAs (isolated from colonies # 47 and # 55) encoded hemoglobin polypeptides, as encountered in the previous screen representing false positive interactions (Table 10). Hemoglobin and immunoglobin polypeptides are likely saturated in the cDNA library as they are expressed in higher abundance in human bone marrow tissue, a hematopoietic tissue which gives rise to blood cells and immunoglobins (182). The third protein identified in this screen as a putative interactor was CAP1 whose partial cDNA was isolated from colony #26 (Table 10). CAP1 (adenylate cyclase-associated protein 1) is related to the yeast CAP protein, a cytoplasmic protein which is involved in the cyclic AMP pathway (183). Since CAP1 is a cytoplasmic protein and all Artemis binding proteins are nuclear proteins, further characterization of this interaction is not required as it represents a "biological" false positive.

Y2H Screening with Artemis

Screening with the first system did not yield successful results due to the poor quality of the Hybrid Hunter™ human lung cDNA library. Subsequent screens with pHybLex/Zeo-NArtemis with the same system were conducted using the DupLEX-A™ Human Fetal Brain cDNA library. Screening with this library resulted in the growth of only 15 colonies; and the putative interactors identified were accounted for as a result of "biological false positives" and the presence of one potential interactor (ATRX protein) which was cloned in out-of-frame with the V5-AD of pGJ4-5 vector. More extensive screening using this system was necessary for further Y2H analysis in search of a promising hit; however, the use

of Zeocin™ was found to be too costly for subsequent screenings with the Hybrid Hunter system. The Matchmaker™ pretransformed interaction mating sytsem was used for further analysis of Artemis and PARP-1 binding factors. Although this system seems initially promising, screenings were hindered due to the low mating efficiency of Artemis expressing pGBKT7 constructs. Although mating efficieny was increased to raise the number of diploids to more sufficient levels, the site-specific recombination of the pGADT7-Rec vector prevented a thorough and complete analysis with this system. Taken into consideration the results of both of these former screenings, it is fitting to propose that the expression of Artemis within the yeast may have imposed toxic side effects or elicited yeast stress response(s). This may be the result of the expression of a protein with endonucleolytic capability within the nucleus. Previous attempts to find Artemis interacting partners using a yeast two hybrid approach have not been reported. The only close encounter was in 2004, in a study conducted by Ishiai *et al*, a Y2H screen was carried out to identify binding partners of SNM1A/PSO2 [an SNM1 homolog which bears homology with the metallo-β-lactamase domain of Artemis (Figure 7C)]. The carboxyl-terminal fragment of chSNM1A (encoding the metallo-β-lactamase domain) was used as bait. In a screen of $8x10^6$ DT40 cDNA library clones, 14 clones were found to specifically interact with chSNM1A, three of which were identical clones encoding hPIAS1 (protein inhibitor of activated STAT,1) protein (85). Although SNM1A and Artemis have unique biological roles and are distinct outside of the β-lactamase domain, both our bait construct and the PSO2 bait employed in this study entailed the same region of homology. Like the results of our Artemis screenings, Ishiai *et al* found a limited number of clones that specifically interacted with SNM1A, and only one clone produing a specific interaction. Most standard Y2H screens produce hundreds of clones that interact with bait of interest, which then lead to identification of several putative interactors.

Y2H screening with PARP-1

PARP-1 was used for Y2H screens not only because it is a protein of great interest in our laboratory but also to serve as a positive control for our systems. Previous studies have reported identification of PARP-1 binding factors using Y2H screens. PARP-1 was shown to interact with the DNA repair factor XRCC1 using a yeast two hybrid assay and co-immunoprecipitation studies. This interaction was demonstrated to occur through the respective BRCT domains of XRCC1 and PARP-1 as well as through the N terminus of PARP-1 (184). The N-terminus of PARP-1 was shown to interact with the DNA repair factor DNA Ligase III via Y2H analysis and confirmed (185). The automodification domain of PARP-1 was also shown to interact with hUbc9 (186) and ribosomal protein S3a (187) by means of Y2H analysis by two independent studies.

Yeast lack all PARP-1 proteins, therefore, it is feasible that the expression of PARP-1 within the pGBKT7 vector in yeast may have been somewhat toxic, or that the protein does not undergo all of the appropriate post-translational modifications within the cell and hence does not fold properly in yeast. It is however, more fitting that the results of our PARP-1 screenings were hindered as a result of technical road blocks. And since PARP-1 served as a positive control for our screens, and a positive hit could not be confirmed, the success of Y2H screening using Artemis as bait is inconclusive.

Classical Y2H system versus Interactive Mating

Both the classical Y2H (Hybrid Hunter™ Two-Hybrid LexA based System) and interactive mating (Matchmaker™ Pre-transformed Yeast Two-Hybrid Gal-4 based system) systems employed in this study offered unique advantages and disadvantages. The biggest disadvantage of the Matchmaker™ pre-transformed system was that high mating efficiencies could not be achieved at sufficient levels to screen the cDNA library until saturation. Also, there is no way to predict that expression of a given bait construct will have a negative impact on the yeast mating efficiency prior to ordering the system. Mating efficiency of at least 2% was required to screen the library until saturation, which was easily achieved using AH109 [pGBKT7-p53] and Y187 [pGADT7-T] control strains. If there was a toxic effect resulting from Artemis or PARP-1 expression in the yeast, it was not exhibited in a way that could have been phenotypically measured, ie. yeast physiology, growth rate, viability, colony size. Mating with both PARP-1 and Artemis bait constructs however, would not yield mating efficiencies higher than 0.03% - 0.06%. Optimizing growth conditions prior to mating resulted in approximately ten fold increase in number of diploids (0.7% and 0.6% for pGBKT7-cPARP-1 and pGBKT7-Nartemis, respectively). A ten fold increase in mating efficiency was still not sufficient to screen the cDNA library until saturation. Double positives (diploids) were significantly easier to achieve with the classical LiAc tranformation system (10^6-10^7 double positives). The main factor that determines the cost, effort and success of a Y2H cDNA screening is the total number of yeast cells that are double positive and expressing the respective fusion proteins.

The cDNA library in the Matchmaker™ system is pre-transformed in the yeast, and is supplied in six 1 ml vials (sufficient for only six library screens). In this manner,

screening is made simpler with less variability than the traditional LiAc transformation techniques; however, only a limited number of screens can be conducted using a single kit. The Hybrid Hunter™ system seems initially more costly than the Matchmaker™ pre-transformed system, as the screening kit and cDNA library must be ordered separately. Subsequent screenings however, can be conducted readily because the cDNA library is transformed in bacteria, and cDNA library amplification can be easily achieved using mega preparation protocols (~10 000 mg of library cDNA each prep using outlined protocol), sufficient for conducting 20 screens. The disadvantage of the Hybrid Hunter™ system is that the pHybLex/Zeo bait vector is engineered with the Zeocin™ resistance gene. Zeocin™ is available exclusively through Invitrogen and is very costly. Selection of yeast transformed with pHybLex/Zeo requires 300 μg/ml of Zeocin™, which costs approximately $250-350 for every library screen (requiring approximately 700-100 large 150 mm plates). There are LexA and GAL4 based systems that employ vector systems that use amino acid prototrophy, where the bait vector can be selected by plating on medium that lacks the appropriate amino acid. This may be beneficial as subsequent screening can be conducted without use of costly Zeocin™.

Future Yeast Two Hybrid Screens

The majority of the road blocks encountered throughout these screenings were technical in nature; therefore it would not be accurate to conclude that future subsequent Y2H screenings would not be successful in identifying biologically relevant interactors. To avoid such technical setbacks, any future screens should be conducted using a classical Y2H system with a vector system that utilizes nutritional based selection (extensive screenings can be conducted with less expenditure). In addition, a normalized cDNA library would be

beneficial (and is now commercially available), which can significantly reduce the number of encountered false positives. Normalized cDNA libraries are "equalized' with respect to the representation of each protein, in contrast to classical cDNA libraries, in which cDNAs from highly abundant mRNAs are overrepresented and cDNAs from less abundant mRNAs are rare, increasing the likelyhood of finding interactions with less abundant mRNAs (188). A number of methods have been developed for normalizing cDNA libraries, some of these methods have also been optimized for enrichment of full-length cDNAs (189). One such method, utilizes new technology called DSN-normalization based on the properties of duplex-specific nuclease (DSN), available through Custom Normalized cDNA Library Construction of Clontech Laboratories. Numerous studies have also developed protocols for enriching open reading frames of a cDNA library and have been shown to increase the number of clones with inserts in the correct reading frame more than fourfold compared to classical cDNA libraries (190, 191, 192, 193).

The binding factors that were identified for Artemis and PARP-1 though initially looked very promising may represent biological false positives. Further elucidation of these identified factors is required. In order to validate a potentially genuine interaction, these cDNAs must be re-cloned back into a prey plasmid (ie. pYESTrp2) in the proper frame and retested for interaction with the bait. Full length cDNA of the putative interactor would not be necessary, as the partial cDNA derived should be sufficient to re-produce the interaction. Once expression of prey in the correct frame is confirmed, the interaction must be confirmed using a combination of other techniques to validate and characterize the interaction. Co-immunoprecipitation are considered to be the gold standard for validating putative protein-protein interations identified in Y2H system, mediated by an antibody against a target antigen that in turn precipitates with the interacting protein.

CONCLUSIONS

In summary, the results presented in this study illustrated the strengths and weaknesses of using Y2H analysis (classical and interaction mating systems) for identification of Artemis and PARP-1 binding factors. Although Y2H analysis is a powerful technique that has led to a significant number of publications over the last decade, it is not suitable for all bait proteins and may not yield reproducible results consistently. Furthermore, many conventional and alternative Y2H systems are readily available promising more efficient results with fewer occurrences of result-polluting false positives. It was found in our study that classical Y2H systems are both more cost efficient and provide less variability than the interaction mating systems. Moreover, vector systems that are based on nutritional selection and that do not employ Zeocin^{TM} resistance for selection are much more cost effective options. The Artemis and PARP-1 binding factors identified in this study show that one can identify two proteins that may share functional and biological significance that "interact" when co-expressed in a cell, while not representing a genuine interaction. Conversely, two completely unrelated and functionally distinct proteins or polypeptides may interact when co-expressed in a cell because the protein interfaces produce an energetically favorable interaction, or the presence of the two proteins may strengthen or stabilize transcriptional machinery leading to reporter gene activation. Conversly, while these false positives may yield strong interactions in a Y2H system, two interactors with strong affinity for one another that interact within the cell may produce a weak interaction in a Y2H system that may not lead to reporter gene activation and not be identified. Therefore, like all methods of protein interaction analysis, a combination of techniques is necessary to verify and characterize a genuine interaction identified via a Y2H system.

