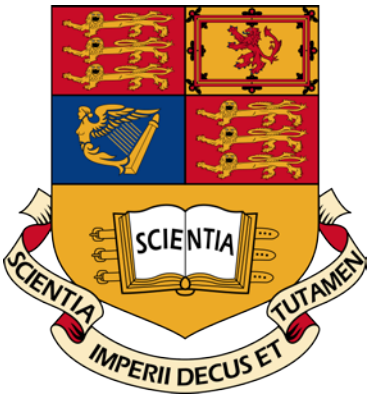


IMPERIAL COLLEGE LONDON
FACULTY OF MEDICINE
DEPARTMENT OF MEDICINE
DIVISION OF EXPERIMENTAL MEDICINE

Using the RISCI Genetic Screening Platform for Elucidating Apoptosis Signalling Network

A thesis submitted for the degree of Doctor of Philosophy

LIN Kuan Chee Bevan

16 September 2012

Supervisors
Professor Stefan Grimm
Dr John Wharton

Declaration of Originality

I, LIN Kuan Chee Bevan, hereby declare that I am the sole author of this thesis, the whole of which constitutes original research work conducted at the Department of Medicine, Imperial College London. Any section which has been previously published or submitted for publication has been duly referenced.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any figures, data, ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the United Kingdom Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and Imperial College London, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Abstract

Considerable development in the field of nanotechnology is increasingly yielding novel applications of nanoparticles. The unique properties of nanoparticles in particular their high aspect ratio (length : width ratio), however could pose potential risks to the user. A high throughput genetic screening platform, RISCI (robotic single cDNA investigation), was previously established for the systematic evaluation of single gene activities. Here, RISCI was utilised to identify pro-apoptotic genes as well as genes involved in the positive and negative regulation of silica nanoparticle-induced cell death.

This project describes the further development of the screening platform by harnessing its capability to screen a cDNA library comprising approximately 30,000 full length, completely annotated, and sequenced human genes for novel regulators of apoptosis. It integrates an extensive skill sets and is broadly organised into three major phases: Setup, Screen and Analysis. The integration of a pro-apoptosis treatment to screen for inhibitors and sensitizers is a novel aspect of the current experimental setup, along with the low redundancy library.

The extensive setup phase focused on technical aspects. The cDNA library, acquired as plasmid DNA, was transformed into a bacterial host for replication and subsequent DNA isolation. A new high-throughput process was developed encompassing the production of competent bacteria and a heat shock transformation protocol, which was subsequently transferred onto the robotic platform. In parallel, the software controlling the robots was redeveloped to allow for execution of user-defined protocols while novel transfection protocols were adapted for automation.

The screen identified 699 apoptosis inducers, 1,141 inhibitors and 626 sensitizers. Bioinformatics analysis revealed that the inducers were highly enriched for cell death associated terms, while the inhibitors were strongly associated with cancer profiles. Both inducers and sensitizers were predominantly achieving the functional effect on the protein level, but inhibitors were mainly transcription based. Enriched metal response genes also suggest that the silica nanoparticles were causing their toxicity through reactive oxygen species generation. Intriguingly, the screen identified many noncoding sequences as being functionally capable of regulating apoptosis. These noncoding candidates are capable of regulating the protein coding counterparts identified from the screen.

The truly interesting part of the project outcome remains those unknown candidates that were implicated in apoptosis regulation for the first time. Dissemination of the consolidated candidate list would help accelerate the experimental validation of these candidates and aid other researchers in deriving novel hypotheses when the candidates are placed in their research context.

Table of Contents

Acknowledgement

This project was generously supported by two awards from Johnson & Johnson and the Medical Research Council, United Kingdom. The Johnson & Johnson grant supported the initial phase including work on the setup of the RISCI screen, establishment of the NITE library under the bacterial host and autophagy. The Medical Research Council supported the work on the nanoparticles toxicity and the implementation of the RISCI screen.

The project was the cumulative result of guidance and support from several individuals, chief among them was my supervisor Professor Stefan Grimm whose dedication and continuous guidance on the project proved invaluable. His passion for research and firm belief in the RISCI screen was crucial in driving the research forward despite the many hurdles which infringed upon the progress. Many thanks go to Dr Georg Sindelar and Dr Volker Kachel for sharing their deep knowledge and experience on working with the pair of unique robots. Dr Kachel was extremely helpful and supported the repair of the robots on numerous occasions when systems failed, often travelling from Germany to assist with the repairs.

The post docs Dr Shazia Irshad, Dr Anthony Lemarié, Dr Laurence Huc-Lemarié, Dr Foy Carpenter, Dr Anne-Laure Mahul and Dr Nick Kassouf whom were always willing to share their expertise and advice on the experiments. Dr Anne-Laure Mahul was imperative in managing the team and keeping the lab in working order. The PhD and summer students of the Apoptosis Group Dr Ryota Iwasawa, Dr Evangelos Pazarentzos, Christoph Datler, Ghada AbuAli, Wanwisa Chaisaklert, Ding Qize, Ming Hwang, Friederike Feldmann, Jan Kroon, Ellen Stelloo, Patricia Jaaks and Birol Çabukusta all of whom helped to keep the group lively and provided the supportive environment to encourage each other along the way.

The affiliates of the Imperial College Bioinformatics Support Service, Dr Derek Huntley, whom was instrumental in helping me along initially as I explored the bioinformatics sphere. Dr Huntley also provided his expertise in the annotation of the NITE library and supported the bioinformatics work for our publication. Arshad Khan and Geraint Barton both of whom shared their knowledge of microarrays with me. Dr Alberto Polleri for sharing his advice on analysis approaches and being our guide to bioinformaticians and statisticians around the Hammersmith Campus. Dr Enrico Petretto of the MRC Clinical Sciences Centre, Imperial College London and his PhD student Sarah Langley both of whom advised and contributed their expertise on microarrays and differential analysis, another key area of analysis for our publication. Dr Petretto was particular thoughtful and on multiple occasions

made his time available for our discussions on the mathematical aspects including the many equations for the candidate selection. His advice was extremely important in maintaining the quality of analysis.

Our collaborators for the MRC funded research on nanoparticles toxicity. Professor Terry Tetley for her advice on use of nanoparticles, Dr Eva Valsami-Jones and Dr Agnieszka Dybowska from the Natural History Museum London for the characterisation of nanoparticles we tested and used, Dr Odu Okoturo and Dr Robert Edwards for their work on proteomics profiling and Professor Alan Boobis for coordinating the project.

This thesis is dedicated to my family. To mum and dad (Yvonne and Simon Lin), your unyielding affection, understanding and support helped me through the darkest moments of my life to make this endeavour a success, and to my brothers Douglas and Arthur both of whom are equally lovable and supportive.

Publications and Conferences

Publications

Genetic cell culture screens reveal mitochondrial apoptosis control.

Grimm S, Lin B.

Front Biosci. 2009 Jan 1;14:1471-8.

Determining signalling nodes for apoptosis by a genetic high-throughput screen.

Lin B, Huntley D, Abuali G, Langley SR, Sindelar G, Petretto E, Butcher S, Grimm S.

PLoS One. 2011;6(9):e25023. Epub 2011 Sep 22.

De-ubiquitinating proteases USP2a and USP2c cause apoptosis by stabilising RIP1.

Mahul-Mellier AL, Datler C, Pazarentzos E, Lin B, Chaisaklert W, Abuali G, Grimm S.

Biochim Biophys Acta. 2012 Aug;1823(8):1353-65. Epub 2012 May 30.

De-ubiquitinating protease USP2a targets RIP1 and TRAF2 to mediate cell death by TNF.

Mahul-Mellier AL, Pazarentzos E, Datler C, Iwasawa R, AbuAli G, Lin B, Grimm S.

Cell Death Differ. 2012 May;19(5):891-9. doi: 10.1038/cdd.2011.185. Epub 2011 Dec 16.

* Pending publication for the work discussed in this thesis.

Posters

Archaeal insights into chaperone-mediated protein folding

Andrew Large, Bevan Lin, Preethy Gowrinathan, Riddhi Shah and Peter Lund

Investigating apoptosis and toxicity signalling with RISCI

Bevan Lin, Odu Okoturo, Ghada AbuAli, Robert Edwards, Alan Boobis and Stefan Grimm

Presented at the following conferences:

12-16 November 2011	ACR-NCI-EORTC International Conference Molecular Targets and Cancer Therapeutics Moscone Center West in San Francisco, CA
21 March 2012	High Throughput/High Content Technology Symposium, Imperial College London
23 April 2012	Young Scientist Day, Department of Medicine, Imperial College London

Abbreviations

List of the Abbreviations and Acronyms commonly used in this thesis. Specific gene names are excluded and will be separately described in the associated text.

A549	Human alveolar basal epithelial cells
aa	amino acid
AIDS	Human immunodeficiency virus infection / acquired immunodeficiency syndrome
Amp	Ampicillin
BCA	Bicinchoninic Acid (Assay)
Bgal	β-galactosidase
BIND	Biomolecular Interaction Network Database
BIR	Baculoviral IAP repeat domain
bp	base pairs
BSA	Bovine serum albumin
CARD	Caspase-recruitment domain
CAT	Chloramphenicol acetyl transferase (Assay)
CFP	Cyan Fluorescent Protein
CMV	Cytomegalovirus (promoter)
CPRG	Chlorophenolred-β-D galacto-pyranoside (Assay)
DD	Death domain
DDJB	DNA database of Japan
DED	Death effector domain
DiOC6	3,3-dihexaoxacarbocyanine Iodide
DISC	Death-inducing signalling complex
DLS	Dynamic light scattering
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
DWE	Distance-weighted estimator
DWP	Deep well plate
EBSS	Earle's Balanced Salt Solution
ELISA	Enzyme-linked immunosorbent assay
ENCODE	The Encyclopedia of DNA Elements Consortium
ER	Endoplasmic reticulum
EST(s)	Expressed sequence tag(s)
FCS	Fetal calf serum
FDR	False discovery rate
FLJ	Full-length long Japan
FRET	Fluorescence resonance energy transfer
FTIR	Fourier transform infrared spectroscopy
GENCODE	Encyclopaedia of genes and gene variants
GFP	Green Fluorescent Protein
GO	Gene ontology

GSEA	Gene set enrichment analysis
GUI	Graphical user interface
HA	Hemagglutinin
HD	Huntington's Disease
HEK	human embryonic kidney
HeLa	Human cervical carcinoma
HSP(s)	Heat shock protein(s)
HTS	High throughput
IAP(s)	Inhibitor(s) of apoptosis
IPA	Ingenuity® Pathway Analysis
IRES	Internal ribosome entry site
LB	Lysogeny broth
LDH	Lactate dehydrogenase (Assay)
lncRNA	Long noncoding RNA
LPS	Lipo-polysaccharides
MINT	Molecular INTeraction database
miRNA	microRNA
MOMP	Mitochondrial outer membrane permeabilisation
MRC	Medical Research Council
mRNA	Messenger RNA
MTP	Microtitre plates
MTT	Dimethyl thiazolyl diphenyl tetrazolium salt (Assay)
NCBI	National Center for Biotechnology Information
ncRNA	noncoding RNA
NITE	National Institute of Technology and Evaluation
NK	Natural killer
NOD	Nucleotide binding and oligomerisation domain
NP	Nanoparticles
OD	Optical density
PBS	Phosphate buffered saline
PCD	Programmed cell death
PCR	Polymerase chain reaction
PEI	Polyethyleneimines
pGL	pGreenLantern-1
PI	Propidium iodide
PIR	Protein Information Resources
PTP	Permeability transition pore
RegRNA	Regulatory RNA Motifs and Elements Finder
RISCI	Robotic single cDNA investigation
RNA	Ribonucleic acid
ROI	Region of interest
ROS	Reactive oxygen species
rRNA	ribosomal RNA
SDS	Sodium dodecyl sulphate
snoRNA	Small nucleolar RNA
snRNA	small nuclear RNA

SRB	Sulforhodamine B (Assay)
SV40	Simian virus 40 (promoter)
TAE	Tris-acetate-EDTA (Buffer)
TCL	Cytotoxic T cells
TEM	Transmission electron microscopy
TNF	Tumour necrosis factor
TNTD	Treated-non-treated difference
TNTR	Treated-non-treated ratio
tRNA	transfer RNA
UCSC	University of California Santa Cruz
UTR	Untranslated region
VEE	Visual Engineering Environment
WRD	WD40 repeat domain
XDR	X-ray diffraction
YFP	Yellow Fluorescent Protein

Chapter 1: Introduction

Apoptosis

History and Highlights

Cell death is a concluding event of any living system, an essential event which requires delicate management within multicellular organisms. Apoptosis represents one of the programmed cell death mechanisms evolved by nature to manage the systematic clearance of dying cells.

Apoptotic phenotypes were first described in principle by the German scientist Carl Vogt in 1842. The concept and eventual term of programmed cell death began taking hold only in the mid-ninetieth century after observations indicating the crucial role of cell death in the development of multicellular organisms [1, 2]. “Apoptosis" was coined in 1972 by Kerr, Wyllie and Currie as a term to describe the processes leading up to the controlled and reproducible pattern of cell death, analogous to the poetic description of “leaves falling away from a tree” which the term draws from its Greek origin [3]. Around the same time, pioneering works in the establishment of the nematode *Caenorhabditis elegans* as a model system by Brenner and Sulston et al. [4, 5], and discoveries in subsequent decades of the mechanistic regulations by cell death (“ced”) genes by Horvitz et al. in the C. elegans system [6, 7] and the Apoptosome complex by Wang et al. [8, 9] established some of the fundamental molecular controls of apoptosis signalling.

Today, apoptosis research is an extremely active field, with more than 70,000 publications [10] since its initial conception detailing the vast signalling networks and physiological and clinical relevance. Cells may decide to undergo apoptosis for the benefit of the organism as a whole, a classical example of which is the activation of apoptosis during embryonic development resulting in proper deletion of cells for morphogenesis. Cells damaged via external insults such as virus or encountering DNA replication or repair impairment may also undergo apoptosis as a way of damage limitation [11]. The process may also be forced upon cells, when employed by the immune system during development of the T cells antigen recognition in the thymus, or by cytotoxic T cells (TCL) and natural killer (NK) cells in the removal of compromised cells [12]. Since the proliferation of cells in multicellular organisms such as humans via mitosis is actively balanced by the selective destruction via apoptosis, inappropriate changes to the rate of apoptosis would negatively impact normal physiology. In fact, numerous disorders such as neurodegenerative disease, AIDS, myocardial infraction and atherosclerosis are associated with excessive apoptosis, while cancer and autoimmune diseases are often the result of suppression of apoptosis.

Expanding on the vast knowledge of the apoptosis network not only offers insights into its molecular regulation and disease implications, but also allows for this knowledge to be exploited for the detection and treatment of the damaging disorders resulting from inappropriate apoptosis within a clinical context.

Apoptosis, Necrosis and Autophagy

Since its discovery, the elegant and controlled disposal of dying cells by apoptosis is often compared with necrosis, another classical form of cell death associated with uncontrolled destruction of the cells, to reinforce the morphological distinction of the two types of cell deaths. Necrosis almost always occurs under pathological conditions, while apoptosis can take place under both healthy and disease states. Specialised publications may refer to necrosis as the end stage phenotype, while the sub lethal process prior to necrosis is referred to oncosis [13]; for simplicity, necrosis is used to describe both the process and final phenotype.

Necrosis is a passive and energy-independent process often the result of excessive fluid influx into the cell; hence the effect of necrosis usually extends beyond individual cells to affect a large area. As a result, the cells undergo uncontrolled swelling, accompanied by swollen organelles such as the mitochondria, endoplasmic reticulum & Golgi, disrupted membranes, ruptured lysosomes, detachment of ribosomes, and eventual destruction of the cell membrane integrity [3, 13, 14]. Cells or tissues undergoing necrosis may appear visibly cloudy as a result of reversible denaturation of proteins. The cell lysis causes cellular constituents to leak into the surrounding environment, leading to the recruitment of immune cells. Such response causes inflammation and extensive cellular damages.

In contrast, cells undergoing apoptosis display several characteristics phenotypes such as shrinkage of the cytosol (in contrast to swelling during necrosis), chromatin condensation and subsequent DNA cleavage, mitochondrial depolarisation, display of phosphatidyl serine on the plasma membrane surface, loss of symmetry and membrane blebbing [3, 15]. Morphologically, apoptosis can be classified into three distinct phases [16].

Phase I begins after initiation of apoptosis, whereby the apoptotic cells detach from adjacent cells and the surrounding matrix. The stimuli leads to molecular changes including the activation of the permeability transition (PT) pore [17], hence the loss of mitochondrial membrane potential and release of cytochrome c [18] and apoptosis-inducing factor [19] from the mitochondria intermembrane space. Surface structures of these detaching cells such as microvilli and junction

complexes are also lost as they round up due to the activated caspases and proteases breaking down their substrates such as the cytoskeleton structures.

Under the light microscope, chromatin condensation can be observed as crescent-shaped bodies, this event is followed by the activation of endonucleases which causes DNA fragmentation. DNA fragmentation is another hallmark of apoptosis, which appears with 180- to 200-bp intervals. DNAse I and H, and a calcium/magnesium-dependent endonucleases may be involved in this event, although the specific combination of cleavage enzymes is dependent of the activation stimuli [20]. Distribution of phosphatidyl serine from to the cell surface also occurs as a signal for macrophages to take up apoptotic bodies [21].

During Phase II, membrane blebbing can be observed as the cells produce pseudopodia containing nuclear fragments which then bud off [22]. What remains are the purported apoptotic bodies characterised by their smooth and round appearance. Phase III concludes the process with the membrane becoming permeable to dyes like Tryphan Blue, and the apoptotic bodies and vesicles being phagocytosed by the surrounding cells or by immune cells such as marcophages.

Apoptosis can be triggered by a variety of stimuli, but the appearance of the same cellular changes indicates a convergence of the signalling pathways for apoptosis into a conserved death effector machinery [23]. This makes it possible to employ such changes as markers of apoptosis and techniques like the detection of the membrane phosphatidyl serine with annexin V [24, 25], measuring mitochondrial enzyme activity with the MTT assay, and staining DNA with intercalating dyes or labelled nucleotides have become common techniques in apoptosis research [26, 27]. Such molecular changes that can be easily exploited as sensitive detection assays are extremely important in a screen.

In recent years however, the distinction between these two major forms of cell death is becoming increasingly blurred. It is now recognised that apoptosis is one of many forms of programmed cell death (PCD). Apoptotic cell death or type I PCD is defined by its main characteristics being the apoptotic phenotypes such as caspase activation, membrane blebbing and DNA fragmentation.

Autophagy or type II PCD differs morphologically from apoptosis as it is independent of caspase activation and other apoptotic phenotypes. The process depends on the formation of double membrane autophagosomes and autolysosomes, and features cells with an intact nucleus. Autophagy functions primarily as the cell survival response to growth factor or nutrient starvation by degrading macromolecules such as proteins. Studies have also indicated that the process is capable of recycling materials from damaged organelles [28], and its molecular mechanisms have been

associated with induction of tumour specific cell death and suppression of neurodegenerative diseases [29]. Both apoptosis and autophagy, while being distinct in morphologies, are evolutionarily conserved. Their signalling networks have become ever more linked through recent studies that indicate intricate and extensive cross-talking between them. In certain instances, both act synergistically, while in others, activation of autophagy can occur only with the suppression of apoptosis [30]. Their close interplay and ultimate control over the cell survival status could be exploited by tumour cells through the same trigger [31], hence also offers valuable targets for treatments [32].

As the understanding of the cell signalling network deepens with the exponential data generation approach offered by high-throughput research, greater overlaps between previously "defined" processes would likely emerge, providing more insights into the coordination of apoptosis and autophagy with other forms of programmed cell death, such as necropoptosis which displays a combination of features from apoptosis and necrosis [33, 34].

Apoptosis Signalling Network

CED-3, CED-4 and CED-9, discovered in *C. elegans*, were some of the first genes to be associated with apoptosis signalling. The human homologues of these genes were found to be caspase-9, apaf-1 and bcl-2 and constitute the core components for the activation of apoptosis [35]. Decades of extensive research have since expanded this into a vast network, which comprises a delicate equilibrium of pro- and anti-apoptotic signals under normal cellular conditions. The flow of the apoptosis pathway begins with a signal trigger, which could be the aggregation of death ligands or a stress response such as after DNA damage or reactive oxygen species accumulation. Upon activation by these signals, initiation complexes and platforms are formed from an ensemble of adaptors and initiation proteins. This in turn activates cascades for signal amplification which ultimately commits the cells to apoptosis with the activation of effector proteins such as proteases and endonucleases for the disassembly of the cells.

Extrinsic and Intrinsic Pathways

The extrinsic pathways of apoptosis are initiated with the binding of the death receptors such as the Fas and TNF receptors to their ligands, by which the apoptosis signals are triggered externally. The extracellular death ligands are homotrimetric, thus the binding to their corresponding receptors would form ligand-receptor complexes that are at least homotrimeric. These complexes in turn recruit, via their death domain (DD), cytosolic factors and adaptor proteins containing the death effector domain (DED), providing a platform which increases the concentration of DED-containing

initiator caspases within the local proximity to enable auto-catalysis in a model known as “induced proximity” [36]. Activated caspases then cleave and activate effector caspases and other substrates to execute apoptosis. An example is the Fas receptor (FasR), which upon binding to the Fas ligand (FasL), recruits FADD and procaspase-8 to form the death-inducing signalling complex (DISC) [37, 38], upon which caspase-8 auto-activates. The DISC complex is functionally homologous to the apoptosome, which is central to the intrinsic pathway, and convergence exists between the extrinsic and intrinsic pathways through for example the cleavage of Bid that is mediated by caspase-8. Cleaved BID (tBID) interacts with Bax and promotes its insertion into the mitochondria to trigger the release of pro-apoptotic factors [36, 39, 40].

The intrinsic pathways, centred on the mitochondria, involve signals triggered from within the cells as a result of stress or oncogene activation, leading up to the loss of mitochondria integrity and release of pro-apoptotic factors. The mitochondria intermembrane space comprises a concoction of pro-apoptotic factors such as cytochrome c, SMAC/DIABLO, endonuclease G, Omi/HtrA2 and AIF. The permeabilisation of the outer mitochondria membrane is tightly regulated by various factors including formation of the permeability transition (PT) pore, members of the Bcl-2 superfamily and mitochondrial lipids [41]. The disruption of the mitochondria outer membrane causes the pro-apoptotic proteins to enter the cytosol to promote apoptosis [42]. The classical route is the binding of the released cytochrome c to the cytosolic APAF1, which causes a conformation change enabling ATP binding and heptamerisation to form the apoptosome [8, 43], which forms the platform on which caspase-9 becomes activated [8, 44] and triggers downstream caspase cascade including caspase-3 activation. Increasingly, mediation of the intrinsic pathway independent of the apoptosome and caspases is also beginning to emerge [45]. The inactivation of the intrinsic pathway is the primary mechanism by which oncogenes achieve apoptosis evasion, hence is an important hallmark of cancer [46].

Caspases

The caspase family of proteins is the centrepiece within the apoptosis signalling response. Caspases are cysteine proteases which cleave their substrate after an aspartate residue [47]. Caspase-1 or interleukin-1 –converting enzyme (ICE) was the first caspase to be discovered in humans [48] while the molecular mechanisms of caspases were determined in *C. elegans* [7]. Caspases are synthesised as catalytically inactive zymogens and can be broadly divided into the initiator and effector caspases. Mammalian initiator caspases consist of caspase-2, -8, -9 and -10 and are activated by apoptotic platforms like the apoptosome to initiate the caspase cascade. Caspases-3, -6 and -7 are the effector

caspases, which are downstream of the caspase cascade activated by the initiator caspases to cleave a broad range of substrates and commit the cell to apoptosis.

Structurally, both initiator and effector caspases exist as homodimers. The initiator caspases contain adaptor motifs such as the caspase-recruitment domain (CARD) and death-effector domain (DED) in their N-terminal prodomain. These adaptor domains are absent in the effector caspases, which have short amino acid sequences of 20 – 30 residues in its place. During apoptosis, the initiator caspases auto-activates upon recruitment to the apoptotic activation platforms, which in turns proceed to proteolytically activate the effector caspases. The molecular mechanism for the activation of effector caspases is derived from structural information of caspase-7 [49]. The core structure of the pro- and activated caspase remained unchanged. Instead, the reconstitution of the catalytic activity is dependent on the loop structures, most importantly the L2 loop. Under the natural conformation of the pro-caspase, the L2 loop is locked in a closed conformation; the cleavage of L2 loop allows it to move into a new conformation which stabilises the active site, thus conferring catalytic activity to the caspase by several orders of magnitude. Substrate recognition is also dependent on these loop structures, which binding pockets frequently found in L1, L3 and L4 loops [50]. Caspases recognise a sequence of four amino acids, P4-P3-P2-P1, cleaving the protein after the P1 position. Aspartate is usually the preferred residue in the P1 position, but it has been found that glutamate can take this position without abolishing catalytic ability in the case of the Drosophila Dronc [51]. Glutamate is also the preferred residue for P3 in all caspases studied [50, 52] while the residue for P4 differs between caspases and is thought to determine the substrate specificity.

Since the activation of caspases ultimately commits the cell to irreversible cell death, the process is tightly regulated especially for the initiator caspases which requires assembly of multimeric complexes. Caspases are regulated transcriptionally and post-translationally, but its catalytic activity can also be inhibited by the Inhibitor of Apoptosis (IAP) family of proteins [53]. Viruses are also known to exploit caspases inhibition to bypass apoptosis such as through the baculoviral protein p35 [54], p49 [55] or the cowpox virus serpin CrmA [56].

Inhibitor of Apoptosis (IAPs) Family

Proteins in the IAPs family possess one to three copies of the baculoviral IAP repeat (BIR) domain, which is a cysteine and histidine rich, zinc chelating protein domain around the N terminus. In addition, several mammalian IAPs also possess addition domains such as the RING zinc-finger (RZF) or CARD around the C terminus. The IAPs comprises a subfamily of the larger BIR containing protein family, with the additional functional capacity to inhibit apoptosis. Notable members of the IAPs family are X-chromosome-linked IAP (XIAP), c-IAP1, c-IAP2, NAIP which possess three BIR domains

and one RZF domain, while single BIR containing members includes Survivin, testis-specific IAP (Ts-IAP) and Livin [57].

IAPs regulate apoptosis inhibition through a number of mechanisms, one of which is the direct binding and inhibition of caspases through the BIR containing region [58, 59]. Each BIR domain consists of the N-terminal "linker" which is unstructured followed by the main globular structure of four to five alpha helices and multiple antiparallel beta sheets which chelate a zinc ion. For IAPs containing multiple BIR domains, the second BIR domain (BIR2) is associated with caspase-3 and -7 inhibition and the third (BIR3) for the inhibition of caspase-9 [59, 60]. For single BIR members, the domain can either exhibit specificity similar to BIR2 or BIR3 by inhibiting both caspase-3 and -7, or only caspase-9 for Survivin and Ts-IAP respectively, or a combination of all three caspases as with Livin [61-63]. The linker region of BIR2 occupies the active site of caspase-3 and -7 in a reversed orientation relative to the caspase substrates or other inhibitors, and prevents catalytic activity by steric hindrance. The main BIR2 domain does not contribute to caspase-3 inhibition, but is required for the stabilisation of the linker in the caspase-7 active site [64-66]. Caspase-9, in contrast to the other initiator caspases, becomes catalytically active upon recruitment to the apoptosome, which induces a conformational change without requirement of proteolytic activation. In addition to the processing of caspase-3, the activated caspase-9 can also commence a self-cleavage event, generating a new amino terminus. The XIAP BIR3 binds to this newly generated region, preventing the caspase-9 from undergoing homodimerisation, hence keeping the caspase in an inactive state [67].

IAPs such as XIAP, c-IAP1 and c-IAP2 can also target IAPs and their interaction partners for proteasomal degradation through the RZF zinc-finger domain. These RZF possessing proteins act as adaptor proteins to recruit their targets to the E2 enzyme ubiquitination complex, the conjugation of the ubiquitin tag subsequently flags their targets for degradation by the proteasome. XIAP and c-IAP1 were demonstrated to undergo auto-ubiquitination and degradation upon apoptosis stimuli [68], while XIAP and c-IAP2 were capable of targeting caspases-3 and -7 for proteasomal degradation [69, 70]. SMAC/DIABLO released from the mitochondria during apoptosis also possesses an tetrapeptide IAP-binding motif, hence is able to competitively bind to IAPs like XIAP thereby lifting their inhibitory effects on the caspases through steric clashes [64, 71] and becoming ubiquinated in their place [72, 73].

Upon the activation of the caspase cascade, the IAPs can be cleaved by caspases. Various caspases such as caspase-3, -6, -7 and -8 are able to cleave XIAP while c-IAP1 can be cleaved by caspase-3. The XIAP cleavage fragment occurs in late stage apoptosis, hence is considered to be a phenotypic

marker of apoptosis rather than the abolishment of its inhibitory effect [57]. Cleavage of c-IAP1 is proposed to convert the anti-apoptotic effect into pro-apoptotic through the release of the C-terminal fragment containing the CARD and RZF domains [74]. Proteins such as XAF1, SMAC/DIABLO and Omi/HtrA2 are able to suppress the anti-apoptotic capacity of IAPs. XAF1 can bind XIAP directly to disrupt its inhibition of caspase-3 and the suppression of XAF1 is proposed as a mechanism of apoptosis suppression in transformed cells [75, 76]. SMAC/DIABLO and Omi/HtrA2 are thought to work in similar mechanism discussed above.

Hence, IAPs as such constitute a checkpoint in apoptosis progression after the release of cytochrome c, determined by the balance between IAPs and its suppressors such as XAF1 and SMAC/DIABlO. This allows for a reversal of cell fate prior to commitment to apoptosis through caspase activation and IAPs sequestration or degradation.

Apoptosome

The initial discovery of caspase-3, a prominent executioner caspase generally regarded as the "point of no return" during apoptosis, raised further questions regarding its mechanism of activation [77]. Procaspase-9, cytochrome c and Apaf-1 were subsequently determined as the molecules responsible for caspase-3 activation [78-80], forming an approximately 1 MDa complex central to mitochondrial-mediated apoptosis known as the apoptosome.

Apaf-1 is a member of the STAND (signal transduction ATPases with numerous domains) family present in the cytosol. Structurally, Apaf-1 comprises the nucleotide binding and oligomerisation domain (NOD) flanked by the CARD domain and WD40 repeat domain (WRD) on the N- and C-termius, respectively. The CARD domains within Apaf-1 and procaspase-9 allow for the interaction to occur while the NOD domain contains the adenine nucleotide binding site important for the oligomerisation of Apaf-1 into the apoptosome. The WRD domain prevents Apaf-1 oligomerisation by holding the NOD domain in place thus avoiding conformational changes necessary for apoptosome formation [81, 82].

Cytochrome c, with its already well defined function as part of the respiratory chain, was identified as the final critical component of apoptosome complex amid great scientific excitement for its dual functions [80]. The localisation of cytochrome c is of particular importance. Within the healthy cells, cytochrome c localised to the mitochondrial intermembrane space functioning as part of the respiratory chain. Upon apoptosis, mitochondrial damage allows the release of cytochrome c into the cytosol where it interacts and bind to Apaf-1, relieving the autoinhibition imposed by the WRD domain [83]. The separation of various components of the apoptosome within different organelles reduces the chance for unintentional activation of the caspase cascade.

The binding between Apaf-1 and cytochrome occurs in a 1:1 proportion and enables Apaf-1 to assume an extended conformation, forming a heptameric structure with seven-fold symmetry resemblance of a wheel termed the apoptosome. Recent modelling and cryo-electron microscopy studies indicate that the NOD domains of Apaf-1 form an inner ring surrounded by an outer ring comprising the WRD domains and cytochrome c. The Apaf-1 CARD domains form a disk resting above the inner ring [84, 85]. Once formed, the apoptosome complex is able to recruit procaspase-9 and bring about its activation by either the induced proximity or induced conformation model as discussed above. The activated caspase-9 then proceeds to activate caspase-3 committing the cell to irreversible apoptosis.

The exchange of an ADP molecule for an exogenous adenosine nucleotide is also important, however the precise type and function of adenosine nucleotide remain the subject of on-going research interest. ATP, dATP and other non-hydrolysable ATP analogs such as AppNHp and AppCp are able to interact with Apaf-1 leading to apoptosome formation, indicating that the ϒ-phosphate is adequate for complex formation [86].