APPENDICES

Appendix I: Composition and use of different growth media

Table 11: Various culture media used to grow *E.* coli, L40 and AH109 strains and select for various plasmids.

Medium	Selection	Purpose
LB	All bacteria strains	Lenox L broth for bacteria culture
LB Zeocin	*E. Coli* + pHybLex/Zeo	Low Salt LB supplemented with 25 μg/ml Zeocin for selection of bait plasmid in bacteria
LB Kan	*E. Coli* + pGBKT7	LB medium supplemented with 100 μg/ml Kanamycin for selection of bait plasmid in bacteria
LB Amp	*E. Coli* + pGADT7-Rec or *E. Coli* + pYESTrp2/pJG4-5	LB medium supplemented with 100 μg/ml Ampicillin for selection of prey plasmid in bacteria
YPD or YPAD	All yeast strains	Complex medium for yeast growth (non-selective) L40 and AH109 require Adenine for growth
YPAD Z300	L40 + pHybLex/Zeo	YPAD supplemented with 300 μg/ml Zeocin selects for bait plasmid
YC	All strains	Can be made selective
YC-W	L40 + pYESTrp2/pJG4-5 or AH109 + pGBKT7	Medium does not contain tryptophan, can select for plasmid that contains the Tryptophan auxotrophic marker
YC-L	AH109 + PGADT7-Rec	Media does not contain Leucine, can select for plasmid that contains the Leucine auxotrophic marker
YC-LW	AH109 + pGADT7-Rec + pGBKT7	Media does not contain tryptophan or Leucine. Selects for leucine and Tryptophan auxotrophic markers
YC-WU Z300	L40 + pYEStrp2/pJG4-5 + pHybLex/Zeo	Media contains 300 μg/ml Zeocin, does not contain tryptophan Selects for Zeocin resistance and Tryptophan auxotrophic marker
YC-WHUK Z300	L40 + pYESTrp2/pJG4-5 + pHybLex/Zeo and interaction	Media contains 300 μg/ml Zeocin, does not contain Tryptophan, Histidine, Uracil or Lycine Selects for presence of Bait, prey plasmids and activation of Reporter gene
YC-WHAL	AH109 + pGBKT7 + pGADT7-Rec and interaction	Media does not contain Tryptophan, Histidine, Adenine or Leucine Selects for presence of bait and prey and activation of 2 reporter genes: Histidine and Adenine

Culture medium

LB Broth (Lenox L broth)

1% Peptone
1% NaOH
0.5% Yeast Extract Granulated
+/- 2% Agar

This mixture of peptone, yeast extract, and sodium chloride (EMD laboratories Inc.) were dissolved in water (1 L). The pH of the medium was adjusted to 7 using NaOH pellets. Agar was added to mixture for LB plates. The media was autoclaved at 121°C for 20 min. Kanamaycin, or Ampicillin were added at a final concentration of 100 μg/ml when appropriate once the autoclaved media was cool to approximately 50 °C.

Low salt LB broth (Zeocin resistant strains)

1% Peptone
5% NaOH
5% Yeast Extract Granulated
+/- 2% Agar

The pH of the medium was adjusted to 7.5 using NaOH pellets. The liquid medium was autoclaved at 121°C for 20 min. Zeocin™ at a final concentration of 25 μg/ml were added to medium once autoclaved media was cooled to approximately 50 °C.

YPD (YPAD) medium

Yeast Peptone Dextrose Medium

2% Peptone

1% Yeast Extract Granulated

2% D-Glucose
+/- 0.01% Adenine
+/- 300 μg/ml Zeocin™
+/- 2% g Agar

Adenine was supplemented as it is required for growth of AH109 and L40 strains. The pH of media was adjusted to 7.5 for Zeocin™ selection using NaOH pellets. Glucose was autoclaved separately and added to medium after sterilization. Media was autoclaved at 121°C for 20 min.

YC minimal medium for Yeast

Yeast were grown in supplemented YC minimal medium to maintain selection for plasmids introduced into the strain and for the integrated reporter constructs. Amino acid(s) were omitted from media for selection of plasmid(s) and integration of reporter construct(s) in screen. Generally strains were maintained in mid log phase growth.

0.12% yeast nitrogen base (**without either** amino acids or ammonium sulfate)
0.5% ammonium sulfate
1% succinic acid
2% glucose
0.01% (adenine, arginine, cysteine, leucine, lysine, threonine, tryptophan, uracil)
0.005% (aspartic acid, histidine, isoleucine, methionine, phenylalanine, proline, serine, tyrosine, valine)
2% agar (for plates)

Amino acids were omitted according to appropriate drop out media for specific selection. The pH was adjusted with NaOH pellets. When making plates the agar was added after dissolving all other reagents. Glucose was autoclaved separately and added to media after sterilization. Zeocin™ (300 μg/ml) and 3-AT (5 mM) were added when appropriate after media has cooled to 50°C. Drop out media can be autoclaved for no longer than 20 minutes at 15 psi, 121°C. Zeocin™ plates are stable for month stored at 4 °C in the dark.

Appendix II: Primers and Oligomers used in PCR reactions

Primer	Sequence
Artemis Forward Primer I	5'-AGTTCTTTCGAGGGGCAGATGGCCGAGTATCCAACT-3'
Artemis Reverse Primer II	5'-ATACTCGAGCTAGTCATCTTCCTCCTCTGAGTCTCGGTGAACTGT-3'
Artemis Forward Primer III	5'-GCTAGAACAGTTCACCGAGACTCAGAGGAGGAAGATGACTAT-3'
Artemis Reverse Primer IV	5'-ATACTCGAGTTAGGTATCTAAGAGTGAGCATTTTCTTTTTTT-3'
Artemis DelA Forward primer I	5'-ACCAAATGCTCACTCTTAGATACC-3'
Artemis DelA primer II	5'-ACCTTCCTCGGGTATATGTGTCTCACTGC-3'
Artemis DelB primer III	5'-ACCAAAAGTTTGCTAAATCTTAGC-3'
Artemis DelB reverse primer IV	5'-ACCCAAGGAAGTAATATCCCCACTG-3'
Artemis DelC forward primer V	5'-ACCGACAAAGCTGACTACAGACC-3'
Artemis DelC reverse primer VI	5'-ACCTGCCACTGTGGAGGAAGGGAAG-3'
Artemis DelD forward primer VII	5'-ACCGGGGGATCTCAGTCACCAAAGC-3'
Artemis DelD reverse primer VIII	5'-ACCTTCACAATCTACAAAGTTTGTG-3'
pYESTrp2 Seq forward primer	5'- AGGGCTGGCGGTTGGGGTTATTCGC-3'
pYESTrp2 Seq reverse primer	5'-GAGTCACTTTAAAATTTGTATACAC-3'
pJG4-5 Seq forward primer	5'-GATGCCTCCTACCCTTATGATGTGCC-3'
pJG4-5 Seq reverse primer	5'-GGAGACTTGACCAAACCTCTGGCG-3'
PARP-1 forward primer I	5'-GGAGGATTACATATGGAGCTAAAGAAAGTGTGTTCAACT-3'
PARP-1 reverse primer II	5'-TCCGGTCGACTGTTCTTCATATAATTTCATGAACTG-3'
Artemis forward primer A	5'-GGAGGAGGACATATGAGTTCTTTCGAGGGGCAGATGGCCGA GTATCCAACT-3'
Artemis reverse primer B	5'TCCCGCTGCAGGCTAGTCATCTTCCTCCTCTGAGTCTCGGTGAAACTGT-3'
Artemis forward primer C	5'-GGAGGAGGACATATGGCTAGAACAGTTCACCGAGACTCAGAGGAG GAAGATGA-3'
Artemis reverse primer D	5'-TCCCGCTGCAGGTTAGGTATCTAAGAGTGAGCATTTTCTTTTTTT-3'
pGADT7-Rec Forward	5'- CTT TAA TAC GAC TCA CTA TAG GGC GAG C -3'
pGADT7-Rec Reverse	5'-CAG TAT CTA CGA TTC ATC TGC AGC TCG AG -3'

Appendix III: Small Scale Yeast Transformation Protocol

A small-scale yeast transformation protocol was used for routine transformations. All overnight yeaset cultures (10 ml in volume) were grown in 50 ml Falcon tubes to allow sufficient room for efficient swirling. Several holes were punctured into Falcon tube cap to ensure that yeast grow aerobically. 10 ml of YPAD was inoculated with a fresh large colony of AH109 or L40. Yeast 10 ml culture must be grown in 50 ml falcon tubes to prevent yeast from settling to the bottom and growing anaerobically. The OD_{600} of the overnight culture was determined and diluted to an OD_{600} of 0.3 with YPAD (or YC drop out medium) and grown for an additional 4-5 hours until culture reached an OD600 ~ 0.8. Cells were centrifuged at 1000 g and resuspended in 40 ml of 1X TE (10mM Tris-HCl, 1 mM EDTA, pH 7.5). The cells were centrifuged again at 1000 g and resuspended in 2ml of 1X LiAc/0.5X TE (100 mM LiAc, 5 mM Tris-HCl, 0.5 mM EDTA, pH 7.5) and incubated at room temperature for 10 minutes. For each yeast transformation, 100 µl of this yeast suspension was mixed with 1µg of plasmid DNA and 100 ng denatured sheared salmon sperm DNA. To this, 700 µl of 1X LiAc/40% PEG-3350/1X TE (100 mM Lithium acetate, 40% PEG-3350, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) was added and mixed by inverting and incubated at 30 °C for 30 minutes. An 88 µl aliquot of DMSO was added to the mixture and heat shocked at 42 °C for 7 minutes (if transforming with a Zeocin™ resistant plasmid, 4 ml of YPAD was added to the solution post heat shock and incubated for 1 hour at 30°C). The mixture was centrifuged at 9000 rpm (microcentrifuge) for 1 min and the supernatant was removed. The pellet was resuspended in 200 µl 1X TE. A 100 µl aliquot of 10x dilution of this suspension was plated on appropriate media for selection of transformed plasmid. Plates were incubated at 30 °C for 2-3 days until plump yeast colonies appeared.

Appendix IV: Colony Grid Pattern

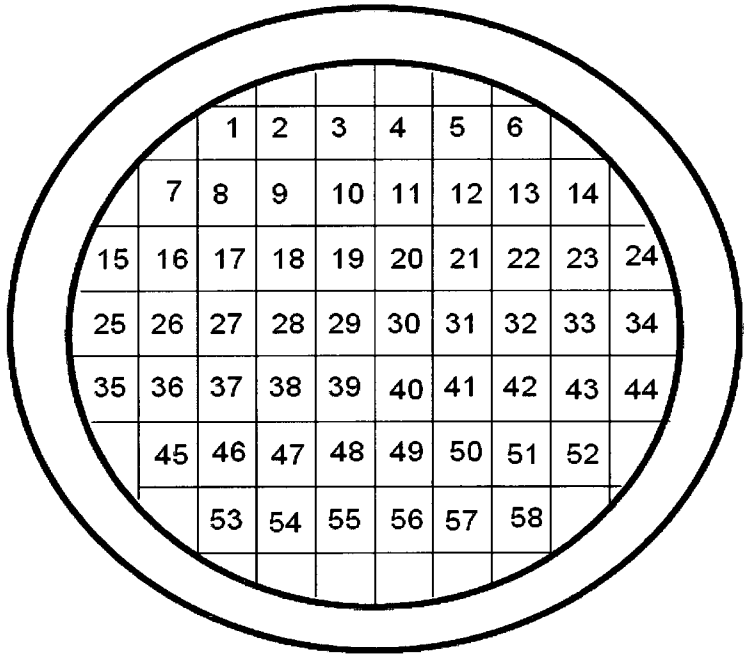
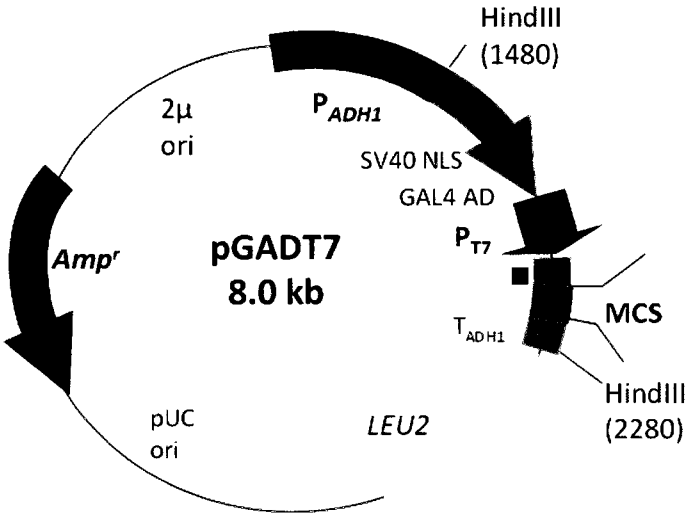


Figure 38: Colony grid pattern used to create master plate of putative interactors. The putative interactors were patched on to master plate and the filter lift assay protocol was employed for detection of β-galactosidase activity.

Appendix V: pGADT7 vector



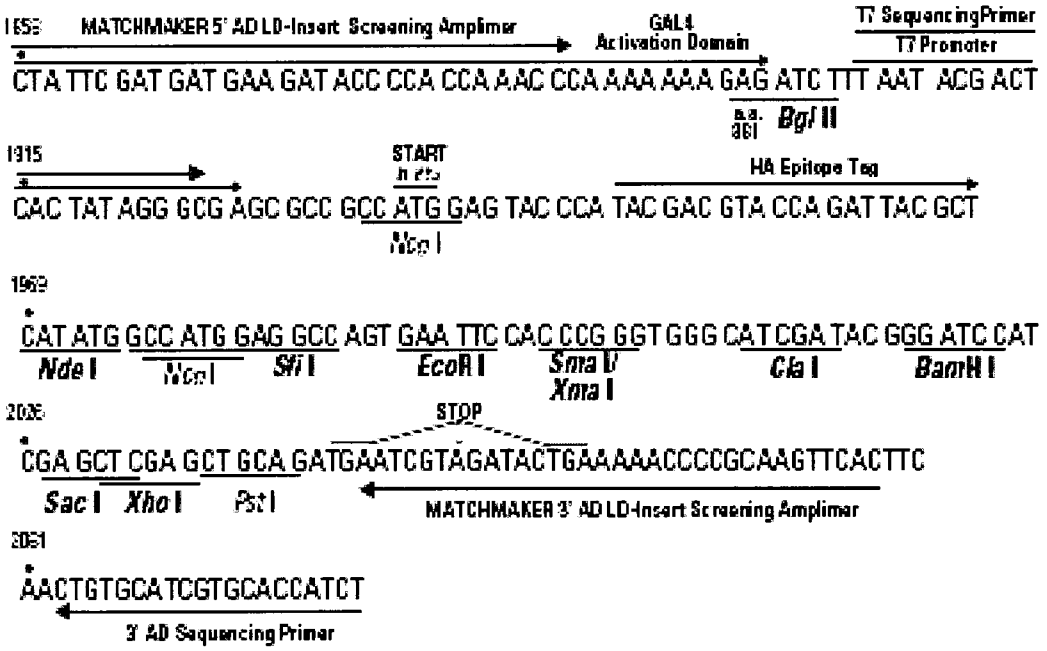


Figure 39: Diagram of the pGADT7 vector. During the yeast two-hybrid screen, the pGADT7 was the positive control plasmid that encoded a fusion of the SV40 large T antigen and the GAL4 AD. The GAL4 AD+ fusions expressed from this vector possess a hemagglutinin (HA) epitope tag (■). The pGADT7-Rec vector contains an antibiotic resistance marker (Ampr) for selection in *E. coli* and a nutritional marker (LEU2) for selection in yeast [Adapted from (134)].

Appendix VI: Amino acid sequences

Artemis amino acid sequence

```
  1 mssfegqmae yptisidrfd renlrarayf lshchkdhmk glraptlkrr lecslkvyly
 61 cspvtkelll tspkyrfwkk riisieietp tqislvdeas gekeeivvtl lpaghcpgsv
121 mflfqgnngt vlytgdfrla qgeaarmell hsggrvkdiq svyldttfcd prfyqipsre
181 eclsgvlelv rswitrspyh vvwlnckaay gyeylftnls eelgvqvhvn kldmfrnmpe
241 ilhhlttdrn tqihacrhpk aeeyfqwskl pcgitsrnri plhiisikps tmwfgersrk
301 tnvivrtges syracfsfhs syseikdfls ylcpvnaypn vipvgttmdk vveilkplcr
361 ssqstepkyk plgklkrart vhrdseeedd ylfddplpip lrhkvpypet fhpevfsmta
421 vsekqpeklr qtpgccraec mqssrftnfv dceesnsese eevgipaslq gdlgsvlhlq
481 kadgdvpqwe vffkrndeit deslenfpss tvaggsqspk lfsdsdgest hissqnssqs
541 thiteqgsqg wdsqsdtvlv ssqernsgdi tsldkadyrp tikenipasl meqnvicpkd
601 tysdlksrdk dvtivpstge pttlssethi peeksllnls tnadsqsssd fevpstpeae
661 lpkrehlqyl yeklatgesi avkkrkcsll dt
```

PARP-1 amino acid sequence

```
  1 mdielpylae yartgratck gckstiskdt lriavmvqsa fhdakvpnwf hktcffknqr
 61 pssvgdiqni gnlrfadqke ltdlveniqe visaqlgkkr skafnlalkd fgieyakssr
121 stcrgceqki nkdlvrlrkt vydtevgmky ggqplwhhle cfaqlrselg wfasgedmpg
181 fqsladddqa kvknaippik seelpdtkra kmelsdtnee gekkqrlkdq ndayfrfrdd
241 iknkmkkkdi dillkfnnqq pvtgdteklf dqtadlltfg aiescsecns cqfivnksgy
301 icngnhsewt kcnkllkept rsacivpkel kalynflntv keipstrifn nfppnkstfs
361 rslltknknn dvlvrptipr ispplynlkf siiglknqhk elrkrienlg gkfevkisen
421 tiaiistele iqkkstrmkf aeelgihivp iefldfvead tegaikyins tcicswgtdp
481 ksripkettk slnsnsiytk smpvsrtfkv kdglavdpds gldiahvyv dsnnkysvvl
541 gltdiqrnkn syykvqllka dkkekywifr swgrigtnig nskleefdts esakrnfkei
601 yadktgneye qrdnfvkrtg rmypieiqyd ddqklvkhes hfftskleis vqnliklifd
661 idsmnktlme fhidmdkmpl gklsahqiqs ayrvvkeiyn vlecgsntak lidatnrfyt
721 liphnfgvql ptliethqqi edlrqmldsl aeievaysii ksedvsdacn pldnhyaqik
781 tqlvaldkns eefsilsqyv knthasthks ydlkivdvfk vsrqgearrf kpfkklhnrk
841 llwhgsrltn fvgilshglr iappeapptg ymfgkgiyfa dmvsksanyc ctsqqnstgl
901 mllsevalgd mmectsakyi nklsnnkhsc fgrgrtmpdp tksyirsdgv eipygetitd
961 ehlkssllyn eyivydvaqv niqylfrmef kysy
```