The apoptosome complex is a major signalling hub in the intrinsic mitochondrial-mediated pathway where pro-apoptotic signals are integrated, activating the death complex.

Tumor Necrosis Factor (TNF) Superfamily

The TNF superfamily comprises an extensive group of cytokines and their associated receptors and is an area of immerse research interest. The family currently has 19 ligands and 30 receptors, each of which is involved in a wide range of cellular functions including cell proliferation, morphogenesis and apoptosis. First isolated from B lymphoblastoid cell line and myeloid cell line [87, 88], subsequent functional and structural analyses revealed significant homology between TNF-α and TNF-β setting the foundation for the large group of associated cytokines. Ligand members involved in the regulation of apoptosis includes TNF-α, FasL, VEGI, TRAIL and TWEAK, with their respectively receptors being TNFR1 & TNFR2, Fas, DR3, DR4 & DR5 and Fn 14. A single ligand may be able to bind to more than one receptors; TRAIL for example is capable of binding to DR4, DR5, DcR1, DcR2 and OPG. It is widely accepted that this ability to bind a range of different receptors serve to reduce the cytokine response, although the precise reason has yet to be completely understood.

Six receptor members possess the death domain (DD) within the cytoplasmic portion, with the DD acting to recruit adaptor proteins to create a platform for caspase activation. However, the DD is not indispensable, with receptors such as TNFR2 lacking a DD but remaining capable of mediating apoptosis [89].

Signalling via the TNF superfamily is the primary basis for the extrinsic apoptotic pathway. A prominent receptor which activates apoptosis through their associated adaptor proteins is TNFR1 (also known as DR2). Upon binding to TNF-α, TNF-α and TNFR1 form a trimeric complex which sequentially recruits TRADD, FADD and then the initiator procaspase-8, which in turns auto-activates and cleaves procaspase-3 into its active form initiating the apoptosis caspase cascade [90, 91]. This is known as the death-inducing signalling complex (DISC, see above). The induction of apoptosis signalling is one of three pathways which may be initiated following TNF-α ligand binding, the other two signalling pathways being activation of NF-κB which mediates transcription of proteins involved in cell survival and anti-apoptotic response, and activation of MAPK signalling involving activation of JNK stress related proteins and subsequent transcription factor activation such as c-Jun and ATF2, which the latter is responsible for cell differentiation, proliferation and pro-apoptotic responses. Other proteins associated with the TNF-α -TNFR1 complex include RIP, TRAF2 and ASK1 [92, 93]. Our group has recently discovered USP2a as the ubiquitin-specific protease that specifically keep in check the levels of RIP by targeting it for proteasomal degradation. TRAF2 in turn was found to be able to inhibit USP2a activity on K48 ubiquitin chains, and hence regulate the sensitivity of the cells to apoptosis [94].

While TNF- α was first discovered to be an anti-cancer agent, it has since become linked with a range of pathophysiological diseases from cancer to neurologic and pulmonary diseases, diabetes and obesity together with other TNF superfamily members. TNF- α-based therapeutics has been a great source of treatment for various diseases and continues to remain an active drug development target. For example, a TNF- α monoclonal antibody was approved for treatment of psoriatic arthritis, rheumatoid arthritis and active ankylosing spondylitis while receptor antagonists such as Atacicept which inhibits B cells activation by TACI are in clinical trials for autoimmune disease treatment [95, 96].

Bcl-2 Protein Family

The Bcl-2 family is a group of apoptosis regulators which governs mitochondrial changes such as the mitochondrial outer membrane permeabilisation (MOMP). Over 30 members have been discovered since the gene, bcl-2, was first found to promote survival of cytokine-dependent hematopoietic cells in the absence of the survival signals [97, 98]. Bcl-2 family members possess at least one conserved Bcl-2 homology (BH) domains. Functionally, the family can be classified into pro-survival and pro-apoptosis. Pro-survival members such as Bcl-2, Bcl-X_L. Bcl-w and Mcl-1 have the capacity for apoptosis inhibition over a wide range of cellular insults, and possess at least two BH domains with the most homologous members possessing all four BH domains [99]. The pro-apoptosis members

can be further classified into multidomain members such as Bax, Bak and Bok which have high structural homology to Bcl-2 proteins, and the single domains BH3-only proteins such as Bid, Bad, Bim, Noxa and Puma which only have the BH3 domain [100].

Within the apoptosis signalling network, the Bcl-2 family members exist to counteract the functions of other family members. The pro-survival members frequently bind the pro-apoptotic members, thereby hampering their ability to mediate release of cytochrome c. In the case of Bcl-2, it can bind to Bak through its exposed N-terminus while preventing Bax translocation to the mitochondria, thus prevents Bax/Bak oligomerisation and inhibits the release of cytochrome c through mitochondrial membrane permeabilisation [101, 102].

Current consensus on the BH-3 only proteins assumes that they can function through the "de-repressors" or "direct activators" models. De-repressor BH-3 only proteins exert their pro-apoptotic effect through the binding of the pro-survival members such as Bcl-2 at their hydrophobic pockets, hence relieving their apoptosis inhibitory effects [100, 103, 104]. The direct activators such as Bid and Bim are also able to induce apoptosis independent of Bcl-2 inhibition through interaction with Bax or direct damage to the mitochondria membrane [105]. For example, Bid is cleaved by active caspase-8 into tBid and undergoes myristoylation at the N-terminus, enabling it to interact with Bax or Bak which leads to Bax/Bak oligomerisation and subsequent apoptosis. The Bax/Bak complexes are hypothesised to exist transiently in a model known as "kiss and run" [106]. However, it is unlikely that activation of either de-repressors or direct activators is sufficient for apoptosis induction; the inhibition of Bcl-2 pro-survival proteins needs to occur together with Bid mediated activation of Bax and Bak [104]. tBid oligomers have also been shown to trigger apoptosis independent of the Bax/Bak complexes [107].

A number of hypothesises exist to explain how the Bcl-2 family proteins governs mitochondrial stability. The first relies on structural similarity between Bid and Bcl-X_L and observations of channel formation by tBid, Bax, Bcl-2 and Bcl-X_L. During apoptosis, translocation of Bcl-2 from the cytosol to the mitochondria and ER has been observed, and the subsequent pore formation is thought to protect or antagonise apoptosis [108, 109]. tBid has also been shown to behave similarly [110]. The second hypothesis suggests Bcl-2 family proteins could form a mitochondria pore termed the mitochondrial apoptosis-induced channel (MAC). While the composition of the channel complex remains to be determined, studies have indicated MAC to be involved in the release of cytochrome c. Bax has been proposed to be one of its components [111, 112]. The third model proposed that the curvature or composition of the mitochondrial lipid bilayer is altered by Bcl-2 family members. Studies have indicated tBid insertion into mitochondria to lead to rupture and overflow of

intermembrane space proteins into the cytoplasm, in a scenario similar to antibiotic polypeptide activity on bacteria [113].

In addition to the signalling via the mitochondria, Bcl-2 family members are also involved in other pathways that determine the fate of the cell. For example, Bcl-2 can enforce a pro-survival effect when expressed at the ER by maintaining calcium homeostasis, regulating caspase and Bax activation, while also inhibiting autophagy through interaction with Beclin1 [114-117].

As vital regulators in the nexus of apoptosis signalling, the activity of the Bcl-2 family proteins are tightly regulated at transcriptional, post-transcriptional and post-translational levels [118, 119]. Understanding how this balance between survival and apoptosis is governed can contribute in the mechanisms of disease disorders such as cancer and autoimmune diseases. Bcl-2 family proteins are the subject of therapeutic development, with several inhibitors of Bcl-2 like proteins in development pipelines as anticancer regimens [120, 121].

Heat Shock Proteins, ER Stress and Calcium Release associated Signalling

Heat shock proteins (HSPs) are produced in response to various stresses encountered by the cell. Under normal conditions, they are involved in various cellular functions including promoting proper folding of nascent polypeptide chains and multi-protein complex assembly, regulation of protein degradation through the proteasome, and protein translocation between various organelles. Their wide range of interaction targets confers a significant role in apoptosis by promoting the balance between pro-survival and pro-apoptosis signals [122-124].

HSP-induced protection is able to rescue from apoptosis, but this survival response is limited by the protein folding capacity which becomes depleted in the presence of significant stress scenarios. Hence HSPs act both as a defence mechanism and stress sensors, serving first as a temporal brake on apoptosis for cell repair subsequently shifting the equilibrium towards apoptosis should damage accumulation becomes critical [125, 126].

Within mammalian systems, four major classes of HSP, Hsp27, Hsp60, Hsp70 and Hsp90 exists to influent apoptosis across various signalling points [127]. Hsp27 has been shown to exert an anti-apoptotic effects by overexpression experiments [128], the mechanism proposed as its direct binding and sequestering of cytosolic cytochrome c from Apaf-1 [129]. Mitochondrial Hsp60 was reported to bind procaspase-3, while its cytosolic counterpart was found to be in complex with Bax. Apoptosis induction leads to dissociation of Hsp60 from its binding partners releasing procaspase-3 into the cytosol, while freeing Bax for apoptosis participation through translocation to the mitochondria. The chaperone function of Hsp60 can also accelerate caspase-3 activation [130, 131].

Hsp70 can influent apoptosis through various interactions. It can directly bind Apaf-1 in an ATP-dependent manner to inhibit apoptosome formation, indicative that its protein folding function is involved in the process [132]. As apoptosis progresses after caspase activation, Hsp70 continues to be able to counteract the process, by limiting the activation of cytosolic phospholipase A2, protection from caspase-3 and managing changes in nuclear morphology [133]. It can also bind and stabilise protein kinase C and Akt [134, 135], and supress JNK activation thereby inhibiting stress induced apoptosis and Bid-dependent apoptosis [136]. In late phase apoptosis where chromosomal fragmentation is effected by the DNase CAD, Hsp70, in conjunction with its co-chaperone Hsp40, is able to regulate the binding of iCAD (inhibitor of CAD) to CAD, hence exerts its control over DNA degradation [137]. Hsp90 can regulate TNF-receptor signalling through stabilisation of receptor interacting protein (RIP) allowing RIP interaction with TNFR-1, which triggers activation of NF-κB and JNK [138]. The presence of Hsp90 in the IKK complex is also required for effecting TNF-induced activation of IKK and NF-κB signalling [139].

While the extensive presence of HSPs across cell organelles allows them to initiate and regulate apoptosis from virtually all major cellular components [127], this section will focus on the ER for its crucial role in protein biosynthesis and maintenance of calcium at homeostatic levels. ER overloading and changes to its oxidative environment are among a range of stresses which triggers ER stress induced apoptosis [139]. This is achieved through two mechanisms, the unfolded protein response and alterations to calcium signalling. Caspase-7 and -12, and the transcription factor CHOP/GADD153 have been demonstrated to be involved in this process [140, 141]. Grp78, induced by signals such as ER calcium depletion, is able to inhibit apoptosis through direct binding of caspase-7 and -12 [142]. Grp78 is also able to activate NF-κB signalling [143]. An increasingly exciting field is the interface between the ER and mitochondria, the central apoptosis signalling hub. This involves junctions between the two organelles through which it is hypothesised that pro- and anti-apoptotic molecules could be exchanged. Bcl-2 has been shown to regulate calcium fluxes via such junctions [144] while pro-apoptotic Bax and Bak were demonstrated to localise to the ER to promote caspase-12 associated apoptosis [145]. Furthermore, our group has discovered a novel platform spanning the mitochondria-ER interface, termed the ARCosome. This comprises the Fis1-Bap31 complex upon which procaspase-8 becomes activated [146].

Calcium release from the ER is an important apoptotic signal alongside cytochrome c release from the mitochondria. Cytosolic calcium is highly toxic to the cells, resulting in immense activation of phospholipases and proteases, hence calcium is largely imported and stored within the ER [147]. Calcium as a chemical messenger coordinating the mitochondria-ER activation of apoptosis was

discovered by Boehning *et al*. During apoptosis, a small amount of cytochrome c was initially released from damaged mitochondria. These became bound to inositol-1,4,5-trisphosphate ($InsP_3$) receptors on the ER membrane, allowing massive calcium export. The large amount of calcium released becomes loaded into mitochondria, causing further release of cytochrome c and other pro-apoptotic proteins [148]. The global increase in cytosolic calcium also activates proteases and phospholipases and is required for nucleases activity, aiding in the destruction and packaging of the cells during apoptosis.

p53 and transcription dependent apoptosis

p53 is a master transcription regulator involved in a wide range of functions from cell cycle, senescence, DNA metabolism, differentiation to apoptosis. It is a critical gene in maintaining proper cellular activity, with more than half of cancer cells containing mutations which inactive p53 [149].

It acts as a switch primarily to determine the cell fate between cell cycle inhibition and apoptosis induction in response to stresses such as DNA damage. p53 can initiate transcription of its target genes by binding their consensus promoter sequences [150], or via non-canonical sequences such as in the case of the p53-induced gene 3 (*PIG3*), which occurs through microsatellite sequence of its untranslated region [151]. A key contributing factor to p53's ability to selectively bind promoters lies in the redox sensing capacity of its DNA binding region. Under oxidative environment, Cys277 of the human p53 is oxidised, reducing its binding affinity to the promoter of *Gadd45*, a DNA repair associated gene, while maintaining its affinity for $p21^{WAF/CIP1}$ (*p21*), a cell cycle regulatory gene. The oxidation of Cys277 is likely to be due to oxygen radicals resulting from high UV treatment [152]. The converse is true where *Gadd45* promoter affinity is increased under reducing conditions, with Cys277 reduction allowing for binding to the cytosine rich region of its promoter by p53. Interaction with transcriptional co-activators can also augment p53 promoter binding affinity. Myc, which prevents growth arrest resulting from irradiation generated DNA damage, can specifically bind to the promoter of *p21* upon irradiation, repressing its activation by p53 [153]. In contrast, apoptosis stimulating proteins of p53 (ASPP) promotes apoptosis and prevents cycle progression by enhancing p53 binding to promoters of pro-apoptotic genes such as *Bax* and *PIG3*, while lowering affinity for *p21* [154]. Such interactions with co-activators and repressors are often cell type specific.

p53 regulates apoptosis through both the extrinsic and intrinsic pathways. It can induce transcription of surface transmembrane receptors such as Fas (the receptor for FasL) [155], DR5 (receptor for TRAIL) [156] and PERP [157]. Fas and DR5 leads to DISC complex formation and caspase-8 activation, the classical extrinsic pathway. PERP is a novel transmembrane member of the PMP-22/gas family associated with regulation of cell growth, and works with E2F-1 for apoptosis induction [157]. Within

the intrinsic pathway, p53 promotes the expression of BH-3 only members of the Bcl-2 family such as Bax, Bid, Noxa and PUMA [158-162], and caspase activation such as caspase-8 and caspase-6 [163-165]. Localisation of p53 to the mitochondria has also proved to be extremely apoptogenic to the cells [166]. It has been found that p53 can directly complex with Bcl-2 and Bcl-X_L leading to permeabilisation of the mitochondrial outer membrane in a transcription independent manner [167, 168].

In addition to its direct pro-apoptotic effects, p53 can also abrogate survival signals in response to excessive stresses. Survival signals such as binding of mitogens and cytokines to their receptors are primarily transduced within the cell by phosphoinositide 3-kinase (PI3K). PI3K is able to activate AKT, a serine/threonine kinase, leading to downstream phosphorylation of a wider target range including Mdm2 (one of the major p53 inhibitors) in favour of cell survival and proliferation [169, 170]. As a result, p53 is destabilised by Mdm2 through the activity of AKT. This signalling route is neutralised under stress induced activation of p53, where p53 promotes caspase activation which subsequently degrades AKT [171, 172]. Furthermore, p53 is able to upregulate *PTEN*, a tumour suppressor gene, and *cyclin G*. *PTEN* encodes a phosphatase protein product that then proceeds to de-phosphorylate PI3K while cyclin G recruits PP2AB' (also a phosphatase) to de-phosphorylate Mdm2 [169, 171]. By counteracting its inhibition on all three major signalling nodes, p53 is able to shift the cellular fate from pro-survival to pro-apoptosis.

Dependence Receptors

A distinct class of receptors, known as dependence receptors, has emerged as part of the apoptosis signalling network. In contrast with pro-survival receptors such as NGF receptors which promotes cell survival upon ligand binding and death receptors like TNFR1 which triggers apoptosis, dependence receptors is capable of a combination of both functions. When a ligand is bound, dependence receptors promotes cell survival, proliferation and differentiation signals, while the absence of the ligand leads to apoptosis induction [173].