ATRX amino acid sequence

```
  1 mcillhdlhe sklntlvqkl hdflahssee seetsspprl amnqntdkis gsgsnsdmme
 61 nskeegtsss ekskssgssr skrkpsivtk yvesddekpl ddetvnedas nensenditm
121 qslpkgtviv qpepvlnedk ddfkgpefrs rskmktenlk krgedglhgi vsctacgqqv
181 nhfqkdsiyr hpslqvlick ncfkyymsdd isrdsdgmde qcrwcaeggn liccdfchna
241 fckkcilrnl grkelstimd ennqwycyic hpeplldlvt acnsvfenle qllqqnkkki
301 kvdseksnkv yehtsrfspk ktssncngee kklddscsgs vtysysaliv pkemikkakk
361 liettanmns syvkflkqat dnseissatk lrqlkafksv ladikkahla leedlnsefr
421 amdavnkekn tkehkvidak fetkarkgek pcalekkdis kseaklsrkq vdsehmhqnv
481 pteeqrtnks tggehkksdr keepqyepan tsedldmdiv svpssvpedi fenletamev
541 qssvdhqgdg ssgteqeves ssvklnissk dnrggikskt takvtkelyv kltpvslsns
601 pikgadcqev pqdkdgyksc glnpklekcg lgqensdneh lvenevslll eesdlrrspr
661 vkttplrrpt etnpvtsnsd eecnetvkek qklsvpvrkk dkrnssdsai dnpkpnklpk
721 skqsetvdqn sdsdemlail kevsrmshss ssdtdineih tnhktlydlk tqagkddkgk
781 rkrksstsgs dfdtkkgksa kssiiskkkr qtqsessnyd selekeiksm skigaarttk
841 kripntkdfd ssedekhskk gmdnqghknl ktsqegssdd aerkqeretf ssaegtvdkd
901 ttimelrdrl pkkqqasast dgvdklsgke qsftslevrk vaetkekskh lktktckkvq
```

```
 961 dglsdiaekf lkkdqsdets eddkkqskkg teekkkpsdf kkkvikmeqq yesssdgtek
1021 lpereeichf pkgikqikng ttdgekkskk irdktskkkd elsdyaekst gkgdscdsse
1081 dkksknqayg rekkrckllg kssrkrqdcs ssdtekysmk edgcnssdkr lkrielrerr
1141 nlsskrntke iqsgssssda eessednkkk kqrtsskkka vivkekkrns lrtstkrkqa
1201 ditsssssdi edddqnsige gssdeqkikp vtenlvlssh tgfcqssgde alsksvpvtv
1261 ddddddndpe nriakkmlle eikanlssde dgssddepee gkkrtgkqne enpgdeeakn
1321 qvnsesdsds eeskkpryrh rllrhkltvs dgesgeekkt kpkehkevkg rnrrkvssed
1381 sedsdfqesg vseevsesed eqrprtrsak kaeleenqrs ykqkkkrrri kvqedsssen
1441 ksnseeeeee keeeeeeeee eeeeeedend dskspgkgrk kirkilkddk lrtetqnalk
1501 eeeerrkria erereereklr evieiedasp tkcpittklv ldedeetkep lvqvhrnmvi
1561 klkphqvdgv qfmwdccces vkktkkspgs gcilahcmgl gktlqvvsfl htvllcdkld
1621 fstalvvcpl ntalnwmnef ekwqeglkdd eklevselat vkrpqersym lqrwqedggv
1681 miigyemyrn laqgrnvksr klkeifnkal vdpgpdfvvc deghilknea savskamnsi
1741 rsrrriiltg tplqnnliey hcmvnfiken llgsikefrn rfinpiqngq cadstmvdvr
1801 vmkkrahily emlagcvqrk dytaltkflp pkheyvlavr mtsiqcklyq yyldhltgvg
1861 nnseggrgka gaklfqdfqm lsriwthpwc lqldyisken kgyfdedsmd efiasdsdet
1921 smslssddyt kkkkkgkkgk kdssssgsgs dndvevikvw nsrsrgggeg nvdetgnnps
1981 vslkleeska tsssnpsspa pdwykdfvtd adaevlehsg kmvllfeilr maeeigdkvl
2041 vfsqslisld liedflelas rektedkdkp liykgegkwl rnidyyrldg sttaqsrkkw
2101 aeefndetnv rgrlfiistk agslginlva anrviifdas wnpsydiqsi frvyrfgqtk
2161 pvyvyrflaq gtmedkiydr qvtkqslsfr vvdqqqverh ftmneltely tfepdllddp
2221 nsekkkkrdt pmlpkdtila ellqihkehi vgyhehdsll dhkeeeelte eerkaawaey
2281 eaekkgltmr fniptgtnlp pvsfnsqtpy ipfnlgalsa msnqqledli nqgrekvvea
2341 tnsvtavriq plediisavw kenmnlseaq vqalalsrqa sqeldvkrre aiyndvltkq
2401 qmliscvqri lmnrrlqqqy nqqqqqqmty qqatlghlmm pkppnlimnp snyqqidmrg
2461 myqpvaggmq ppplqrappp mrsknpgpsq gksm
```

Caspase-7 amino acid sequence

```
   1 maddqgciee qgvedsaned svdakpdrss fvpslfskkk knvtmrsikt trdrvptyqy
  61 nmnfeklgkc iiinnknfdk vtgmgvrngt dkdaealfkc frslgfdviv yndcsvtqag
 121 vqrrdlgrlq pppprfkqys clsllsswdy rhapprlanf cifrregfrh vgqaglklla
 181 ssdpptsgsq sagitgvslh tqpqimkvsl rhcntficfa
```

"Unknown protein" amino acid sequence

```
   1 fkptkikrdk eghyimvkgs iqqeeltiln iyapntgapr fikqvlsdlq rdldshtlim
  61 gdfntplstl drstrqkvnk dtqelnsalh qadlidiyrt lhpksteytf fsaphhtysk
 121 idhivgskal lskckrteii tnylsdhsai klelriknlt qsrsttwkln nlllndywvh
 181 nemkaeikmf fetnenkdtt yqnlwdafka vcrgkfialn aykrkqersk idtltsqlke
 241 lekqeqthsk asrrqeitki raelkeietq ktlqkinesg swfferinki drplarlikk
 301 kreknqidti kndkgdittd pteiqttire yykhlyankl enleemdtfl dtytlprlnq
 361 eeveslnrpi tgseivaiin slptkkspgp dgftaefyqr ymeelvpfll klfqsiekeg
 421 ilpnsfyeas iilipkpgrd ttkkenfrpi slmnidakil nkilanriqq hikklihhdq
 481 vgfipgmqgw fnirksinvi qhinrakdkn hmiisidaek afdkiqqpfm lktlnklgid
 541 gtyfkiirai ydkptaniil ngqkleafpl ktgtrqgcpl spllfnivle vlarairqek
 601 eikgirlgke evklslfadd mivylenpiv saqnllklis nfskvsgyki nvqksqafly
 661 tnnrqtesqi mgelpfvias krikylgiql trdvkdlfke nykpllkeik edtnkwknip
 721 cswvgriniv kmailpkviy rfnaipiklp mtfftelekt tlkfiwnqkr ariaksilsq
 781 knkaggitlp dfklyykatv tktawywyqn rdidqwnrte pseimphiyn ylifdkpekn
 841 kqwgkdslfn kwcwenwlai crklkldpfl tpytkinsrw ikdlnvkpkt iktleenlgi
 901 tiqdigvgkd fmsktpkama tkdkidkwdl iklksfctak ettirvnrqp ttwekifaty
 961 ssdkglisri ynelkqiykk ktnnpikkwa kdmnrhfske diyaakkhmk kcssslaire
1021 mqikttmryh ltpvrmaiik ksgnnrcwrg cgeigtllhc wwdcklvqpl wksvwrflrd
1081 leleipfdpa ipllgiypne yksccykdtc trmfiaalft iaktwnqpkc ptmidwikkm
1141 whiytmeyya aikndefisf vgtwmkleti ilsklsqeqk tkhrifslig gn
```

REFERENCES:

1. Hanawalt, P.C., P.K. Cooper, A. K. Ganesan, R. S. Lloyd, C. A. Smith, and M. E. Zolan. 1982. Repair Responses to DNA Damage: Enzymatic Pathways in *E. coli* and Human cells. *J. Cell. Biochem.* 18:271-283.

2. J.W., and S. J. Elledge. 2007. The DNA Damage Response: Ten Years After. *Mol. Cell.* 28: 739-745.

3. Vispe, S., and M.S. Satoh. 2000. DNA repair patch-mediated double strand DNA break formation in human cells. *J. Biol. Chem.* 275: 27386-92.

4. Shivastavi, M., L.P. De Haro, and J.A. Nickoloff. 2008. Regulation of DNA double-strand break repair pathway choice. *Cell. Res.* 18: 134-147.

5. Weterings, E., and D.J. Chen. 2008. The Endless tale of non-homologous end joining. *Cell Res.* 18: 114-124.

6. Schulte-Uentrop, L., R.A. El-Awady, L. Schliecker, H. Willers, and J. Dahm-Daphi. 2008. Distinct roles of XRCC4 and Ku80 in non-homologous end-joining of endonuclease-and ionizing radiation-induced DNA double-strand breaks. *Nuc. Acid. Res.* 36(8): 2561-2569.

7. Haince, J.F., S. Kozlov, V.L. Dawson, T.M. Dawson, M.J. Hendzel, M.F. Lavin, and G.G. Poirier. 2007. Ataxia Telangiectasia Mutated (ATM) Signalling Network Is Modulated by a Novel Poly(ADP-ribose)-dependent Pathway in the Early Response to DNA- damaging Agents. *J. Biol. Chem.* 282(22): 16441-16453.

8. Morio, T., and H. Kim. 2008. Ku, Artemis, and ataxia-telangiectasia-mutated; Signalling networks in DNA damage. *Int. J. Biochem. Cell. Biol.* 40: 598-603.

9. Lavin, M.F. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. 2007. *Oncogene.* 26: 7749–7758.

10. Williams, R.S., J. S. Williams, and J.A. Trainer. 2007. Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signalling, and the chromatin template. *Biochem. Cell. Biol.* 85: 509-520.

11. Huen, M.S.Y. and J. Chen. 2008. The DNA damage response pathways: At the crossroad of protein modification. *Cell. Res.* 18:8-16.

12. Durocher, D. and S.P. Jackson. 2001. DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr. Opin. Cell. Biol.* 13:225-231.

13. Lavin, M.F., G. Birrell, P. Chen., S. Kozlov, S. Scott, and N. Gueven. 2005. ATM signalling and genomic stability in response to DNA damage. *Mutat. Res.* 569(1-2):123-32.

14. Bakkenist, C.J., and M. B. Kastan. 2003. DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. *Nature*. 421: 499-506.

15. Berkovich, E., RJ. Jr. Monnat, and M.B. Kastan. 2007. Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair. *Nat. Cell Biol.* 9: 683–690.

16. Eggo, P.A. and M. Löbrich. 2006. Contribution of DNA repair and cell cycle checkpoint arrest to the maintenance of genomic stability. *DNA. Repair.* 5(9-10):1192-8.

17. Kuo, L.J. and L.X. Yang. 2008. Gamma-H2AX - a novel biomarker for DNA double-strand breaks. *In. Vivo*. 22(3):305-9.

18. Iijima, K., M. Ohara, R. Seki, H. Tauchi. 2008. Dancing on Damaged Chromatin: Functions of ATM and RAD50/MRE11/NBS1 Complex in Cellular Responses to DNA Damage. *J. Radiat. Res.* 49: 451-464.

19. Kolas, N.K., J.R. Chapman, S. Nakada, J. Ylanko., R. Chahwan, R., F.D. Sweeney, S. Panier, M. Mendez, J. Wildenhain, T.M. Thomson, L. Pelletier, S.P. Jackson, and D. Durocher. 2007. Orchestration of the DNA damage response by the RNF8 ubiquitin ligase. *Science*. 318: 1637-1640.

20. Berkovich, E., R.J. Jr. Monnat, and M.B. Kastan. 2007. Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair. *Nat. Cell. Biol.* 9: 683-690.

21. Goodarzi, A.A., A.T. Noon, D. Deckbar, Y. Ziv, Y. Shiloh, and M. Lobrich. 2008. ATM Signalling Facilitates Repair of DNA Double-Strand Breaks Associated with Heterochromatin. *Mol. Cell.* 31: 167-177.

22 Ziv, Y., D. Bielopolski, Y. Galanty, C. Lukas, Y. Taya, D.C. Schultz, J. Lukas, S. Bikker-Jensen, J. Bartek, Y. Shiloh. 2006. Chromatin Relaxation in response to DNA double strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell. Biol.* 8: 870-876.

23. Riches, L.C., M. L. Anthony, and N.J. Gooderham. 2008. Early events in the Mammalian Response to DNA double-strand breaks. *Mutagenesis*. 23(5):331-339.

24. Paull, T.T., E.P. Rogakou, V. Yamaziki, C.U. Kirchgessner, M. Gellert, and W.M, Bonner. 2000. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* 10(15):886-95.

25. Furuta, T. H. Takemura, Z.Y. Liao, G.J. Aune, C. Redon, O.A Sedelnkova, D.R. Pilch, E.P. Rogakou, A. Celeste, H.T. Chen, A. Nussenzweig, M.I. Aladjem, W.M.

Bonner, and Y. Pommier. 2003. Phosphorylation of histone H2AX and activation of Mre11, Rad50, and Nbs1 in response to replication-dependent DNA double-strand breaks induced by mammalian DNA topoisomerase I cleavage complexes. *J. Biol. Chem.* 278(22):20303-12.

26. Abraham, R.T. and R. S. Tibbetts. 2005. Guiding ATM to Broken DNA. *Cell Biol.* 308: 510-511.

27. Zhou, C., Z. Li, H. Diao, Y. Yu, W. Zhu, Y. Dai, F.F. Chen, and J. Yang. 2006. DNA damage evaluated by gammaH2AX foci formation by a selective group of chemical/physical stressors. *Mutat. Res.* 604: 8–18.

28. MacPhail, S. H., J.P.Banath, Y. Yu, E. Chu, and P.L. Olive. 2003. Cell cycle-dependent expression of phosphorylated histone H2AX: reduced expression in unirradiated but not X-irradiated G1-phase cells. *Radiat. Res.* 159: 759–767.

29. Tanaka, T., A. Kurose, X. Huang, W. Dai, and Z. Darzynkiewicz. 2006. ATM activation and histone H2AX phosphorylation as indicators of DNA damage by DNA topoisomerase I inhibitor topotecan and during apoptosis. *Cell. Prolif.* 39:49–60.

30. Rogakou, E. P., W. Nieves-Neira, C. Boon, Y. Pommier, and W.M. Bonner. 2000 Initiation of DNA fragmentation during apoptosis induces phosphorylation of H2AX histone at serine 139. *J. Biol. Chem.* 275:9390–9395.

31. McGowan, C.H. 2003. Regulation of the eukaryotic cell cycle. *Prog. Cell. Cycle. Res.* 5:1-4.

32. Woo, R.A., and R.Y. Poon. 2003. Cyclin-dependent kinases and S phase control in mammalian cells. *Cell. Cycle.* 2(4):316-24.

33. Nillson, I., and I. Hoffmann. 2000. Cell cycle regulation by the Cdc25 phosphatase family. *Prog. Cell. Cycle. Res.* 4:107-14.

34. Sancar, A., L.A. Lindsey-Boltz, K. Unsal-Kacmaz K, and S. Linn. 2004. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* 73:39-85.

35. Grallert, B., and E. Boye. 2008. The multiple facets of the intra-S checkpoint. *Cell. Cycle.* 7(15):2315-20

36. Lindahl, T., M.S. Satoh, G.G. Poirier, and A. Klungland. 1995. Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks. *Trends Biochem. Sci.* 20:405–411.

37. D'Amours, D., S. Desnovers, I. D'Silva, and G.G. Poirier. 1999. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem. J.* 342 (Pt 2): 249-68.

38. Woodhouse, B.C., and G.L. Dianov. 2008. Poly ADP-ribose polymerase-1: An international molecule of mystery. *DNA. Repair.* 7: 1077-1086.

39. Herceg, Z., and Z. Q. Wang. 2001. Functions of poly(ADP-ribose) polymerase (PARP) in DNA repair, genomic integrity and cell death. *Mutat. Res.* 477(1-2):97-110.

40. Kim. M.Y., T. Zhang., and W.L. Kraus. 2005. Poly(ADP-ribosyl)ation by PARP-1: 'PAR-laying' NAD+ into a nuclear signal. *Genes. Dev.* 19(17):1951-67.

41. Dantzer, F., J.C. Amé, V. Screiber, J. Nakamura, J.M. Murcia, and G. De Murcia. 2006. Poly(ADP-ribose) Polymerase-1 Activation During DNA Damage and Repair. *Methods. Enzymol.* 409:493-510.

42. Ahel, I., D. Ahel, T. Matsusaka, A.J. Clark, J. Pines, S.J. Boulton, S.C. West. 2008. Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature.* 451: 81–85.

43. Pleschke, J.M., H.E. Kleczkowska, M. Strohm, and F.R. Althaus. 2000. Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.* 275(52):40974-80.

44. Berger, N.A. 1985. Poly(ADP-ribose) in the cellular response to DNA damage. *Radiat. Res.* 101: 4-15.

45. Carson. D.A., S. Seto, B. Wasson, and C.J. Carrera. 1986. DNA strand breaks, NAD metabolism, and programmed cell death. *Exp. Cell. Res.* 164: 273-281.

46. Oliver, F.J., J. Menissier-de Murcia, and G. de Murcia. 1999. Poly(ADP-Ribose) Polymerase in the Cellular Response to DNA Damage, Apoptosis, and Disease. *Am. J. Hum. Genet.* 64: 1282-1288.

47. Nosseri, C., S. Coppola, and L. Ghibelli. 1994. Possible involvement of poly(ADP-ribosyl) polymerase in triggering stress-induced apoptosis. *Exp. Cell. Res.* 212: 367–373.

48. Scovassi, A. I. and G. G. Poirier. 1999. Poly(ADP-ribosylation) and apoptosis. *Mol. Cell. Biol.* 199: 125-137.

49. Ying, W., C.C. Alano, P. Garnier, and R.A. Swanson. 2005. NAD+ as a metabolic link between DNA damage and cell death. *J. Neurosci. Res.* 79: 216–223.

50. Yang, J., ZP. Xu, Y. Huang, H.E. Hamrick, P. J. Duerksen-Hughes, and YN. Yu. 2004. ATM and ATR: Sensing DNA damage. *World. J. Gastroenterol.* 10(2): 155-160.

51. Lees-Miller, S.P and K. Meek. 2003. Repair of DNA double strand breaks by non-homologous end joining. *Biochimie.* 85(11): 1161-73.

52. Burma, S., B.P.C. Chen, and D.J. Chen. 2006 Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA. Repair.* 5: 1042-1048.

53. Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., and Takeda, S. 2006. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA. Repair.* 5: 1021–1029.

54. Griffin, C.S., and J. Thacker. 2004. The role of homologous recombination repair in the formation of chromosome aberrations. *Cytogenet. Genome. Res.* 104(1-4):21-7.

55. Kurimasa, A., S. Kumano, and N.V. Boubnov. 1999. Requirement for the kinase activity of human DNA-dependent protein kinase catalytic subunit in DNA strand break rejoining. *Mol. Cell. Biol.* 19:3877-3884.

56. Kienker, L.J, E.K Shin, and K. Meek. 2000. Both V(D)J recombination and radioresistance require DNA-PK kinase activity, though minimal levels suffice for V(D)J recombination. *Nucleic. Acids. Res.* 28: 2752–2761.

57. Dahm, K. 2007. Functions and regulation of human Artemis in double strand break repair. *J. Cell. Biochem.* 100:1346-1351.

58. Dianov, G.L., P. O'Neil, and D.T. Goodhead. 2001. Securing genome stability by orchestrating DNA repair: removal of radiation-induced clustered lesions in DNA. *Bioessays.* 23: 745-749.

59. Mahaney, B.L., K. Meek, and S.P. Lees-Miller. 2009. Repair of ionizing radiation-induced DNA double-strand breaks by non-homologous end-joining. *Biochem.* 417: 639-650.

60. Povirk, L.F., T. Zhou, R. Zhou, M.J. Cowan, and S.M. Yannone. 2007. Processing of 3'- phosphoglycolate-terminated DNA double strand breaks by Artemis nuclease. *J. Biol. Chem.* 282: 3547-3558.

61. Ma, Y., U. Pannicke, K. Schwarz , and M.R. Lieber. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell.* 108:781-794.

62. Poinsignon, C., R. de Chasseval, S. Soubeyrand, D. Moshous, A. Fischer, R.J. G. Hache´ and J.P de Villartay. 2004. Phosphorylation of Artemis following irradiation induced DNA damage. *Eur. J. Immunol.* 34: 3146–3155.

63. Soubeyrand, S., L. Pope, R. de Chasseval, D. Gosselin, F. Dong, J.P. de Villartay, and R.J.G. Haché. 2006. Artemis Phosphorylated by DNA-dependent Protein Kinase Associates Preferentially with Discrete Regions of Chromatin. *J. Mol. Biol.* 358(5):1200- 11.

64. Goodarzi, A.A., Y. Yu, E. Riballo, P. Douglas, S.A. Walker, R. Ye, C. Harer, C. Marchetti, N. Morrice, P. Jeggo, and S. P. Lees-Miller. 2006. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. *EMBO. J.* 25(16):3880-9.

65. Ahnesorg, P., P. Smith, and S.P. Jackson. 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end joining. *Cell.* 124:301-313.

66. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature.* 302: 575–581.

67. Soulas-Sprauel, P., P. Rivera-Munoz, L. Malivert, G. Le Guader, V. Abramowski, and J.P. de Villartay. 2007. V(D)J and immunoglobulin class switch recombination: a paradigm to study the regulation of DNA end-joining. *Oncogene.* 26: 7780–7791.