Identified members of dependence receptors class include neurotrophin receptor $P75^{ntr}$ [174], DCC receptors [175], UNC5H [176], androgen receptor (AR) [177], Patched [178], neogenin [179] and integrins like $a_v\beta_3$ and $a_5\beta_1$ [180]. There is no structural and functional similarity between the various dependence receptors, but general consensus holds that dependence receptors are likely to possess the dependence receptor transmembrane (DART) motif [181] and caspase substrate recognition sequences within their cytoplasmic regions.

The precise mechanism for the activation of apoptosis by dependence receptors is the subject of current research, but much evidence have indicated the involvement of caspase-3 [182], -7 and -8

[178] as the protease that release a pro-apoptotic domain. How caspases become activated by an unbound receptor is unknown as they lack the death domains which recruit caspase adaptor proteins but a model similar to induced proximity as in the case of the death receptors has been proposed. In dependence receptor induced apoptosis, the common theme is for caspase-3 to mediate one or two cleavage within the cytoplasmic region of the receptors in the absence of their ligands, which either release a pro-apoptotic fragment (UNC5H) or expose a receptor region (DCC) which then propagate the apoptotic signal [175, 176, 178, 183]. It also does not require Apaf-1 or cytochrome c which are well established players of the intrinsic apoptosis pathway, hence it is proposed that dependence receptor induced apoptosis via a distinct mechanism.

Our group has previously isolated CD82 (Kai1), a known metastasis suppressor, as an apoptosis inducer and have demonstrated its activity as a dependence receptor.

Clinical Implications

As the primary mechanism for controlled cell elimination, apoptosis is fundamental to normal tissue and cellular function. It is involved in a diverse range of processes such as tissue homeostasis, organogenesis, maturation of the immune system, and elimination of compromised cells such as virus-infected or cancer cells. Apoptosis is managed through a delicate balance of pro- and anti-apoptotic signals, and when disrupted would give rise to various diseases such as Alzheimer's, Parkinson's [184], autoimmune diseases, cardiovascular diseases such as myocardial ischemia and hypoxia [185]. Apoptosis signalling may also be abused by viruses and cancer cells for proliferation means and to evade cell death resulting from their invasion. This acquired ability to evade apoptosis is famously included as one of the hallmarks of cancer [46].

In Alzheimer's disease, cortical neurones undergo apoptosis as a result of β-amyloid accumulation within the brain [186], while in Parkinson's disease excessive dopamine generated from dopaminergic neurons [184]. Within the immune system, apoptosis removes immature B cells producing antibodies against the host proteins, and is employed by the cytotoxic T cells to terminate their target cells among other functions [187, 188]. Failure to remove the defunct B cells would preserve the anti-self antibodies which inappropriately tag the host cells as foreign causing the immune system to attack the host, while excessive activity by the cytotoxic T cells may lead to apoptosis in normal host cells. Dysfunctional apoptosis within the immune system has indeed been linked to rheumatoid arthritis, inflammatory bowel disease and diabetes [189, 190]. Excessive apoptosis though various signalling nodes such as TNF-α and caspase-1 have also been associated with cardiovascular diseases [185].

While the above-mentioned diseases are caused by up-regulation of pro-apoptosis signals, viruses often inhibit apoptosis after infection to extend the life of their host for viral replication to occur. Viruses like baculoviruses encode IAPs which are able to inhibit caspase activity while cowpox serpin crmA is also able to inhibit apoptosis, leading to an increase tendency of cancer formation [191]. Adenovirus and papiloma virus instead employ the use of a p53 inhibitor [192]. Interestingly, the human immunodeficiency virus (HIV) triggers activation of apoptosis in CD4+ cells, while producing HIV-1 transactivating protein (Tat) in infected cells. Tat is then taken up by other uninfected T cells, sensitising them to apoptosis through lowering of intracellular antioxidant levels [193, 194].

Apoptosis inhibition is critical to cancer formation, and cancer cells have acquired an arsenal of mechanisms at all major signalling pathways through accumulated spontaneous mutations and virus infections to shift the balance towards survival and proliferation. The master regulator p53 is one of the most frequently and heavily mutated genes in cancers, with 26,000 known mutations [195] capable of conferring apoptosis inhibitory function on the cells. Another common strategy is the up-regulation of anti-apoptotic proteins in cancer cells, such as c-FLIP which can bind to the adaptor protein FADD preventing caspase activation via the extrinsic pathway, and Bcl-2 which blocks cytochrome c release central to the intrinsic mitochondrial pathways [119, 196]. Decoy receptors may also be up-regulated by up to 50% to diminish the effects of death receptors signalling from the immune cells [197]. Other anti-apoptotic proteins up-regulated includes IAPs and surviving, while pro-apoptotic proteins such as caspases are down-regulated [198]. The net effect is the dampening of pro-apoptotic signalling in favour of anti-apoptotic factors and activation of survival and proliferation pathways.

The vast apoptosis signalling network also offers a myriad of opportunities in development of drugs and therapeutic strategies against these diseases and cancers. Antagonists, inhibitors and antibodies against the major signal regulators such as Bcl-2, TNF receptors, HSPs, caspases and p53 are in development for treatment. For example, soluble TNF receptors have been used into the treatment of Cohn's disease by acting as a sink to bind the excess TNF that is triggering apoptosis, while peptides derived from p53 have been employed in cancer treatment [90, 127, 199]. Other strategies include the use of gene therapy or small molecules to correct for the defects or mutations, such as in the case of p53 where the small molecule PhiKan083 was shown to be able to reactivate mutant p53 and restore its function [200]. The use of cytokines, ligands and their analogues to trick cancer cells into suicide is also a promising area. TRAIL was also shown to be able to bind to both the death receptors DR4 and DR4 to form a complex leading to the destruction of tumour cells [201] while TNF-α is used with other chemotherapy with success [202].

Nanoparticles

Definition, Characteristics and Applications

Nanotechnology is a rapidly evolving field, and in the past decades efforts have become focused on expanding the development and applications of nanomaterials. Nanomaterials possess unique properties conferred by their size and are used in most industries today ranging from drug delivery and therapeutic agents to textiles, food and cosmetics. While nanomaterials are often well characterised in terms of chemistry and applications, the toxicity assessment of such materials is often lacking. As these materials become commonly used, people would be increasingly exposed either through direct contact in the workplace, through airborne particles or materials intended for health or therapeutic use. Hence, critical toxicity assessments are required to balance their potential applications and ensure the safety of the consumers and workplaces.

A precise definition of nanoparticles is currently lacking and dependent on its context of use. A strict definition dictates that nanoparticles have dimensions between 1 – 100 nm, within which they can exhibit novel properties resulting from quantum effects. Their properties become similar to atomic and molecular interactions below 1 nm, while approximately above 100 nm they are thought to behave as with bulk properties of the materials [203, 204]. However, the upper size limit is arbitrary and is extendable up to 1000 nm.

Characteristics such as particle size, size distribution, aspect ratios, shape, surface, catalytic activity, core chemistry, redox potential and porosity are important for interpretation of experimental data and risk assessments [205]. These characteristics are often affected by experimental conditions. For example, the presence of serum proteins in cell culture media is known to aggregate nanoparticles [206].

Each nanoparticle possesses a unique combination of these characteristics, which form the basis of their industrial applications. Aluminium nanoparticles are widely used as energetics, coatings & alloys, incendiary devices and sensors. The use of aluminium and other nanoscale metals together with metal oxides for energetics is based on their ability to release large amount of energy on oxidation [207]. Some metallic nanoparticles and their oxides such as silver and zinc are well established antimicrobial agents capable of killing or slowing the growth of bacterial without toxicity to surrounding cells and tissues [208]. Sulphur nanoparticles have also been recently discovered to possess efficient antimicrobial activity [209]. Nanoparticles are also increasingly marketed as key and active ingredients in cosmetics offering improved UV protection, “long-lasting” effects, increased colour and quality finish by leading manufacturers. Mineral-based nanoparticles such as titanium

dioxide and zinc oxide are some of the key ingredients in sunscreen [210]. Carbon nanotubes are commonly used nanoparticles in production of electronic, sports, aircraft and textile products due to their strength and lightness. Nanoparticles are also finding extensive applications in biomedical and biotechnology. Carbon nanostructures have recently been demonstrated as scaffolding structures in the repair of large gaps in damaged nerves [211]. Novel formulation of nanomaterials are exploited as drug delivery vehicles as nanocarriers and nanocapsules where the drug molecules are encapsulated. Dextran-camptothecin nanoparticles have thus far been shown to be having the longest circulation half-life of above 300 hours in humans with little toxicity observed. Nano-liposomes are another form of non-viral drug delivery systems [212]. These nanocarriers can have their cancer recognition affinity improved further through functionalization with cancer-specific antibodies and other biomolecules [213]. Such active tumour targeting and delivery of drugs could significantly reduce the side effects associated with traditional chemotherapy.

Silica nanoparticles are among the most widely used particles due to the ease of preparation and inexpensive production. They can be easily functionalised to provide good drug separation while their porosity characteristics make them good drug delivery systems. They have been used to stabilised aspirin tablets [214] and often used as pharmaceutical excipients which are inactive substance used as a carrier for delivery of active ingredients [215]. The established DNA binding ability of silica also makes them suitable candidates for vector delivery in gene therapy [216]. Colloidal anhydrous silicon dioxide is commonly used as a non-toxic excipient for oral and topical delivery of pharmaceutical products, but injection-based delivery may trigger local tissue reactions and toxicity [217].

Nanoparticles: Signalling and Apoptosis

The diversity of nanoparticles is likely to induce different signalling and toxicity pathways specific to the characteristics of each nanoparticle type. Such characteristics, along with its surface, method of dispersions and experimental preparation determines the type of biomolecules especially proteins that binds to the nanoparticles upon their entry into the human system. The biological molecules form the protein corona, which has an essential role on cell surface receptors and lipids interactions with the nanoparticles, thereby determining its entry into the cells [218].

The increased surface area of nanoparticles is another important indicator of toxicity, as non-toxic materials of the micrometre range are often capable of inducing a toxic response in cells and organs exposed to similar materials in the nanometre range [219]. Insoluble nanoparticles like titanium dioxide and carbon black have been shown to trigger pulmonary inflammation or intratracheal instillation in animal models after inhalation [220].

Other *in vivo* studies have indicated that carbon nanotubes behave similarly to fibrous asbestos as a result of their high aspect ratios (length to width ratio), inducing pleural granuloma and mesothelioma as a result of its shape rather than chemistry [221]. Nano-silver is thought to trigger toxicity as a result of free radical formation and the resulting oxidative stress [208]. High concentrations of nanoparticles are associated with necrosis [222] while lower doses induce apoptosis. Toxicity effects from nanoparticles vary markedly. Aluminium nanoparticles (40 – 47 nm diameters) were found to induce toxicity through apoptosis in Neuro-2A cells by affecting the mitochondrial integrity [223]. Similar nano-aluminium (with diameters between 8 – 12 nm) inhibits apoptosis and triggers proliferation of mouse epithelial JB6 cells [224]. This inhibitory effect involves regulations of AP-1 and increased Bcl-X_L expression through the increased expression of SIRT-1.

While studies are increasingly demonstrating apoptosis induction upon treatment with nanoparticles, the molecular signalling remains to be largely unexplained. Nanoparticles induced apoptosis are hypothesised to be the consequence of various signalling points from the death receptors to mitochondria and lysosomes. Silica nanoparticles-induced apoptosis was shown to involve lysosomal permeabilisation while the JNK/P38 pathway and subsequent Bid and caspase-8 activation were associated with titanium dioxide nanoparticle apoptosis in lymphocytes [225, 226]. The JNK pathway apoptosis was instead inhibited by nano-fullerene in cerebral microvasculature endothelial cells [227, 228]. Other apoptosis related proteins that were discovered to be changed through expression and proteomics studies includes cyclin dependent kinase inhibitor 1A (CDNK1), alpha synuclein, Bcl-2 modifying factor, NF-κB, transitional ER ATPase, VDAC-1, MAPK/ERK kinase kinase 1, DAG kinase, FADD protein, chemokine (C-X-C motif) ligand 1, HSP70 2, IL7 and MDR/TAP, all of which were changed in a pro-apoptotic manner [229].

Pro-survival pathways were also triggered as a response to nanoparticles presence. In RLE-6TN rat lung epithelial cells treated with carbon black nanoparticles, an ERK1/2 signalling mechanism involving EGFR and β1-integrin activation of PI3K and Akt cascade was shown to induce apoptosis [230]. NF-κB signalling was also activated in A549 exposed to water-soluble nanoparticles generated from propane combustion [231].

High Throughput Technology

High Throughput Screening Concept

The high throughput concept took hold in the early 1990s as an approach for lead compound discovery among pharmaceutical and biotechnological companies. This approach focuses on

processing thousands of samples per day through lab automation and robotics to identify candidates for the drug discovery pipeline.

Traditionally, high throughput screening (HTS) is performed in 96-well microplates, generating 10,000 to 100,000 data points daily. Such preference has shifted towards further miniaturisation and higher plate densities in recent years, in favour of the 384-well and 1536-well plate formats [232]. Miniaturisation increases the daily throughput in excess of 200,000 samples (termed ultra-high throughput screening or uHTS) and addresses the key issue of costs and time. Cost is a major inhibitory factor in adoption of HTS technology both in terms of initial hardware investment such as robots and the consumables costs down the pipeline.

Increasingly efficient systems and miniaturisation has made the technology affordable in academic research, leading to widespread popularity of high throughput techniques in the past five to six years. In a simplistic sense, HTS involves systematically screening a collection of chemical or biological molecules, known as the library, for candidate effectors of a phenotype-of-interest. The approach is often divided into various key phases: Assay development & validation, Library generation, Process optimisation, Screen Implementation, Data Capture & Analysis and Secondary Validation & lead generation [233, 234].

HTS has evolved into a mature discipline since its early implementation as part of the pharmaceutical lead compound discovery process. Efficient automation and miniaturisation combined with commercial availability of large chemical and gene libraries and high quality read-outs have established HTS as an indispensable tool for both commercial and academic research. While past efforts have focused heavily on synthetic and natural small molecules chemical compounds, the future trend will become directed at genes as novel targets such as transporters, ion channels and other protein interaction based molecules [235].

A screen is different from a selection in which any cell not displaying the phenotype-of-interest, termed background, is removed using some form of selective pressure such as chemical resistance. A screen identifies the target cells in a background of cells [233]. This is an important distinction in apoptosis research since the phenotype-of-interest, apoptosis, removes the cells, making it difficult to select for apoptotic genes.

Libraries and Assays

The libraries, or collections of small molecules- or nucleic acid- based compounds, are often exclusively generated by each research group in the early days of HTS. The size of the collection could vary widely, from hundreds strong collections at academic institutions to collections with

millions of compounds at pharmaceuticals companies developed over the years [234]. Depending on the purpose of the screen, the library would either comprise of small molecules as in the case of lead discovery or, RNAi or cDNA plasmids as part of a functional genomics screen. The compound libraries have become expanded to include a myriad of molecules derived from plant or other natural sources, complementing traditional sources such as combinatorial chemistry and automated synthesis [236]. The Human Genomic Project has provided large collections of cDNA libraries which are sequenced and defined [237], while RNAi libraries are often designed and synthesised by commercial companies. In recent years, the increasing adoption of HTS has led chemical and biotech companies to develop such libraries for commercial purposes, making this key component of HTS widely and readily available for research and development.

HTS assays for a primary screen are sensitive read-outs critical in hit discovery and minimisation of false positives. They are less quantitative and often used only in duplicates, although many companies are increasingly using single measurements to reduce costs. The generally accepted approach is to use a fast and sensitive primary assay to generate a list of positives or “hits” which are then subjected to more rigorous and accurate secondary assays. A variety of assays types is employed in HTS, including proliferation/cytotoxicity assays, ELISA, reporter enzyme based assays and binding assays, all of which are adapted into automated robotic formats [234].

Assays could be classified based on the involvement of cells. Cell-free assays, on the one hand, assess the interaction of proteins with other peptides or nucleic acids directly in the assay buffers in the absence of cells. This could be employed in screens for antibodies against a particular antigen in ELISA-style assays, to study protein-protein interactions using fluorescence resonance energy transfer (FRET) based assays [238], or even identification of protein inhibitors [239].