68. Shatz, D.G. 2004. V(D)J recombination. *Immun. Rev.* 200: 5-11.

69. Moshous, D., I. Callebaut, R. De Casseval, B. Corneo, M. Cavazzana-Calvo, F. M. Cavazzana-Calvo, F. Le Deist, I. Tezcan, O. Sanal, Y. Bertrand, N. Philippe, A. Fischer, and J.P. de Villartay. 2001. Artemis, a Novel DNA Double-Strand Break Repair/V(D)J Recombination Protein, Is Mutated in Human Severe Combined Immune Deficiency. *Cell.* 105: 177-186.

70. Janeway, C.A., P. Travers, M. Walport, and M.J., Shlomchik. 2005. Immunobiology: The immune system in health and disease. Garland Science: New York and London.

71. Nicolas, N., D. Moshous, M. Cavazzana-Calvo, D. Papadopoulo, R. De Casseval, F. Le Deist, A. Fischer, J.P. de Villartay. 1998. A Human Severe Combined Immunodeficiency (SCID) Condition with Increased Sensitivity to Ionizing Radiations and Impaired V(D)J Rearrangements Defines a New DNA Recombination/Repair Deficiency. *J. Exp. Med.* 188: 627-634.

72. Gu, Y., S. Jin, Y. Gao, D.T. Weaver, and F.W. Alt. 1997. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding

activity, and inability to support V(D)J recombination. *Proc. Natl. Acad. Sci.* 94: 8076–8081.

73. Blunt, T., D. Gell, M. Fox, G.E. Taccioli, A.R. Lehmann, S.P. Jackson, and P.A. Jeggo. 1996. Identification of a nonsense mutation in the carboxyl-terminal region of DNA dependent protein kinase catalytic subunit in the SCID mouse. *Proc. Natl. Acad. Sci.* 93: 10285–10290.

74. Shin, E.K., L.E. Perryman, and K. Meek. 1997. A kinase634 V(D)J Recombination and DNA Repair negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. *J. Immunol.* 158:3565–3569.

75. Frank, K.M., J. M. Sekiguchi, K.J. Seidl, W. Swat, G. A. Rathbun, H. L. Cheng, L. Davidson, L. Kangaloo, and F.W. Alt. 1998. Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature.* 396: 173–177.

76. Fischer, A., M. Cavazzana-Calvo, G. de Saint Basile, J.P. Di Santo, C. Hivroz, F. Rieux- Laucat, F. Le Diest. 1997. Naturally occurring primary deficiencies of the immune system. *Annu. Rev. Immunol.* 15: 93-124.

77. Taccioli, G.E., G. Rathbun, E. Oltz, T. Stamato, P. A. Jeggo, and F.W. Alt. 1993. Impairment of V(D)J recombination in double-strand break repair mutants. *Science.* 260(5015): 207-211.

78. Moshous, D., L. Li, R. de Chasseval, N. Philippe, N. Jabado, M.J. Cowan, A. Fischer, and J.P. de Villartay. 2000. A new gene involved in DNA double-strand break repair and V(D)J recombination is located on human chromosome 10p. *Hum. Mol. Genet.* 9: 583-588.

79. Callebaut, I., D. Moshous, J.P. Mornon, and J.P. de Villartay. 2002. Metallo-β-lactamase fold within nucleic acids processing enzymes: the β-CASP family. *Nucleic. Acids. Res.* 30(16): 3592-3601.

80. Callebaut, I., D. Moshous, J.P. Mornon, and J.P. de Villartay. 2002. Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. *Nucleic. Acids. Res.* 30(16):3592-601.

81. Moshous, D., I. Callebaut, R. de Chasseval, C. Poinsignon, I. Villey, A. Fischer, and J.P. de Villartay. 2003. The V(D)J recombination/DNA repair factor Artemis belongs to the metallo-beta-lactamase family and constitutes a critical developmental checkpoint of the lymphoid system. *Ann. N. Y. Acad. Sci.* 987:150-7.

82. Poinsignon, C., D. Moshous, I. Callebaut, R. de Chasseval, I. Villey, and J.P. de Villartay. 2004. The metallo-beta-lactamase/beta-CASP domain of Artemis constitutes the catalytic core for V(D)J recombination. *J. Exp. Med.* 199(3):315-21.

83. Ma, Y., U. Pannicke, H. Lu, D. Niewoli, K. Schwarz, and M.R. Lieber. 2005. The DNA-dependent protein kinase catalytic subunit phosphorylation sites in human Artemis. *J. Biol. Chem.* 280(40):33839-46.

84. Pannicke, U., Y. Ma, K.P. Hopfner, D. Niewolik, M.R. Lieber, and K. Schwarz. 2004. Functional and biochemical dissection of the structure-specific nuclease Artemis. *EMBO. J.* 23(9):1987-97

85 Ishiai, M., M. Kimura, K. Namikoshi, M. Yamazoe, K. Yamamoto, H. Arakawa, K. Agematsu, N. Matsushita, S. Takeda, J.M. Buerstedde, and M.Takata. 2004. DNA Cross-Link Repair Protein SNM1A Interacts with PIAS1 in Nuclear Focus Formation. *Mol. Cell. Biol.* 24 (24): 10733–10741.

86. Dronkert, M.L., J. de Wit, M. Boeve, M.L. Vasconcelos, H. van Steeg, T.L. Tan, J.H. Hoeijmakers, and R. Kanaar. 2000. Disruption of mouse SNM1 causes increased sensitivity to the DNA interstrand cross-linking agent mitomycin C. *Mol. Cell. Biol.* 20: 4553–4561.

87. Jeggo, P., and P. O'Neill. 2002. The Greek Goddess, Artemis, reveals the secrets of her cleavage. *DNA. Repair.* 1: 771-777.

88. Geng, L., X. Zhang, S. Zheng, and R.J. Legerski. 2007. Artemis Links ATM to G2/M Checkpoint Recovery via Regulation of Cdk1-Cyclin B. *Mol. Cell. Biol.* 27(7): 2625-2635.

89. Dahm, K. 2007. Functions and Regulation of Human Artemis in Double Strand Break Repair. *J. Cell. Biochem.* 100:1346–1351.

90. Drouet, J., P. Frit, C. Delteil, J.P. de Villartay, B. Salles, and P. Calsou. 2006. Interplay between Ku, Artemis, and the DNA-dependent Protein Kinase Catalytic Subunit at DNA Ends. *J. Biol. Chem.* 281(38): 27784–27793.

91. Moshous, D., C. Pannetier., R. Rd. Chasseval., F. Fl. Deist, M. Cavazzana-Calvo, S. Romana., E. Macintyre., D. Canioni, N. Brousse, A. Fischer, J.L. Casanova, J.P. Villartay. 2003. Partial T and B lymphocyte immunodeficiency and predisposition to lymphoma in patients with hypomorphic mutations in Artemis. *J. Clin. Invest.* 111:381-387.

92. Rooney, S., J. Sekiquichi, C. Zhu., H.L. Cheng, J. Manis, S. Whitlow, J. DeVido., J. Chaudhuri, D. Lombard., and F.W. Alt. 2002. Leaky scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. *Mol. Cell.* 10:1379-1390.

93. Riballo, E., M. Kuhne, N. Rief, A. Doherty, G.C. Smith, M.J. Recio, C. Reis, K. Dahm, A. Fricke, A. Krempler, A.R. Parker, S.P. Jackson, A. Gennery, P.A. Jeggo,

and M. Lobrich. 2004. A Pathway of Double-Strand Break Rejoining Dependent upon ATM, Artemis, and Proteins Locating to -H2AX Foci. *Mol Cell.* 16(5):715-24.

94. Chen, B.P. C., N. Uematsu, J. Kobayashi, Y. Lerenthal, A. Krempler, H. Yajima, M. Löbrich, Y. Shiloh, and D. J. Chen. 2006. Ataxia Telangiectasia Mutated (ATM) Is Essential for DNA-PKcs Phosphorylations at the Thr-2609 Cluster upon DNA Double Strand Break. *J. Biol. Chem.* 282(9): 6582-6587.

95. Zhang, X., J. Succi, Z. Feng, S. Prithivirajsingh, M.D. Story, and R.J. Legerski. 2004. Artemis Is a Phosphorylation Target of ATM and ATR and Is Involved in the G2/M DNA Damage Checkpoint Response. *Mol. Cell. Biol.* 24(20): 9207–9220.

96. Deckbar, D., J. Birraux, A. Krempler, L. Tchouandong, A. Beucher, S. Walker, T. Stiff, P. Jeggo, and M. Löbrich. 2007. Chromosome breakage after G2 checkpoint release. *J. Cell. Biol.* 176(6): 749-755.

97. Krempler, A., D. Deckbar, P.A Jeggo, and M. Lobrich. 2007. An imperfect G2M checkpoint contributes to chromosome instability following irradiation of S and G2 phase cells. *Cell. Cycle.* 6(14):1682-6.

98. Audebert, M., B. Salles, and P. Calsou. 2004. Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.* 279: 55117–55126.

99. Wang, H., B. Rosidi, R. Perrault, M. Wang, L. Zhang, F. Windhofer, and G. Iliakis. 2005. DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer. Res.* 65: 4020–4030.

100. Wang, M., W. Wu, B. Rosidi, L. Zhang, H. Wang, and G. Iliakis. 2006. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic. Acids. Res.* 34: 6170–6182.

101. Windhofer, F., W. Wu., M. Wang., S.K. Singh., J. Saha., B. Rosidi., and G. Iliakis. 2007. Marked dependence on growth state of backup pathways of NHEJ. *Int. J. Radiat. Oncol. Biol. Phys.* 68: 1462-1470.

102. Yan, C. T., Boboila, C., Souza, E. K., Franco, S., Hickernell, T. R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A. A., Manis, J. P. *et al.* 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature.* 449:478–482.

103. Kaufmann, S.H., S. Desnoyers, Y. Ottaviano, N.E. Davidson, G.G. Poirier. 1993. Specific proteolytic cleavage of poly(ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. *Cancer. Res.* 53:3976–3985.

104. Morrison, C., G.C. Smith, L. Stingl, S.P. Jackson, E.F. Wagner, and Z.Q. Wang. 1997. Genetic interaction between PARP and DNA-PK in V(D)J recombination and tumorigenesis. *Nat. Genet.* 17: 479–482.