Cell-based assays, on the contrary, use various cell types as the primary system to assess drug and gene expression effects. Examples include the Sulforhodamine B (SRB) assay to quantify cytotoxicity [240], chloramphenicol acetyl transferase (CAT) assays for study promoter activity [241], and other enzyme and fluorescent protein based assays involving luciferase and GFP.

The quality of the assay is intricately linked to the quality of the hits, and an extensive library improves the chances of hit generation. While cellular systems are easy to work with, it may not be an accurate representation of the human system where the drug ultimately works. As the complexity of screens develop beyond cell-based assays, researchers are increasingly using higher organisms like zebrafish as the host systems for their physiological and morphological similarities to

mammals [242]. Such phenotypic screens have yielded compounds such as Digoxin, Isoniazid and even Sildenafil for the treatment of a wide range of diseases [242].

Types of screens

Compound screens were the basis of HTS in the lead discovery process. Such screens involve individual treatment of cells with small molecules from the library and assaying for cytotoxicity or other phenotype-of-interest. They are easy to deploy and useful for lead discovery but many notable issues remain such as compound solubility, the active concentration and physiological relevance.

As the screen design evolves, high throughput techniques began to be combined with DNA sequencing. These high-throughput sequencing approaches in combination with immunoprecipitation have been used in the study of DNA methylation and promoter expression alteration in various breast cancer cell lines [243]. Deep sequencing (where the region is sequenced for multiple fold coverage) has been employed in many systematic screens such as the Cancer Genome Project, with studies revealing the extent of mutations in major oncogenes and regulators. Such screens have revealed frequent mutations to genes like PREX2, MAP2K1 and MAP2K2 in melanomas [244, 245], SPOP, FOXA1 and MED12 in prostate cancer [246] and components of the ubiquitin-mediated proteolysis pathways in renal cancer [247]. These screens sequence only the target regions. Whole genome sequencing is more common in yeast and bacteria with smaller genome size, along with some forms of directed evolution [248, 249].

HTS has further been adapted for functional annotation of genes. These screens, termed functional genetic screens, transiently introduce cDNA expression plasmids or RNAi into mammalian cell lines, after which the phenotype-of-interest such as apoptosis is assayed for.

This introduction of genetic materials can be performed by co-precipitation of the plasmids with calcium phosphate or polyethyleneimines which are then taken up by the cells via endocytosis [250], with lipid-based transfection methods forming DNA-liposomes that can fuse with the cell membrane [251], or by packaging and transduction with viruses. Recently, paramagnetic nanoparticles are also increasingly being exploited in transfection of expression plasmids which claims to achieve higher levels of transfection and low toxicity [252].

cDNA libraries often result in a gain-of-function screen where the phenotype-of-interest is conferred upon the cells by ectopic expression while RNAi libraries lead to loss-of-function screens by suppressing the expression of the target genes. By systematically over-expressing or knocking down each gene, knowledge about its functional activity could be derived [253-255]. Indeed, various functional screens were successfully designed and implemented, identifying novel regulators of

cellular functions like telomerase activity, acetylation-deacetylation of histones, genome maintenance, oxidative stress response and cancer development [256-259]. Synthetic lethality is a common adaption of functional screens. It refers to redundancy within a cellular signalling network, where two genes are described as synthetically lethal if mutation of both genes results in cell death while mutation of the individual genes do not. Exploiting the concept of synthetic lethality with a cancer-associated mutation could provide a cancer-specific framework for drug development [260]. Such adaptations involve the use of mutation or ineffective concentration of drugs in combination with individually transfected cDNA or RNAi to determine which genes, genomic loci, proteins or other conditions then sensitise the cells to cell death. Synthetic lethal screens are particularly important in cancer research as a tool to study the anti-apoptosis ability of cancer cells and identify conditions capable of sensitising these cells to chemotherapy induced toxicity [255].

While cell-based "*in vitro*" functional screens are effective at dissecting cellular signalling pathways, it is impossible to study complex mechanisms and diseases at higher levels of organisations such as tissues and organs. Hence, as mentioned above, HTS was adapted to the use of animal models and has been demonstrated in amphibians like *Xenopus tropicalis* [261], invertebrate *Ciona intestinalis* with high genomic similarity with humans [262] and Zebrafish [242]. For example, a screen for suppressors of crb mutant was performed in Zebrafish. crb mutants give rise to more mitotic cells with the heterozygotes having increased cancer risks. The screen was achieved at a rate of 1,000 compounds per week, completing a 16,320 compounds library within 16 weeks, identifying suitable compounds as anti-cancer agents [263]. Transgenic Zebrafish were also established as excellent cancer models for melanoma and leukaemia [264-266] and the study of the nervous system [267].

Apoptosis Related Screens

Early screens were dependent on internally generated libraries and often labour intensive due to the prohibitive initial investment required. Albayrak *et al.* implemented a manual screen using a cDNA library generated from mouse mRNA, leading to the identification of 72 apoptosis inducing clones. The screen was performed in human embryonic kidney (HEK) 293T cells and in excess of 10,000 cDNA clones were screened all of which were visually inspected for typical apoptotic phenotypes like membrane blebbing. The candidate clones were later subjected to quantitative assays by flow cytometry counting of propidium iodide positive staining [268].

A similar screen for apoptosis inducers was later performed by Park *et al.*, processing a small library of 938 cDNA using automated fluorescence microscopy to image transfected cells nuclei stained with Hoechst 33342. The nuclear images were later analysed using computer algorithm yielding three positive clones [269]. Another study by Wan *et al.* employed commercially available high-throughput

systems and transfection reagents to screen 29,910 clones from an in-house cDNA library was achieved at high financial costs. The study focused on cancer development by detecting a change in cell colony counts after G418 selection, identifying 8,237 positive clones with 49 associated with apoptosis based on bioinformatics analysis [254].

Reverse transfection cell-based microarray [270] is an alternative to rapidly screen large libraries cost efficiently. This is achieved through miniaturisation where the gene libraries are printed onto slides in the presence of transfection reagents as it is used with standard microarrays. Cells are then added, which grow over and take up the DNA mix, expressing the genetic material. Such screens by Palmer *et al.* [271] and Mannherz *et al.* (2006) [272] identified 3/1959 and 10/382 clones positive as apoptosis inducers, respectively.

Compound screens were also commonly exploited to detect the pro-apoptotic activity of a compound since apoptosis is tightly associated with cancer development. A compound screen by Johnson *et al.* screened over 9,000 compounds in prostate cancer 3 (PC-3) cells with multiplexed PCR primers for toxicity markers such as Hsp70, Gadd153 and O6-methylguanine-DNA methyltransferase, and survivin as marker for apoptosis inhibition. Compounds that specifically reduce the mRNA of survivin and other prostate cancer markers but not the toxicity markers are anticancer drug candidates [273]. Another study by Tian *et al.* combines a library of compounds isolated from herbal medicine with an innovative FRET-based assay to identify anticancer compounds. The reporter, YFP fused with CFP through a caspase-3 cleavage sequence, was stably expressed in HeLa cells and generated the FRET effect due to close proximity of the YFP and CFP. When a compound induced apoptosis, capase-3 becomes activated and cleaved the reporter causing the loss of the FRET signal. This screen identified compounds from the tanshinone and podophyllotoxin family as candidates for drug development [274]. Compound screens were also used to identify novel caspase-3 inhibitors [239] and agents that specifically kill cancer stem cells [275].

Screening technology was also taken *in vivo* to the Drosophila model, with a library of 13,071 RNAi targeting the whole genome. This study identifies 42 positive apoptosis inducing candidates including a novel regulator Tango7. Tango7 silencing was able to inhibit caspase-dependent apoptosis, a function also presence in its human homologue PCID1. PCID1 functions through caspase-9 and offers a new development in the regulation via the apoptosome [276].

Automation and Robots

Laboratory automation has been around since the early 1980s when robots were used for simple repetitive tasks like setting up sequencing or PCR reactions and colony picking of bacterial clones from agar plates. Automation technology was critical to the Human Genome Project and its rapid

development contributed to the early completion by almost a decade. Large-scale public genome projects and private drug discovery pipelines continue to drive innovative applications of automation in research, applying rigorous productivity and consistency to research data generation [277].

Frequently cited advantages include the elimination of laborious tasks, improvement in time efficiency with overnight automation, quality and safety of work, and cost-efficiency [277, 278]. Return on investment reminiscent of corporate accounting culture has been calculated to compare the initial costs to the savings and workload generated by robots, and on average, it is suggested that the cost is recovered within two years of initial purchase [278]. Technology has progressed to the stage where even the most challenging or delicate task such as generating protein crystals and cell cultures could be automated, but such advances represent a trade-off between achieving insights into ever increasing, statistically viable datasets and the loss of techniques by the researcher [279].

Beyond the allure of automation remain many issues that are known to researchers but less often discussed. The cost is a major inhibitory factor. Robots are expensive to purchase and operate. While costs may be justifiable, many research groups simply lack the funding to make the initial purchase. Commercial efforts have been made for cost reduction, such as making the robots with their liquid handling platforms more modular, allowing the researcher to add only the function required. Specialised workstations are also available to automate specific processes, such as the Qiagen BioRobot series, which is designed for high-throughput genomics and protein purification for structural analysis. Also, operational costs could be reduced by using less reagents with newer liquid handling platforms capable of dispensing nanolitres using electromagnetic valves, and miniaturisation of assays that support “lab-on-chip” technologies [279].

After the investment, the setup process offers a whole new series of challenges. Adapting a standard protocol to an automated platform is difficult given the required optimisation of liquid handling and assay reproducibility. This could take months before the robot produces a stable run, sometimes even longer for it to achieve its full potential. In addition, fast robots are not sufficient; bottlenecks in the entire process need to be identified and optimised to ensure the highest efficiency in the workflow [280].

Normal laboratory procedures have to be scaled up. Simple issues such as freezer and storage space, and reagent ordering become complex. A moderate library could easily fill up an entire freezer and the sheer volume of reagents required accumulates from microlitres into hundreds of litres. Information and data management have to be integrated into the workflow, where data could

amount to terabytes especially with image acquisitions. Manual analysis is no longer viable, requiring specialised algorithms to identify patterns and the lack of such software easily adds to delays. Thermal gradient across high-throughput plates has been shown to be more significant than in individual tubes [281], and these require proper equilibration to minimise variations between wells [282].

Hence, high-throughput systems, especially those that involve compound screens, are often cited to be capable of screening hundreds of thousands of compounds per day, but achieving this in reality requires sustained funding for consumables and determination. Automation technology forms the core lead discovery pipeline of major pharmaceutical companies today, most of which would find it impossible to operate in the absence of such technologies [279].

Genetic screens involving different techniques are being used to elucidate the apoptosis signalling network, but the full high-throughput potential is often reduced by library redundancy, non-specificity of the assay, delays in automation workflow and the costs among other factors.

Chapter 2: Project Information

Project Overview

The involvement of apoptosis in various disease states and carcinogenesis is well established. Expanding our understanding of the signalling network through high-throughput screening would help assign gene functions to the sequences produced by the various genome projects, revealing candidates that could be targets for anticancer drug development or even gene therapy.

Current screening efforts have identified many new open reading frames as pro-apoptotic gene candidates such as C10orf61, MGC 26717, and FLJ13855 [269] but high library redundancy remains a key issue. Previous screens by our group have identified over a hundred clones as inducers of apoptosis including Orctl3 which is tumour specific [268, 283].

The popularity of nanomaterials and their expanding application in consumer products and production technology requires toxicity assessments and regulations to ensure any potential risks is managed. While nanomaterials are known to cause necrosis at high doses, a variety of nanomaterials have become implicated with toxicity via apoptosis with each able to induce apoptosis via a different mode of action conferred by their unique characteristics. By integrating the treatment of nanoparticles with our high-throughput screening platform, we aim to elucidate the toxicological properties of these nanomaterials through pathway analysis.

This project attempts build on the success of previous screens [268, 284] through the use of a fully sequenced cDNA library instead of previously generated in-house libraries, with the expectation of substantially lower redundancy and hence improved hit quality and counts. Furthermore, this project also attempts to identify positive and negative regulators of apoptosis, as opposed to only apoptosis inducers as in our previous screens.

The project will utilise the RISCI (robotic single cDNA investigation) screen (Figure 1) that employs a specific method for extraction of high purity DNA plasmids with reduced endotoxin levels in combination with an unique beta-galactosidase based assay for apoptosis detection designed for high-throughput screens [285]. This assay exploits the membrane permeablisation in late apoptosis to allow the interaction of the expressed beta-galactosidase with the substrate in the cell to generate a quantifiable signal. DNA extraction and transfection are performed by a pair of custom-built robots, which are capable of 2,000 individual reactions per day.

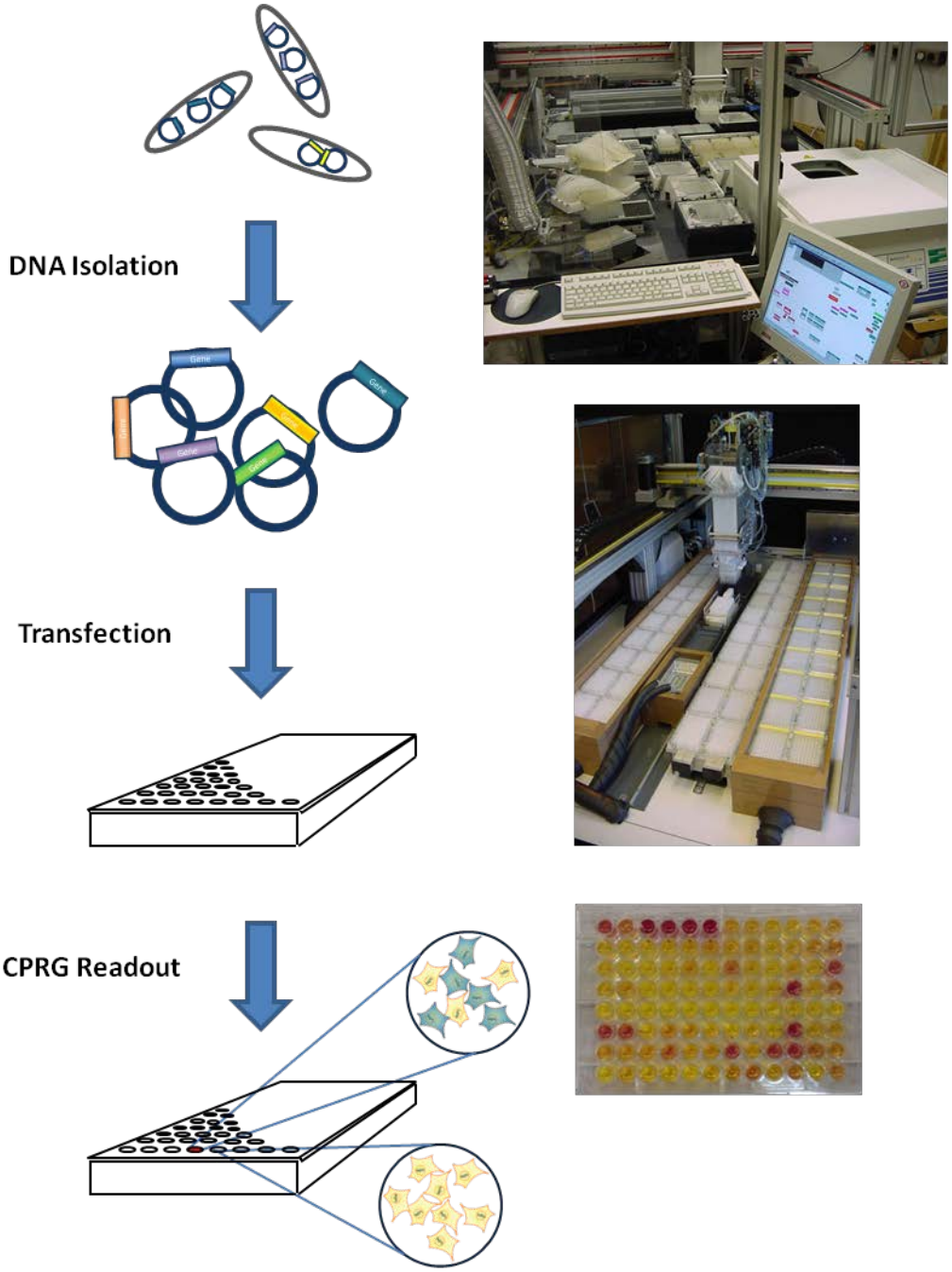


Figure 1: Workflow of the RISCI Screen. RISCI is a high-through screen with a pair of robots automating the DNA-isolation and Transfection procedures. The robots are able to generate a daily maximum throughput of 2,304 unique samples. The intensity of apoptosis is quantified and normalised by the CPRG assay, a functional readout exploiting the degradation of membrane integrity in late apoptosis.