105. Ariumi, Y., M. Masutani, T.D. Copeland, T. Mimori, T. Sugimura, K. Shimotohno, K. Ueda, M. Hatanaka, M. Noda. 1999. Suppression of the poly(ADP-ribose) polymerase activity by DNA-dependent protein kinase in vitro. *Oncogene.* 18:4616–4625.

106. Bryant, H.E., N. Schultz, H.D. Thomas, K.M. Parker, D. Flower, E. Lopez, S. Kyle, M. Meuth, N.J. Curtin, T. Helleday. 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature.* 434:913–917.

107. Farmer, H., N. McCabe, C.J. Lord, A.N. Tutt, D.A. Johnson, T.B. Richardson, M. Santarosa, K.J. Dillon, I. Hickson, C. Knights, N.M. Martin, S.P. Jackson, G.C. Smith, A. Ashworth. 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature.* 434: 917–921.

108. Nevaldine,B., J.A. Longo, and P.J. Hahn. 1997. The SCID defect results in much slower repair of DNA double-strand breaks but not high levels of residual breaks. *Radiat. Res.* 147, 535–540.

109. DiBiase, S.J., Z.C. Zeng, R. Chen, T. Hyslop, W.J. Curran, and G. Iliakis. 2000. DNA- dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. *Cancer. Res.* 60: 1245–1253.

110. Audebert, M., B. Salles, and P. Calsou. 2004. Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. *J. Biol. Chem.* 279(53):55117-26.

111. Audebert, M., B. Salles, M. Weinfeld, and P. Calsou. 2006. Involvement of polynucleotide kinase in a poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway. *J. Mol. Biol.* 356(2): 257-65.

112. Balbo, A., and P. Schuck. 2005. Protein-Protein Interactions. Golemis, E., and P.D. Adams, editors. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, New York: 253-277.

113. Phizicky, E.M., and S. Fields. 1995. Protein-protein interactions: methods for detection and analysis. *Microbiol. Rev.* 59(1): 94-123.

114. De Las Rivas, J., and A. de Luis. 2004. Interactome Data and Databases: Different Types of Protein Interaction. *Comp. Funct. Genomics.* 5(2): 173–178

115. Waksman, G. (Ed.) 2005. Proteomics and Protein-Protein Interactions: Biology, Chemistry, Bioinformatics, and Drug Design. *Prot. Reviews.* 3: 1-19.

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

117. Ma, J. and M. Ptashne. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments *Cell.* 48: 847–853.

118. Bartel, P.L., and S. Fields. 1995. Analyzing protein-protein interactions using two-hybrid system. *Methods. Enzymol.* 254: 241-263.

119. Uetz, P., L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, v. Narayan, M. Sirnivasan, P. Pochart, *et al.* 2000. A comprehensive analysis of protein-protein interactions in *Saccaromyces cerevisiae. Nature.* 403: 623-627.

120. Criekinge, W.V., and R. Beyaert. 1999. Yeast Two-Hybrid: State of the Art. *Biol. Proced. Online.* 4(2): 1-38.

121. Vidalain, P.O., M. Boxem, H.G. Siming, and M. Vidal. 2004. Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods.* 32: 363–370.

122. McAlister-Henn, L., N. Gibson, and E. Panisko. 1999. Applications of the Yeast Two Hybrid System. *Methods.* 19(2): 330-337.

123. Ito, T., K. Ota, H. Kubota, Y. Yamaguch, T. Chiba, K. Sakuraba, and M. Yoshida. 2002. Roles for the two-hybrid system in exploration of the yeast protein interactome. *Mol. Cell. Proteomics.* 1(8):561-6.

124. Golemis, E.A, and R. Brent. 1992. Fused protein domains inhibit DNA binding by LexA. *Mol. Cell. Biol.* 12(7): 3006-14.

125. Evans, P.M., L. Woodbine, E. Riballo, A.R. Gennery, M. Hubank, and P.A. Jeggo. 2006. Radiation-induced delayed cell death in a hypomorphic Artemis cell line. *Hum. Mol. Genet.* 15(8): 1303-1311.

126. Clontech Laboratories Inc. 2007. Matchmaker™ Pretransformed Libraries User Manual. Protocol Number: PT3183-1. Version Number: PR6Y2132. http://www.clontech.com

127. Chen, E.H., E. Grote, W. Mohler, and A. Vignery. 2007. Cell-cell fusion. *FEBS. Lett.* 581(11):2181-93.

128. Invitrogen Life Technologies. 2003. Hybrid Hunter User manual. A two-hybrid system for analysis of protein-protein interactions in the yeast, *Saccharomyces cerevisiae.* http://www.invitrogen.com

129. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head bacteriophage T4. *Nature*. 227: 680-685.

130. Burnette, W, N. 1981. Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate—polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* 112: 195-203.

131. Kanazawa, S., M. Driscoll, and K. Struhl. 1988. ATR1, a *Saccharomyces cerevisiae* Gene Encoding a Transmembrane Protein Required for Aminotriazole Resistance. *Molec. Cell. Biol.* 8: 664-673.

132. Breeden, L., and K. Nasmyth. 1985. Regulation of Yeast HO Gene. Cold Spring Harbor Symp. *Quant. Biol.* 50: 643-650.

133. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215(3):403-10.

134. Clontech Laboratories Inc. 2007. Matchmaker™ Construction & Screening Kits User Manual. Cat. No. 630445 Prot. No. PT3955-1. Prot. Ver. PR792376. http://www.clontech.com

135. Soellick, T.R., and J.F. Uhrig. 2001. Development of an optimized interaction-mating protocol large-scale yeast two hybrid analyses. *Genome. Biol.* 2(12): RESEARCH 0052.1–0052.7.

136. Thermo Fisher Scientific Inc. Instructions: Y-DER® Yeast DNA Extraction Reagent Kit. Cat. No. 78870. http: www. piercenet.com

137. Marchler-Bauer, A., J.B. Anderson, M.K. Derbyshire, C. DeWeese-Scott, N.R. Gonzales, M. Gwadz, L. Hao, S. He, D.I. Hurwitz, J.D. Jackson, Z. Ke, D. Krylov, C.J. Lanczycki, C.A. Liebert, C. Liu, F. Lu, S. Lu, G.H. Marchler, M. Mullokandov, J.S. Song, N. Thanki, R.A. Yamashita, J.J. Yin, D. Zhang, and S.H. Bryant. 2007. *CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res.* 35: D237-40.

138. Li, J., and J. Yuan. 2008. Caspases in apoptosis and beyond. *Oncogene*. 27: 6194–6206.

139. Natesan, S., E. Molinari, V.M. Rivera, R.J. Rickles, and M. Gilman. 1999. A general strategy to enhance the potency of chimeric transcriptional activators. *Proc. Natl. Acad. Sci. U S A.* 96(24):13898-903.

140. Ruden, M.D. 1992. Activating regions of yeast transcription factors must have both acidic and hydrophobic amino acids. *Chromosoma.* 101: 342-348.

141. Ma, J., and M. Ptashne. 1987. A New Class of Yeast Transcriptional Activators. *Cell.* 51: 113-119.

142. Ruden, D.M., J. Ma, Y. Li, K. Wood, and M. Ptashne. 1991. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature.* 350: 250-252.

143. Cantin, G.T., J.L. Stevens, and A.J. Berk. 2003. Activation domain–mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc. Natl. Acad. Sci. U S A.* 100 (21): 12003-8.

144. Torroni, A., A. Achilli, V. Macaulay, M. Richards, and H.J. Bandelt. 2006. Harvesting the fruit of the human mtDNA tree. *Methods.* 32(4): 363-70.

145. Shoubridge, E.A. 2001. Cytochrome c oxidase deficiency. *Am. J. Med. Genet.* 106(1):46-52.

146. Chen, B., D. Zhong, and A. Monterio. 2006. Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. *BMC. Genomics.* 7 (156): 1-19.

147. Gibbons, R.J., D.J. Picketts, L. Villard, and D. R. Higgs. 1995. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with a-thalassemia (ATR-X syndrome). *Cell.* 80: 837–845.

148. Picketts, D.J., D. R. Higgs, S. Bachoo, D. J. Blake, W.J. O. Quarrell, and R. J. Gibbons. 1996. ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum. Mol. Genet.* 5(12): 1899–1907.

149. Xue,Y., R. Gibbons, Z. Yan, D. Yang, T. L. McDowell, S. Sechi, J. Qin, S. Zhou, D. Higgs, and W. Wang. 2003. The ATRX syndrome protein forms a chromatin-remodelling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc. Natl. Acad. Sci. U S A.* 100(19): 10635–10640.

150. Neely, K.E., and J. L. Workman. 2002. The complexity of chromatin remodelling and its links to cancer. *Biochem. Biophys. Acta.* 1603(1):19-29

151. Cardoso, C., S. Timsit, L. Villard, M. Khrestchatisky, M. Fontes, and L. Colleaux. 1998. Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. 7(4): 679-84.

152. Le Dourain, B., A.L. Nielsen, J.M. Garnier, H. Ichinose, F. Jeanouqin, R. Losson, and P. Chambon. 1996. A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. *EMBO. J.* 15(23): 6701-15.

153. Luijsterburg, M.S., C. Dinant, H. Lans, J. Stap E. Wiernasz, S. Lagerwerf, D.O. Warmerdam, M. Lindh, M.C. Brink, J.W. Dobruck, J.A. Aten, M.I Fousteri, G. Jansen, N.P. Dantuma, W. Vermeulen, L.H. Mullenders, A.B. Houtsmuller, P.J. Veschure, and R. van Driel. 2009. Heterochromatin protein 1 is recruited to various types of DNA damage. *J. Cell. Biol.* 185(4):577-86.

154. Varambally, S., S.M. Dhanasekaran, M. Zhou, T. R. Barrette, C. Kumar-Sinha, M. G. Sanda, D. Ghoshk, K. J. Pienta, R. G. A. B. Sewalt, A. P. Otte, M. A. Rubin, and A. M. Chinnaiyan. 2002. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature.* 419: 624-629.

155. Vidal, M. and P. Legrain. 1999. Yeast forward and reverse 'n'-hybris systems. *Nucleic. Acids. Res.* 27(4): 919-929.

156. Formont-Racine, M. J.C. Rain, and P. Legrain. 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat. Genet.* 16: 277-282.

157. Hughes, M.A., C.W. Lee, C.F. Holm, S. Ghosh, A. Mills, L.A. Lockhart, S.L. Reed, and B.J. Mann. 2003. Identification of Entamoeba histolytica thiol-specific antioxidant as a GalNAc lectin-associated protein. *Mol. Biochem. Parasitol.* 127(2):113-20.