The screen will be primarily carried out with the NITE cDNA library, a collection of over 30,000 full length and fully sequenced cDNA acquired from the National Institute of Technology and Evaluation (NITE) in Japan and a smaller RNAi library. As the NITE library is fully sequenced, redundancy is expected to be low and a screen of unique clones on this scale will be representative of the human genome.

Hypothesis and Objectives

We hypothesise that through the screening of a comparatively large collection of clones equivalent to the predicted gene counts from the human genome, the resulting candidates would be representative of a substantial part of the cellular apoptosis signalling network. This large candidate count would allow for bioinformatics, pathway, cluster and enrichment analyses with the prospects of identifying novel and unique signalling responses. The use of nanoparticles to induce toxicity would further enable the identification of various classes of apoptosis regulators, including pro-apoptotic “inducers”, anti-apoptotic “inhibitors” and “sensitizers” which predispose the system to apoptosis. The combined approach would be the foundation for the eventual identification of signalling pathways involved in the apoptosis response, in particular nanoparticles induced toxicity.

The project may be broadly divided into three phases associated with common HTS approaches: Setup, Screen and Analysis.

The Setup phase comprises intensive optimisation of a set of experiments with the objective of identifying the most appropriate conditions to execute the screen. Major experiments include the transformation of the cDNA library (acquired as plasmid DNA) into a suitable bacterial host, identifying a good method of transfection suitable for automation, optimisation of the robotic systems and characterisation of the nanoparticles.

This is followed by the Screen phase which implements the actual high-throughput screen under the optimised conditions. This phase of the research also discusses the control experiments prior to going high-throughput and the candidate threshold determination, and is heavily dominated by repetition and extensive use of the robotic systems. The primary screen probes the entire NITE cDNA library for regulators of apoptosis, and potential candidates are validated in a secondary screen.

The final phase involves extensive analysis of the validated candidates, predominantly focused on gene ontology (GO) clustering, gene set enrichment and pathway analyses, supported by general bioinformatics data mining for available information such as interaction partners and microarray

differential expression analysis of disease datasets. The final aim is to identify the key signalling pathways and hubs active in normal cells and nanoparticles induced toxicity, and possibly the integration of the analysis into a signalling map.

The Robots: History and Descriptions

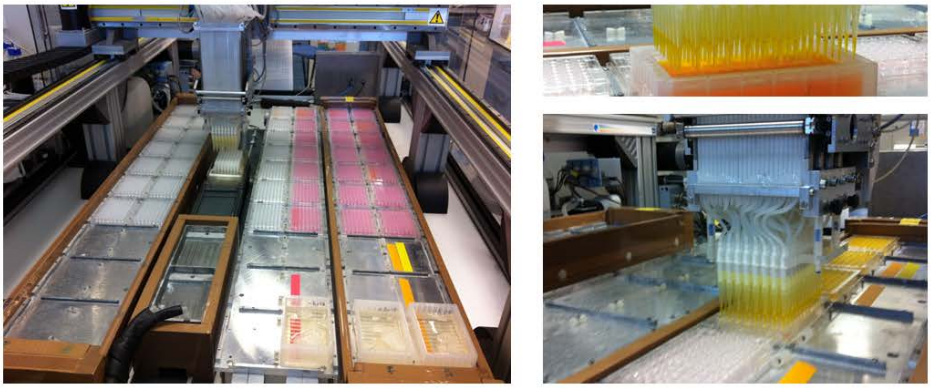
The robotic system comprising of a pair of robots was originally designed and developed at the Max-Plank Institute for Biochemistry in Germany for the specific purpose of implementing the high-throughput screening protocols developed. Both robots operate on similar three axes (x-y-z) gantry on which the main pipetting head is mounted; x and y axes determine the horizontal position of the pipetting head on the platform, while the z axis determines its vertical height.

The Transfection Robot, officially designated as “TFA”, was conceived as a prototype due to its simple requirements of liquid handling functions. Its layout is minimalistic with platforms positions for 96-well microplates arranged in a six by ten layout surrounding the central washing station and four reagent vessels. The washing station cleanses the pipette tips after contact with each reagent, minimising cross contamination and consumable costs. It is capable of processing up to twenty plates of DNA transfection or 1920 samples in a single transfection run [285]. Its minimal layout also allows for easy customisation and implementation of automated protocols.

The DNA Isolation Robot, designated “BASY”, was developed following the success of TFA to specifically handle the DNA isolation protocol adapted from standard alkaline lysis. This protocol differs from the standard in two ways: 1) it involves the co-precipitation of lipo-polysaccharides (LPS) or endotoxins with SDS in isopropanol and 2) it removes residual impurities by washing with acetone. The protocol yields ultra-pure transfection grade plasmids. Hence BASY is much more complex, featuring an integrated computer controlled centrifuge to pellet the cell debris and DNA-silica complexes, dedicated pipetting stations for media removal and acetone wash, integrated chilling platforms at 4°C and -20°C, six reagent vessels with robot-operated lids, shaking platforms and heating platforms in addition to the wash station [286]. The result was that BASY was designed with rigidity and its 4.5-hour automated routine was solely dedicated to implement the DNA isolation protocol. It is capable of processing eight deep well plates (DWP) per run with a yield of 768 samples, but multiple runs could be achieved daily.

The robot pair combines to generate an average throughput of over 2,000 samples, and together with the FLUOstar OPTIMA (BMG Labtech) spectrophotometer forms the automation technology available at our group.

A) Transfection Robot



B) DNA Isolation Robot

Figure 2: Transfection and DNA Isolation Robots. A) Layout of the TFA setup for the addition of transfection complexes to the cells and the additional of CPRG substrate during assay and B) Overview of the DNA Isolation layout, dedicated pipetting stations for media removal and acetone wash, and transport of DWP using the robotic gripper to the integrated centrifuge.

The Library

The NITE Library was a sub-library of sequenced human cDNAs from the "full-length long Japan" (FLJ) collection, which was generated with the aim of providing a foundation for the functional genomics research since 1999. It was acquired from the Japanese NITE Institute and consists of approximately 30,021 individual cDNA clones. Collection, construction and annotation of part of the FLJ collection was published by Ota *et al.* in 2004 [237] and indicated a significant portion (14,490 / 21,243, 68.2%) of the published clones to be unique to FLJ collection. The library consists of a mix between protein coding sequences and non-coding RNA candidates. The FLJ collection was generated such that it includes only full-length and completely sequenced cDNAs, in order to address the short-comings of extensive public sources such as the IMAGE collection or in-house generated libraries which are mostly partially sequenced and annotated using express-sequence tags (ESTs).

The FLJ collection was generated from diverse sources, with the published collection from 61 tissues, 21 primary cell cultures and 16 cell lines. The cDNAs were generated using oligo-capping method, which uses the tobacco acid pyrophosphatase (TAP) to remove the RNA capping structures of matured RNA and ligating a synthetic oligoribonucleotide tag in its place with T4 RNA ligase. The reaction was specific for matured RNA sequences as alkaline phosphatase was used to remove any exposed 51-phosphates of non-capped RNAs prior to the TAP reaction [287, 288]. The synthesised cDNA were clone predominantly into the vector pME-18SFL3 under the control of a mammalian SRalpha promoter, which is a derivative based on the SV40 promoter [289].

Over one million clones were randomly selected and subjected to one-pass sequencing. The sequences matching RefSeq entries were excluded, leaving only those matching human ESTs or without matches. These were then completely sequenced using the standard primer walking method with an accuracy of 99.99%, and all sequences were deposited at the DNA database of Japan (DDJB) [237].

While redundancy is expected within the NITE library in that multiple sequences may lead to the same protein products, each individual cDNA sequence was transcribed in its biological sources, and hence may be the source of splice isoforms generation or novel regulatory mechanism.

β-galactosidase CPRG Colorimetric Assay

Apoptosis assays have placed emphasis on the involvement of caspases as the primary markers of apoptosis, but studies have revealed many caspase-independent apoptosis signalling. Hence, the use of more general phenotypes such as the degradation of cell membrane integrity in late stage apoptosis would avoid screening for candidate regulators which are only caspase-dependent.

Assays which utilise reporters such as luciferase and GFP are also frequently dependent on a single signal whether luminance or fluorescence, with the signal strength correlating with the intensity of the phenotype-of-interest. This type of assay signal is excellent when the reporters are constitutively expressed across all cells in a normal versus compound-treated scenario. However, functional genetic screens are dependent on transient transfection, its efficiency of which frequently varies between experiments. This causes fluctuation in the data generated in the absence of normalisation.

Our choice of readout for apoptosis is the β-galactosidase CPRG Colorimetric Assay, shortened as the CPRG assay, which addresses the two major weaknesses mentioned above. A plasmid containing the β-galactosidase reporter enzyme (pcDNA3.1-BetaGal) is co-transfected with the library cDNA in each transfection, constitutively expressing the reporter only in successfully transfected cells. Hence the amount of reporter present and its signal intensity are positively correlated with the quality of transfection. The β-galactosidase substrate, chlorophenolred-β-D galacto-pyranoside or CPRG, is added to the culture media at the point of assay. The enzyme-substrate interaction is prevented by an intact plasma membrane present only in normal healthy cells, while the “leaky membrane” as a result of membrane blebbing and cellular packaging in late apoptosis allows the CPRG substrate to enter the cells initiating the reaction. This provides a convenient colorimetric assay where the CPRG is hydrolysed, giving a visual colour change from yellow to deep purple quantifiable at absorbance 590 nm. A mild detergent (1% triton) is added to induce total cell lysis once the reaction reaches equilibrium, allowing the total reporter signal to be quantified (Figure 3). The ratio between the two signals indicates the intensity of cell death. However, the CPRG assay is also sensitive to other forms of cell death or damages whereby the membrane integrity is affected, hence a secondary assay to validate apoptosis activity of the candidates will be necessary.

Equation 1: Calculation of CPRG Ratio, an estimator of cell death

$$\text{CPRG Ratio} = \frac{CPRG\ Signal\ "Apoptosis"}{Triton\ Signal\ "Total\ Transfection"}$$

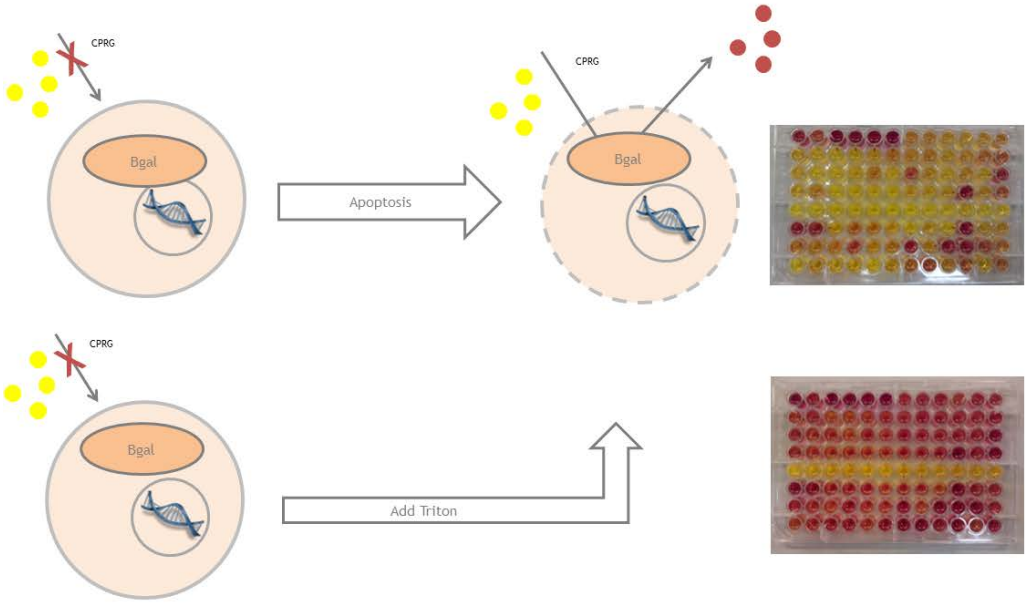


Figure 3: Mechanism of the CPRG Assay. The CPRG Assay exploits the membrane permeablisation in late apoptosis to allow the interaction of the expressed β-galactosidase with the substrate in the cell to generate a quantifiable signal. The signal before lysis indicates the levels of cell death while the manual lysis with a mild detergent gives the total reporter signal from the transient transfection, hence the ratio between the signal before and after manual lysis indicates the intensity of cell death occurring.

Chapter 3: Materials and Methods

Reagent List

Chemicals and Kits

Sigma Aldrich

2-mercaptoethanol (M7522), Acrylamide (40% solution, A7168), Ammonium persulfate (A3678), Ampicillin sodium salt (A9518), Agarose (A9539), Bovine serum albumin (BSA, A7906), Bicinchoninic acid kit (BCA1-1KT), Calcium chloride (C7902), Caspase-3 assay kit (CASP-3-C), Dimethyl sulfoxide (D2650), DMEM AQmedia™ (D0819), Earle's Balanced Salt Solution (E2888), Fetal calf serum (F7524), LB-broth (L3022), L-glutamine (G7513), Magnesium sulphate heptahydrate (63138), Polybrene® (H9268), Polyethylenimine Branched (MW 25,000) (408727), Propidium iodide solution (P4864), Puromycin dihydrochloride (P8833), Sodium azide (S8032), Sodium butyrate (30,341-0), Sodium citrate (S1804), Sodium deoxycholate (30970), Sodium pyruvate (S8636), TEMED (T9281), Triton X-100 (T8532), TWEEN® (P5927) and Yeast extract (Y1625)

Biotium

GelRed™ Nucleic Acid Gel Stain, 10,000X in DMSO (41002)

Fluka

Polyvinyl alcohol mounting medium with DABCO(10981)

Invitrogen

10x PBS (14200-067), DAPI (D3571), DiOC6 (D273), DMEM (21969-035), Earle's Balanced Salt Solution (24010-043), Library efficiency® DH5αTM chemocompetent bacteria (18263-012), Mito Tracker® Red CMXRos (M7512), Penicillin/streptomycin (100U/ml and 100µg/ml, respectively, 15070-063), PicoGreen (P7589), pLenti7.3/V5-TOPO® TA cloning kit (K5310-00), PureLinkTM HiPure plasmid filter purification kit (K2100-07), TMRE (T669), Trypsin-EDTA solution (0.5%, 15400-54) and UltraPure™ DNA Typing Grade® 50X TAE Buffer (24710-030).

All primer oligos were purchased from Invitrogen.

Fermentas

5x protein loading buffer (R0891), 50x TAE buffer (#B49), 6X DNA Loading Dye (R0611), Bradford (500-0006), GeneRule™ 1KB Plus DNA Ladder (SM1333), GeneRule™ 1KB DNA Ladder (SM0314), FastAP Thermosensitive Alkaline Phosphatase (EF0651), PageRulerTM prestained protein ladder (SM1811 or SM0671), ProteoBlock™ protease inhibitor cocktail (R1321) and T4 DNA Ligase HC (EL0013).

All restriction enzymes used were FastDigest® enzymes purchased from Fermentas.

Bio-Rad Laboratories

0.5M Tris-HCl pH 6.8 solution (161-0799), 1.5M Tris-HCl pH 8.8 solution (161-0798), 10x TGS buffer (161-0772) and SDS solution (10% w/v, 161-0416).

GE Healthcare

Hyperfilm™ high performance chemoluminescent film (28906837)

Promega

Wizard® SV Gel and PCR Clean-Up System (A9282) and Caspase-Glo® 3/7 Apoptosis Quantification Kit (G8091).

Polysciences

Polyethylenimine, Linear (MW 25,000) (23966-2)

QIAGEN

Attractene Transfection Reagent (301005) and Effectene Transfection Reagent (301425).

OZ Biosciences

LipoMag Transfection Kit (LM80000)

New England Biolabs (Finnzymes)

Phusion High-Fidelity DNA Polymerase (F-530L)

Clontech

In-Fusion™ Advantage PCR Cloning Kit (639635)

Millipore

CaspaTag Caspase 3,7 In Situ Assay Kit Fluorescein (APT423), CaspaTag Caspase 8 In Situ Assay Kit Sulforhodamine (APT503) and PVDF membrane (IPVH00010).