158. Battersby, A., A. Csiszar, M. Leptin, and R. Wilson. 2003. Isolation of Proteins that Interact with the Signal Transduction Molecule Dof and Identification of a Functional Domain Conserved between Dof and Vertebrate BCAP. *J. Mol. Biol.* 329: 479-493.

159. El Housni, H., I. Vandenbroere, D. Perez-Morga, D. Christophe, and I. Pirson. 1998. A Rare Case of False Positives in a Yeast Two-Hybrid Screening: The Selection of Rearranged Bait Constructs That Produce a Functinoal Gal4 Activity. *Anal Biochem.* 262(1):94-6.

160. Erhart, E. and C.P. Hollenberg. 1983. The Presence of a Defective LEU2 Gene on 2 μ DNA Recombinant Plasmids of *Saccharomyces cerevisiae* Is Responsible for Curing and High Copy Number. *J. Bacteriol.* 156: 625-635.

161. Hsu, Y. P., and G. B. Kohlhaw. 1982. Overproduction and control of the LEU2 gene product 3-isopropylmalate dehydrogenase in transformed yeast strains. *J. Biol. Chem.* 257:39-41.

162. Loison, G., A. Vidal, A. Findeli, C. Roitsch, J.M. Balloul, Y. Lemoine. 1989. High level of expression of a protective antigen of schistosomes in Saccharomyces cerevisiae. *Yeast.* 5(6): 497–507.

163. Gunge, N. 1983. DNA Plasmids. *Ann. Rev. Microbiol.* 37: 253-76.

164. Hollenberg, C. P., A. Degelmann, B. Kustermann-Kuhn, and D. H. Royer. 1976. Characterization of 2-μ.m DNA of Saccharomyces cerevisiae by restriction fragment analysis and integration in an Escherichia coli plasmid. *Proc. Natl. Acad. Sci.* 73:2072-2076.

165. Gerbaud, C., P. Fournier, H. Blanc, M. Aigle, H. Heslot, and M. Guerlneau. 1979. High frequency of yeast transformation by plasmids carrying part or entire 2-μm yeast plasmid. *Gene*. 5:233-253.

166. Struhl, K., D. T. Stinchcomb, S. Scherer, and R. W. Davis. 1979. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Nati. Acad. Sci.* 76: 1035-1039.

167. Hohmann, S., and P.W.H. Mager. 2003. Yeast Stress Responses. © Springer-Verlag Berlin Heidelberg. *Top. Curr. Genet.* 1:1-9.

168. Royer, W.E.Jr., J.E. Knapp, K. Strand, and H.A. Heaslet. 2001. Cooperative hemoglobins: conserved fold, diverse quaternary assemblies and allosteric mechanisms. *Trends. Biochem. Sci.* 26(5): 297-304.

169. Littman, G.W., J.P. Rast, M.J. Shamblott, R.N. Haire, M. Hulst, W. Roess, R.T. Litman, K.R. Hinds-Frey, A. Zilch, and C.T. Amemiyag. 1993. Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. *Mol Biol Evol.* (1): 60-72.

170. Slee, E.A., C. Adrain, and S.J. Martin. 2002. Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis. *J. Biol. Chem.* 279(2): 1030-1039.

171. Walsh, J.G., S. P. Cullen, C. Sheridan, A. U. Lüthi, C. Gerner, and S. J. Martin. 2008. Executioner caspase-3 and caspase-7 are functionally distinct proteases. *Proc. Natl. Acad. Sci. U S A.* 105(35): 12815-12819.

172. Araya, R., R. Takahashi, and Y. Nomura. 2002. Yeast two-hybrid screening using constitutive-active caspase-7 as bait in the identification of PA28g as an effector caspase substrate. *Cell. Death. Differ.* 9: 322-328.

173. Stennicke, H.R., M. Renatus, M. Meldal, and G.S. Salvesen. 2000. Internally quenched fluorescent peptide substrates disclose the subsite preferences of human caspases 1, 3, 6, 7 and 8. *Biochem. J.* 350:563–568.

174. Srinivasula, S.M., M. Ahmad, M. MacFarlane, Z. Luo, Z. Huang, T. F. Alnemri, and E. S. Alnemri. 1998. Generation of Constitutively Active Recombinant Caspases-3 and -6 by Rearrangement of Their Subunits. *J. Biol. Chem.* 273(17):10107-10111.

175. Tina, K.G., R. Bhadra, and N. Srinivasan. 2007. PIC: Protein Interactions Calculator. *Nucleic. Acids. Res.* 35: W473–W476.

176. Pakula, A.A., and R.T. Sauer. 1989. Genetic analysis of protein stability and function. *Anna. Rev. Genet.* 23: 289-30.

177. Petsko, G.A., and D. Ringe. 2004. Protein Structure and Function. London, New Science Press LTD.

178. Dlakic, M. 2000. Functionally unrelated signalling proteins contain a fold similar to Mg^{2+}-dependent endonucleases. *Trends. Biochem. Sci.* 25(6): 272-3.

179. Xiong, Y., and T. H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO. J.* 9(10): 3353–3362.

180. Parseval, N. De. and T. Heidmann. 2005. Human endogenous retroviruses: from infectious elements to human genes. *Cytogenet. Genome. Res.* 110: 318–332.

181. Kazazian, H.H. Jr. 2004. Mobile elements: Drivers of genome evolution. *Science.* 303: 1626-32.

182. Müller-Sieburg, C.E., R. H. Cho, M. Thoman, B. Adkins, and H. B. Sieburg. 2002. Deterministic regulation of hematopoietic stem cell self-renewal and differentiation. *Blood.* 100 (4): 1302-1309.

183. Matviw, H., G. Yu, and D. Young. 1992. Identification of a human cDNA encoding a protein that is structurally and functionally related to the yeast adenylyl cyclase-associated CAP proteins. *Mol. Cell. Biol.* 12(11): 5033-40.

184. Masson, M., C. Niedergang, V. Schreiber, S. Muller, J. Menissier-de Murcia, and G. de Murcia. 1998. XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA-damage. *Mol. Cell. Biol.* 18:3563–3571.

185. Leppard, J.B., Z. Dong, Z.B. Mackey, and A. E. Tomkinson. 2003. Physical and Functional Interaction between DNA Ligase III and Poly(ADP-Ribose) Polymerase 1 in DNA Single-Strand Break Repair. *Mol. Cell. Biol.* 23(16):5919-27.

186. Masson, M., J. Menissier-de Murcia, M.G. Mattei, G. de Murcia, and C.P. Niedergang. 1997. Poly(ADP-ribose) polymerase interacts with a novel human ubiquitin conjugating enzyme: hUbc9. *Gene.* 190(2):287-96.

187. Song, D., S. Sakamoto, and T. Taniguchi. 2002. Inhibition of Poly(ADP-ribose) Polymerase Activity by Bcl-2 in Association with the Ribosomal Protein S3a. *Biochemistry.* 41(3): 929–934.

188. Soares, M.B., M.F. Bonaldo, P. Jelene, L. Su, L. Lawton, and A. Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. U S A.* 91(20): 9228–9232.

189. Bogdanova, E.A., D.A. Shagin, and S.A. Lukyanov. 2008. Normalization of full-length enriched cDNA. *Mol. Biosyst.* 4(3):205-12.

190. Holz, C., A. Leuking, L. Bovekamp, C. Gutjar, N. Bolotina, H. Lehrach, and D.J. Cahill. 2001. A human cDNA expression library in yeast enriched for open reading frames. *Genome. Res.* 11(10): 1730-5.

191. Suzuki, Y. and S. Sugano. 2001. Construction of full-length-enriched cDNA libraries. The oligo-capping method. *Methods. Mol. Biol.* 175:143-53.

192. Das, M., I. Harvey, L.L. Chu., M. Sinha, and J. Pelletier. 2001. Full-length cDNAs: more than just reaching the ends. *Physiol. Genomics.* 6(2): 57-80.

193. Davis, C.A. and S. Benzer. 1997. Generation of cDNA expression libraries enriched for in-frame sequences. *Proc. Natl. Acad. Sci. U S A.* 94(6): 2128-32.

CURRICULUM VITAE

Rahin Farzadfar

EDUCATION

M.Sc. candidate in Biochemistry (specialization in human and molecular genetics)

University of Ottawa, Faculty of Medicine
Department of Biochemistry, Microbiology and Immunology
Ottawa, Ontario

- Thesis: Identification of Artemis and PARP-1 interacting factors: A yeast two hybrid screening approach.

B.Sc. with Honours in Biochemistry

Carleton University, Faculty of Science
Institue of Biochemistry
Ottawa, Ontario.

- Graduated April 2005.
- Thesis: Identification of novel genes involved in DNA double strand break repair in yeast *Saccharomyces cerevisiae.*

ABSTRACTS AND POSTERS

- Farzadfar, R., Soubeyrand, S., and R.J.G. Haché. (2008). Identification of Artemis and PARP-1 interacting factors: A yeast two hybrid screening approach.

- Farzadfar, R., Soubeyrand S., and R.J.G. Haché. (2007) Identification of Artemis and PARP-1 interacting factors: A yeast two hybrid screening approach.

- Farzadfar, R., and A. Golshani. (2005). Identification of novel genes involved in DNA double strand break repair in yeast *Saccharomyces cerevisiae.*

AWARDS AND SCHOLARSHIPS

- Carleton University Admission Scholarhip 2000
- Hyman Soloway Scholarship 2001

RESEARCH EXPERIENCE

Ottawa Health Research Institute
University of Ottawa
Ottawa, Ontario

Position:	M.Sc. candidate
Duties:	Identification of Artemis and PARP-1 interacting factors to gain insight about their function and regulation, using both classical and interactive mating yeast two hybrid screening systems. Artemis and PARP-1 are both proteins with implications in the DNA Damage Response network.
Supervisor:	Dr. Robert Haché

Functional Genomics and Systems Biology
Carleton University
Ottawa, Ontario

Position:	B.Sc. candidate
Duties:	Identification of novel factors involved in the DNA repair and the maintenance of genomic integrity in the budding yeast, *S. cerevisiae* A synthetic genetic array (SGA) and an *in vivo* plasmid assay were used to identify novel genes involved in the non-homologous end joining (NHEJ) pathway of yeast.
Supervisor:	Dr. Ashkan Golshani

REFERENCES

Available Upon Request