Roche

Cell Death Detection ELISA^{PLUS} 10X (1920685) and Chlorophenolred-β-D galacto-pyranoside (10884308001).

MP Biomedicals

zVAD-fmk (O103)

OXOID Ltd.

LB-agar (852323)

Pierce

Enhanced chemoluminescent reagent (#32106)

Thermo Scientific

Paraformaldehyde (28908)

Merck

Colorimetric caspase-3 substrate I (Merck 235400) and fluorogenic caspase-3 substrate II (Merck 235425).

Plasmid Vectors

The standard mammalian expression system employed by our group is the pcDNA3.1 vector (Invitrogen) driven by the CMV promoter. Our group has created a number of derivatives from pcDNA3.1. pcDNA3Δ has the neomycin resistance cassette deleted, its size reduction contributes to improved transfection efficiency. pcDNA3-HA, pcDNA- tCFP(NheI)TAA(NotI) and pcDNA3-(NheI)YFP were used for labelling of proteins with the hemagglutinin (HA) or fluorescence tag. Unless otherwise stated, cloning was performed using the commercially available pcDNA3.1.

The NITE Library was cloned using the pME18SFL3 vector, an SV40-based promoter system for mammalian expression. Apoptosis positive and negative controls such as RIPK1, Caspase-2 and GFP were cloned under the pME18SFL3 vector.

Viral production uses the pLenti7.3 vector (Invitrogen) and a derivative pLenti7.3puro where the GFP cassette was replaced with the puromycin cassette allowing for stable clone selection. Plasmids encoding viral proteins required for co-transfection include group antigen and reverse transcriptase (pGag.Pol), stomatitis virus glycoprotein (pVSV-G) and reverse transcriptase (pRev). pAdVantage (Promega) was co-transfected with the viral plasmids to enhance protein expression.

Other plasmids containing reporter proteins routinely used include pGreenLantern-1 (Invitrogen), pEGFP-C1 (Clontech), pcDNA3-DsRed, pcDNA3-Luciferase and pcDNA3-β-galactosidase.

Molecular Biology

Bacterial Culture

Bacteria containing plasmids were cultured in Lysogeny broth plus (LB^+) media (2% w/v LB, 1% w/v Yeast Extract) in the presence of 100ug/mL ampicillin. Bacteria was inoculated in 25 or 300mL media depending on the scale of DNA isolation and incubated at 37°C and 260 rpm for 24 hours. For large-scale experiments, 50mL starter cultures were used to inoculate 5L of media and cultured at 37°C until the culture reaches optical density of 0.6 at absorbance 600 nm.

Library cDNA clones were cultured in 96-well DWP. 1.4mL LB^+ media per well was dispensed by the robot and each DWP manually inoculated with approximately 5ul of glycerol stocks. The DWP was then sealed with gas permeable adhesive seal (ABgene, AB-1450) and incubated at 37°C and 260 rpm for 24 hours. Plate inoculation may be automated on the BASY robotic platform.

Glycerol stocks (30% v/v glycerol) of bacterial culture were prepared and keep at -80C for long term storage.

Preparation of Competent

A starter culture of the desired *Escherichia coli* strain (DH5α was used) was inoculated and allowed to grow overnight at 37°C, 260rpm. The overnight culture was diluted with fresh media at a 100X factor, and allowed to grow at 37°C, 260rpm until the culture density reached an O.D. of 0.6 at 600nm absorbance. The cells were harvested by centrifugation at 2,000rpm and resuspended in 1/5 volumes of cold 100mM calcium chloride (Sigma-Aldrich). The cell suspension was incubated for 20

minutes on ice. The harvesting step was repeated and the cells resuspended in 1/10 volume of cold 100mM calcium chloride and incubated for 60 minutes on ice. Finally, the cells were harvested and resuspended in 1/10 volume of cold 100mM calcium chloride with 15% glycerol. The cells could be used directly or aliquot into smaller volumes for long term storage at -80°C.

Transformation of Plasmids

Escherichia coli DH5α were used as the bacterial host for plasmid replication. 5ul of plasmid DNA was combined with 50ul of chemo-competent cells and incubated on ice for 15 minutes. The mix was then heat shocked by incubating in 42°C waterbath for 45 seconds before returning to ice for a further 2 minutes incubation. 900ul LB^{+} media was then added for each reaction and the cells allowed to recover at 37°C and 260 rpm incubation for 60 minutes. The cells were then pellet and fresh media added to the desired dilution and plated onto LB agar plate containing the appropriate antibiotic resistance.

This protocol was adapted for high-throughput mode using the BASY platform, with the robot managing all pipetting steps.

Ultra-pure Plasmid DNA Isolation

This protocol was adapted from standard alkaline lysis to yield high quality supercoiled and endotoxin-free plasmids suitable for transfection in mammalian cells.

1.4mL of bacterial cultures were centrifuged at 4,300 rpm (Sigma Robotic Centrifuge 4K15) and the pellet resuspended in 250ul of P1 Buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.0, 100 µg/ml RNase A (Fermentas)). The cells were lysed with the addition of 250ul P2 Buffer (200 mM NaOH, 1% SDS), gently mixed then incubated at ambient temperature for 5 minutes. 250ul of P3 Buffer (3.0M potassium acetate, pH5.5) chilled to 4°C was added next to neutralise the reaction and the mixture left to incubate at ambient temperature for a further 5 minutes. Next, the mixture was centrifuged at 4,300 rpm for 15 minutes after which the supernatant is transferred into new reaction vessels. 120ul P4 Buffer (2.5% SDS in isopropanol) was mixed with the supernatant and incubated at 4°C for 15 minutes then -20°C for another 15 minutes to precipitate the endotoxin. The chilled mixture was centrifuged at 4,300 rpm for 15 minutes and the supernatant transferred into new reaction vessels. 150ul of silica (50mg/mL in water) was added and the mixture allowed to incubate at ambient temperature for 15 minutes for DNA binding to occur. The mixture was pelleted and washed twice with acetone to remove any residual contaminants before drying at 65°C for 15 minutes. Water was finally added to elute the plasmid DNA with the silica pelleted and discarded.

DNA Isolation with Commercial Kits

Plasmid DNA for small-scale experiments were extracted from bacterial using column-based commercially available kits like the Plasmid Mini or Maxi Prep kit (QIAGEN) or the PureLinkTM HiPure plasmid filter purification kit (Invitrogen). All extractions were performed according to manufacturer's protocol.

Quantification of DNA Concentration

DNA extracted using commercial kits was quantified with the NanoDrop (ThermoScientific), using traditional measurement of DNA and protein absorbance at wavelengths 260nm and 280nm respectively, in the absence of any nucleic acid staining.

PicoGreen (Molecular Probes) is a fluorescence nucleic acid stain which binds specifically to double stranded DNA molecules and is sensitive up to the picometer range. DNA was diluted in 100ul TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 7.5). 100ul of a working solution of PicoGreen (1:200 in TE Buffer) was added per well, and the mixture set aside in the dark for 5 minutes incubation at ambient temperature. UV-transparent 96-well plates specifically for DNA measurements were used and the fluorescence signal was quantified using the FLUOstar OPTIMA (BMG Labtech) at excitation 480nm and emission 520nm. The signal was used to derive the DNA concentration using a standard curve. The PicoGreen protocol was employed for samples generated using the ultra-pure plasmids isolation method where DNA was captured using silica binding, and for samples requiring higher degree of sensitivity. The presence of minimal silica residues in the DNA solution interferes with NanoDrop quantification due to similar absorbance wavelength between DNA and silica at approximately 260nm.

Restriction Enzyme Reactions

All restriction enzyme digests were performed using the Fermentas FastDigest® line of restriction enzymes. 1ug of plasmids were mixed with 1ul of the appropriate enzyme and 2ul of 10X FastDigest® Reaction Buffer, and the final reaction volume made up with nuclease-free water to 20ul. Double and multiple digests may be scaled up proportionally, with the combined volume of enzymes not exceeding 10% of total volume. The reaction was incubated at 37°C for 20 minutes and heat-activated at 80°C for 5 minutes prior to use or gel electrophoresis.

DNA Gel Electrophoresis

1% w/v agarose gel was prepared in TAE Buffer (40 mM Tris-acetate, 1 mM EDTA, pH of 8.3) with the nucleic acid stain GelRed (Biotium) added according to manufacturer's protocol, and set in a casting apparatus. DNA samples were mixed DNA loading buffer in a 5:1 ratio and loaded onto the gel. 5ul of

GeneRule™ 1KB Plus DNA Ladder (Fermentas) was loaded as the DNA size marker. The gel was run in excess TAE Buffer at 90V for 45 – 60 minutes.

Cloning

Polymerase chain reaction (PCR) was performed with the Phusion High-Fidelity DNA Polymerase (Finnzymes). The reaction mix was prepared by mixing 1ul of dNTP (10mM), 10ul Phusion HF Buffer, 7.5ul each of forward and reverse primers (200uM), 0.5ul of Phusion polymerase, 1ul of DNA template (200ng), 1.5ul of DMSO and the reactions made up with water to a final volume of 50ul. The PCR reaction was initiated with denaturation (98°C, 59 seconds), then cycled through the amplification phase with denaturation (98°C, 30 seconds), annealing (50°C, 20 seconds), extension (72°C, 30 seconds per 1000 bases), before a final extension at 72°C for 5 minutes.

The appropriate fragment was isolated after gel electrophoresis and purified using the Wizard® SV Gel and PCR Clean-Up Kit (Promega), and subjected to the appropriate restriction digest. The vector was cut with the same restriction enzymes and 1ul of FastAP Thermosensitive Alkaline Phosphatase (Fermentas) added at the final 5 minutes of the 37°C incubation. Both reactions were then purified again using the PCR Clean-Up Kit before ligation.

For ligation, Vector:Insert were mixed in a 1:5 ratio, together with 2ul of 10X T4 Ligase Buffer and 1 unit of T4 DNA Ligase (Fermentas), and the reaction made up to a final volume of 20ul with water. The reaction was allowed to incubate for 10 minutes at ambient temperature before heat inactivation at 70°C for 5 minutes. The reaction mix was then used for heat shock transformation into bacterial host as previously detailed.

Difficult cloning was achieved with the In-Fusion™ Advantage PCR Cloning Kit (Clontech). The In-Fusion enzyme was able to create single-stranded regions at the complementary ends of the linearised vector and PCR insert, the allowing the DNA molecules to anneal together; the strands were repaired upon transformation into bacterial hosts. The kit was used according to manufacturer's protocol.

Mammalian Cell Culture and Transfection

Cell Culture

Human embryonic kidney (HEK293T) cells, 293FT (Invitrogen; for viral production), adenocarcinomic human alveolar basal epithelial cells (A549) and human cervical carcinoma (HeLa, ATCC) cells were cultivated in high glucose Dulbecco's modified eagle medium (DMEM) or DMEM AQmedia™ (Sigma-Aldrich) supplemented with heat inactivated foetal calf serum (FCS, Sigma-Aldrich), 2mM sodium pyruvate (Sigma-Aldrich), penicillin-streptomycin (100U/mL, Sigma-Aldrich) and 2mM L-glutamine (Sigma-Aldrich). HEK293T cells were cultivated in the presence of 5% v/v FCS while 10% v/v FCS was used for 293FT, A549 and HeLa cells. Cells were cultured in 10cm tissue culture dishes, T75 or T175 vented flasks dependent on application at 5% atmospheric CO2 and 37°C in a humidified incubator.

Cells were resuspended in freezing media (supplemented DMEM with 20% FCS and 10% DMSO) and frozen at -80°C for long term storage.

Cells were seeded at 50 – 60% confluency for transfection experiments and 70 – 80% for chemical treatment experiments. For the screening experiments, HEK293T cells were seeded at 14,000 cells per well in a 96-well microplate.

Calcium Phosphate Transfection

For 96-well format, 15 µl of 0.25M calcium chloride mixed with 15 µl of plasmids (300ng plasmid DNA), then 10 µl of Hepes buffer saline (HBS, pH 7.1) and 15 µl of 2mM chloroquine were added immediately. 15 µl of the transfection mix was added to cells without incubation. The cells were incubated in the presence of the transfection mix at 37°C for 6 hours, after which media was replaced. This protocol may be proportionally scaled upwards for 24- and 6- well format by a factor of 4 and 20 respectively.

Polyethylenimine (PEI) Transfection[1]

Linear PEI (Polysciences, MW 25,000 Da) stock solution was prepared by completely dissolving 1g of linear PEI in 100ml (10mg/mL) of sterile hot water. Once the solution has equilibrated to ambient temperature, the solution is neutralised to pH 7.0 using 37% hydrochloric acid (Sigma-Aldrich). The solution was then passed through a 0.22um filter.

Branched PEI (Sigma-Aldrich, MW 25,000 Da) stock solution was prepared by dissolving 2g of branched PEI in 100ml (20mg/mL) of sterile water, neutralised to pH 7.0 using 37% hydrochloric acid

[1] The use of PEI for transfection is currently covered by intellectual property rights including US Patent 6,013,240, European Patent 0,770140, and other foreign equivalents. Polyplus-transfection™ remains its exclusive worldwide licensee. All PEI reagents used in this project are acquired from commercial sources solely for academic research purposes. No PEI preparations were made available for commercial benefits.

(Sigma-Aldrich). The solution was then passed through a 0.22um filter before further dilution with sterile water to 2mg/mL.

Prior to transfection a working stock of the PEI solution was prepared by diluting the stock with water; the linear PEI is diluted ~300X and the branched PEI diluted ~135X in 150mM sodium chloride (Sigma-Aldrich). For transfection in 96-well format, 300ng of plasmids DNA was prepared in 30ul of 150mM sodium chloride. 30ul of the PEI working solution was then added to the DNA solution, and the mixture immediately vortexed or mixed by pipetting. The transfection mix was incubated at ambient temperature for 30 – 60 minutes. The media in the well was changed prior to addition of the transfection mix. 40ul of media was added to the transfection mix and the entire 100ul volume added to the cells. The cells are incubated in the presence of the transfection mix for 5 – 6 hours before the media was replaced. The protocol was proportionally scalable to 24- and 6-well plate formats.

PolyPlus jetPEI Transfection Kit

For 96-well format, 250ng plasmid DNA was diluted in 150mM sodium chloride to a final volume of 50ul; the solution was gently mixed to disperse the DNA. Separately, 0.5ul of the jetPEI transfection reagent was diluted in 150mM sodium chloride to a final volume of 50ul. The jetPEI solution was next added to the DNA solution, mixed and allowed to incubate at ambient temperature for 15 – 30 minutes. The media in each well of the plate is replaced with fresh media containing the transfection mix, and the cells incubated for 5 – 6 hours before media change.

Other Commercial Transfection Kits

Commercially available transfection kits were employed for a range of cell lines that were less efficiently transfected. QIAGEN Effectene and Attractene transfection reagents were employed for HeLa cells transfection, while OZ Biosciences LipoMag transfection kit was used for A549 cells. All commercial kits were used according to the manufacturer's recommended protocol.

Production of Stable Cell Line

The desired cell line was seeded in 6-well plates 15 hours prior to transduction. On the next day, the cells were incubated in the presence of the viral supernatant supplemented with 1ug/mL of Polybrene® (Sigma-Aldrich) for 6 hours. The medium was then replaced with fresh media and incubated further for two to three days, after which the medium was replaced with one containing a lethal dose of puromycin (the dose determined based on a kill curve prior to experiment). The infected cells were maintained in the selection media for three days. After selection, the cells could be passaged and maintained in the appropriate media containing 1ug/mL of puromycin to retain the resistance.

Production of Lentiviral Particles

Viral Production

pLenti7.3puro containing the protein-of-interest was co-transfected with three plasmids containing group antigen and reverse transcriptase (pGag.Pol), stomatitis virus glycoprotein (pVSV-G) and reverse transcriptase (pRev), and pAdVantage (Promega). 20ug of pLenti7.3puro, 7.8ug of pGag.Pol, 5.4ug of pVSV-G, 3.8ug of pRev and 9ug of pAdVantage were combined and made to 437.5ul with water. The DNA solution was transfected using calcium phosphate transfection protocol into 293FT viral production cell line (Invitrogen).

The culture media was replaced with fresh media containing 1mM sodium butyrate (Sigma-Aldrich) the next morning. The supernatant was harvested two to three days post transfection and centrifuged at 500g and at ambient temperature for 10 minutes to remove the cell debris. The supernatant, which contains the viral particles, was then filtered with 0.45um sterile filter; 0.22um filter should be avoided so as not to damage the viral outer glycoproteins.

Viral Titre Calculation

The cell line of interest was seeded on a 12-well plate at 30% confluency. On the day of transduction, the viral stock was diluted at varying dilutions in 400ul of DMEM containing 6ug/mL polybrene® and introduced to the cell line of interest for 48 hours during which integration into the genome and protein expression should occur. The number of cells was counted prior and post selection with 5ug/mL puromycin for 24 hours.

Viral titre was calculated using Equation 2:

Equation 2: Calculation of viral titre

$$TU\ ml^{-1} = (F \times N \times D)/V$$

Where TU is the transforming unit (=number of virion), F is the percentage of surviving cells, N is the number of cells at the time of transduction, D is the dilution factor, and V is the volume of diluted viral sample added into each well in ml.

Assays and Imaging

Propidium Iodide (PI) Staining

The supernatant and the adherent cells were harvested, centrifuged (1500rpm, 5 minutes) and resuspended in lysis buffer (0.1% sodium citrate and 0.1% Triton X-100 in PBS) containing propidium iodide (PI, 20µg/ml). The cell nuclei released by the lysis were acquired with flow cytometer (FACSCalibur™, BD Biosciences) using FL-2 channel to quantify the percentage of cell population with subG1 DNA content. The pan-caspase inhibitor, zVAD-fmk (MP Biomedicals), was used at a concentration of 50µM. Propidium iodide is a DNA stain and hence this assay may be used as an estimator of apoptosis based on quantity of DNA present. This draws on the relationship between DNA quantity and apoptosis, where nucleases activated during apoptosis results in the degradation of DNA.

3,3-dihexaoxacarbocyanine Iodide (DiOC6) Staining

The supernatant and the adherent cells were harvested, centrifuged (1500rpm, 5 minutes) and resuspended in 150ul of PBS containing 40nM 3,3-dihexaoxacarbocyanine iodide (DiOC6) and PI (6µg/ml) for 30 minutes in the incubator and further incubated for 30 minutes at room temperature in dark. Cells were then analysed using a flow cytometry using FL-1 channel for DiOC6 staining and FL-3 for PI staining.

This DiOC6 is retained in intact mitochondria and a decrease in its signal is an indicator of mitochondrial outer membrane depolarisation. PI is excluded from intact cells, and its staining of the cell nuclei represents a degradation of the plasma membrane. In combination, this assay allows the simultaneous quantification of two markers of apoptosis, degradation of membrane integrity in the mitochondria and plasma membrane, hence is capable of estimating populations of early and late apoptosis.

CaspaTag Staining for Caspase Activation (Millipore)

CaspaTag kits (Millipore) uses carboxyfluorescein (FAM) or sulforhodamine (SR) labeled fluoromethyl ketone (FMK) inhibitor probes. These fluorescently labelled inhibitors binds irreversibly to activated caspases via their substrate recognition sequences, caspase-3 & -7 (DEVD), caspase-8 (LETD), caspase-9 (LEHD) and all caspases (VAD), thereby labelling cells undergoing apoptosis with activated caspases. The kit was used according to the manufacturer's recommended protocol, and the FMK or SR labelled probes could be detected with fluorescence microscopy or quantified with flow cytometry using FL1 and FL3 channels respectively.

CPRG Assay

The CPRG assay is dependent on the co-transfection of the β-galactosidase reporter and was performed 48 hours post transfection. For 96-well format, 10ul of CPRG substrate (13mM, Roche) was added per well containing 200ul of phenol-red free media. The reaction was allowed to incubate at ambient temperature for 60 minutes after which the signal intensity was quantified at 590nm. 20ul of 1% Triton lysis buffer was then added. The cells were incubated at ambient temperature for another 60 minutes and the signal intensity quantified at 590. The ratio between the two readings is an indicator of apoptosis (Equation 1).

Cell Death Detection ELISA^{PLUS} (Roche)

The Cell Death Detection ELISA^{PLUS} (Roche) is a quantitative sandwich ELISA assay directed against DNA and histones. The assay specifically detects and quantifies nucleosomes released into the cytosol during apoptosis. Biotin-labelled anti-histone antibodies bind to the streptavidin-coated MTP walls, and the presence of nucleosomes captured by the anti-histone is detected using POD-labelled anti-DNA antibody. The POD reporter enzyme then reacts with the substrate 2,2'-Azino-di[3-ethyl-benz-thiazolin-sulfonat] (ABTS) generating a colorimetric signal that is positively correlated with the apoptosis intensity. The kit was used according to manufacturer's protocol.

Caspase-Glo® 3/7 Apoptosis Quantification Kit (Promega)

The Caspase-Glo® 3/7 kit (Promega, G8091) is a luminescent-based assay that quantifies caspase-3 and -7 activities. The substrate, aminoluciferin, is covalently modified and inactivated using the DEVD tetrapeptide sequence. Luciferase can only catalyse the production of light in the presence of activated caspases which cleave the tetrapeptide, thereby activating the substrate for luciferase cleavage. The kit was used according to manufacturer's protocol.

PARP Cleavage Immunoblotting

Poly (ADP-ribose) polymerase (PARP) is a substrate of caspase-3 and is cleaved during apoptosis. Cells were harvested and lysed in 100ul of RIPA buffer, incubated on ice for 20 minutes. SDS-PAGE/western blot was done using 12% gel and anti-PARP (Cell Signalling #9542, 1:1000 in 5% milk-TBST) probed for 24 hours incubation at 4C; 50ug of proteins was loaded per well. Secondary probe & washing was performed as normal, developing was done with standard ECL.

Full-length PARP is observable at 116kDa while the activation of apoptosis generates a second smaller band at 85kDa. The amount of PARP cleaved may be quantified using imageJ [290] and the apoptosis intensity estimated using Equation 3:

Equation 3: Estimation of apoptosis intensity (%) from PARP cleavage

$$\text{Apoptosis (\%)} = \frac{\text{Intensity Cleaved PARP at 85kDa}}{\text{Intensity of both cleaved and uncleaved PARP at 85 and 116kDa}} \times 100$$

Lactate Dehydrogenase (LDH) Cytotoxicity Assay

The LDH Cytotoxicity Assay (BioChain) is a colorimetric assay that quantifies the amount of LDH, a stable cytosolic enzyme, which is released into the cell culture supernatant upon damage to the plasma membrane. LDH oxidises lactate to pyruvate, which in turns reacts with tetrazolium salt INT forming formazan. Formazan is a water-soluble dye quantifiable at 490nm absorbance and hence the increase in formazan formation is correlated with an increased in number of lysed cells, which is an indication of necrosis. The kit was used according to manufacturer's protocol.

Fluorescence Microscopy

Cells were seeded in wells containing coverslips then transfected and treated accordingly. At the point of microscopy, the cells were fixed at 4°C in 4% para-formaldehyde (PFA, Thermo Scientific) in PBS for 20 minutes. DAPI Mito Tracker® Red CM-H2XRos (MitoTracker) and tetramethylrhodamine ethyl ester (TMRE) (Molecular Probes) staining were used according to the manufacturer's protocol. After staining, the dyes were removed and the coverslips washed three times with water, then mounted onto glass plates with polyvinyl alcohol mounting medium with DABCO (Sigma-Aldrich). Cells were observed and imaged under the fluorescence microscope using the appropriate filters.

Confocal Live Cell Imaging

The confocal laser-scanning microscope (Leica) was setup and equilibrated one hour prior to imaging. Cells were seeded in glass bottom 24-well plates (BD Biosciences) and subjected to the desired experimental conditions were mounted onto the confocal microscope stage, and the system programmed to image between three to five positions per well at an interval of 10 – 15 minutes for a period between 16 – 24 hours at the desired wavelength. Pictures were acquired with LAS AF software and analysed with LAS AF Lite software.

Live cell imaging was used for FRET-based assays such as the caspase-3 FRET sensor generated by our group and imaging of cells undergoing autophagy for time point experiments.

Protein SDS-PAGE Gel Electrophoresis

Preparation of Whole Cell Lysate

Cells were harvested by trypsinisation along with the supernatant depending on the experiment type. The cell suspension was centrifuged at 1,500rpm for 5 minutes and lysed by adding the appropriate lysis buffer (RIPA buffer was commonly used). The cells could also be lysed directly in the plates after media removal and washing. The lysate was incubated on ice for 10 minutes before centrifugation at 13,000 rpm and 4°C for 10 minutes. The supernatant containing soluble proteins was transferred to fresh eppendorf tubes for immediate use or storage at -20°C.

Protein Quantification

Protein concentration was determined using the Bicinchoninic Acid Assay Kit (Sigma-Aldrich). The BCA working solution was prepared by combining 1 part of 4% (w/v) copper(II) sulfate pentahydrate with 50 parts of bicinchoninic acid solution (bicinchoninic acid, sodium carbonate, sodium tartrate, and sodium bicarbonate in 0.1 N NaOH). 10ul of protein samples were added per well in a 96-well plate; a BSA standard curve with a range of 200 – 1000ug/mL was also prepared. 200ul of the BCA working solution was then dispensed to each well, and the reaction mix allowed to incubate at 37°C for 30 minutes or until colour change is visible. The reaction may be accelerated by incubating at 60°C for 15 minutes. Absorbance was measured at 590nm and the sample concentration determined according to the standard curve.

Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

20 – 100ug of the whole cell lysate was resolved by SDS-PAGE (stacking gel: 4.8% polyacrylamide, 125mM Tris-HCl pH6.8, 0.1% SDS, 0.1% APS, 0.1% TEMED; resolving gel: 10-15% polyacrylamide, 375mM Tris-HCl pH8.8, 0.1% SDS, 0.1% APS, 0.05% TEMED) in TGS-running buffer (25mM Tris-HCl, 192mM Glycin, 0.1% w/v SDS). PageRuler^{TM} prestained protein ladder (Fermentas) was used as a size marker.

Proteins were transferred to PVDF membrane (Millipore) in transfer buffer (20% methanol in TGS buffer); semi-dry transfer was performed at 20V for 45 minutes for a single gel.

Western Blot

The membranes were blocked with blocking buffer (5% milk (Sigma-aldrich) in Tris Buffered Saline Tween (TBST; 20 mM Tris, 0.9% NaCl, pH 7.4 and 0.1% Tween-20)) for 30 minutes. The blocking buffer was replaced with the primary antibody and incubated for the desired duration. The membranes were washed with TBST three times, for 10 minutes per cycle. The horseradish peroxidise-conjugated secondary antibody was next added and incubated for 60 minutes. The

membranes were washed three times as before. Enhanced chemoluminescent reagent (Pierce) was applied to the membrane to initiate the HRP reaction and captured using high performance chemoluminescence film (GE Healthcare).

Antibodies used included PARP (Cell Signalling 9542), LC3 (Novus Biologicals NB100-2220), β-actin (Sigma-Aldrich A2228), GAPDH (Santa Cruz SC-32233), GFP-HRP (Santa Cruz SC-8334), Mouse HRP (Molecular Probes G21040), Rabbit HRP (Sigma-Aldrich) and Goat HRP (Sigma-Aldrich A5420).

Analysis

Statistical Analysis

Statistical analysis was performed using unpaired student's t-test. Data were regarded as statistically significant if $p<0.05$ based on the one-sided t-test with unequal variance. Error bars displayed in all figures are indicative of the standard deviation of the sample size of each experiment.

DAVID functional annotation

Functional annotation of the candidate genes was performed using the DAVID functional annotation tool with the default settings, which identify functional clusters enrichments based on gene ontology (GO) terms for biological processes, cellular localisation and molecular functions. DAVID also yielded enrichment of protein interaction, protein domains, tissue specificity and signalling pathways. Biomolecular Interaction Network Database (BIND), Molecular INTeraction database (MINT) and Reactome were used for protein interactions while InterPro, SMART and Protein Information Resources (PIR) databases were used for protein domain enrichment. Tissue specificity used the UniProt tissue specificity annotation.

Gene Set Enrichment Analysis (GSEA)

Gene Set Enrichment Analysis (GSEA) V2.0 was used to identify enriched profiles using three curated Molecular Signatures Database v3.0 collections: : 1) c2 curated gene sets comprising of pathway databases, 2) c3 motif gene sets consisting of conserved cis-regulatory motifs and 3) computational gene sets defined with expression neighbourhoods focused on 380 known cancer genes. A ranked list for the non-treated dataset was generated using the CPRG ratio while the nanoparticles treated dataset were ranked using the normalised TNTD score calculated based on Equation 12. The non-treated dataset corresponds to the inducer candidates while the treated dataset was used for enrichment associated with inhibitors and sensitizers.

Pathway Analysis

Ingenuity® Pathway Analysis (IPA, Ingenuity® Systems, www.ingenuity.com) was used to identify curated signalling pathways from the IPA knowledgebase. This was combined with signalling pathways enriched by DAVID from the KEGG, PANTHER and RECTOME databases, GSEA enriched pathways from the C2 collection, and GeneMANIA network analysis.

Noncoding Sequence Analysis

Noncoding candidate sequences were submitted to the Rfam database to identify structural RNAs including noncoding RNA and cis-regulatory elements, and to RegRNA to identify homologs of regulatory RNA motifs and elements within the noncoding sequences.

The noncoding sequences were also searched for the presence of microRNA using a custom-coded Excel macro based on absolute matches. If the microRNA sequences were derived from IPA, only 7bp seed sequences were available while matured microRNA sequences from the miRBase database comprises of full length microRNA between 18 - 24bp. Information on the microRNA was mined using miRMaid while TargetScanHuman Release 6.2 was used to predict the potential targets of the microRNAs.

Chapter 4: Setup of the RISCI Screen

Background Information

High-throughput screens involve vast repetitions of a specific experimental setup to identify targets-of-interest. It is important that prior to subjecting an experimental setting to such repetitive screening, the optimal conditions are achieved and any potential pitfalls identified. A proper setup translates into an efficient screen which saves on reagents and generates high quality data points within a short period of time.

The RISCI screen comprises the DNA isolation, transfection and assay component phases, all of which needs to be optimal. This section discusses the setup of the RISCI screen for the identification of apoptosis inducers, and inhibitors and sensitizers of nanoparticles induced cell death. The setup phase forms the bulk of the project and is critical to the success of a screen.

Results

High-Throughput Transformation of the NITE cDNA Library

Bioreactor Design and Assembly

The quantity of competent bacterial cells required for the transformation of the NITE Library with 30,000 unique cDNA clones was estimated between 2.5 to 3.0 litres, which needs to be derived from over 30 litres of bacterial cultures. This necessitates the use of a bioreactor for large-scale culture.

To avoid the costs associated with commercial systems, a bioreactor was assembled from commonly used laboratory glassware and consumables. A large 5L volume glass bottle was used as the main culture tank. Prior to assembly, it was filled with 4L water, autoclaved before being equilibrated overnight at 37°C. All other components were sterilised with 100% ethanol.

The 5L glass bottle was set on a magnetic stirrer and covered with a coil of silicone tube connected to a waterbath to maintain the temperature during culture growth. Next, two tubes were introduced into the main culture tank. The air supply tube was inserted from the top and connected a 0.22 µm filter and an air pump, while the sampling tube was inserted through the side entry at the bottle neck. The side entry is connected to a secondary waste flask through a waste tube, leading any

overflow resulting from gas production into the waste flask containing virkon. 1L of 5X LB^{+} media, a starter culture of approximately 100mL *E. coli* DH5α and a thermometer were introduced into the main culture tank, and the content was sealed with a rubber plug (Figure 4A). The cells were allowed to grow until the culture reached a density between 0.5 – 0.6 prior to harvesting.

Sampling of the culture density was achieved via the sampling tube which connected the sterile culture content to a 15mL falcon tube. The pressure within the 15mL falcon tube could be lowered by drawing out the air with a syringe, which causes the bacterial culture to flow into the tube for absorbance assay without disruption to the main culture tank setup. Culture density was determined at 600nm absorbance.

A comparison was made between the growth characteristics of culture derived from the bioreactor setup and the standard culture flask by inoculation of the same amount of starter culture per media volume. 100 µl of culture from each condition was measured every 30 minutes over a three-hour period. The growth rate was reduced by a factor of 3.5 in the bioreactor compared to standard shaking flask culture condition (Figure 4B), but this did not affect the quality of competent cells.

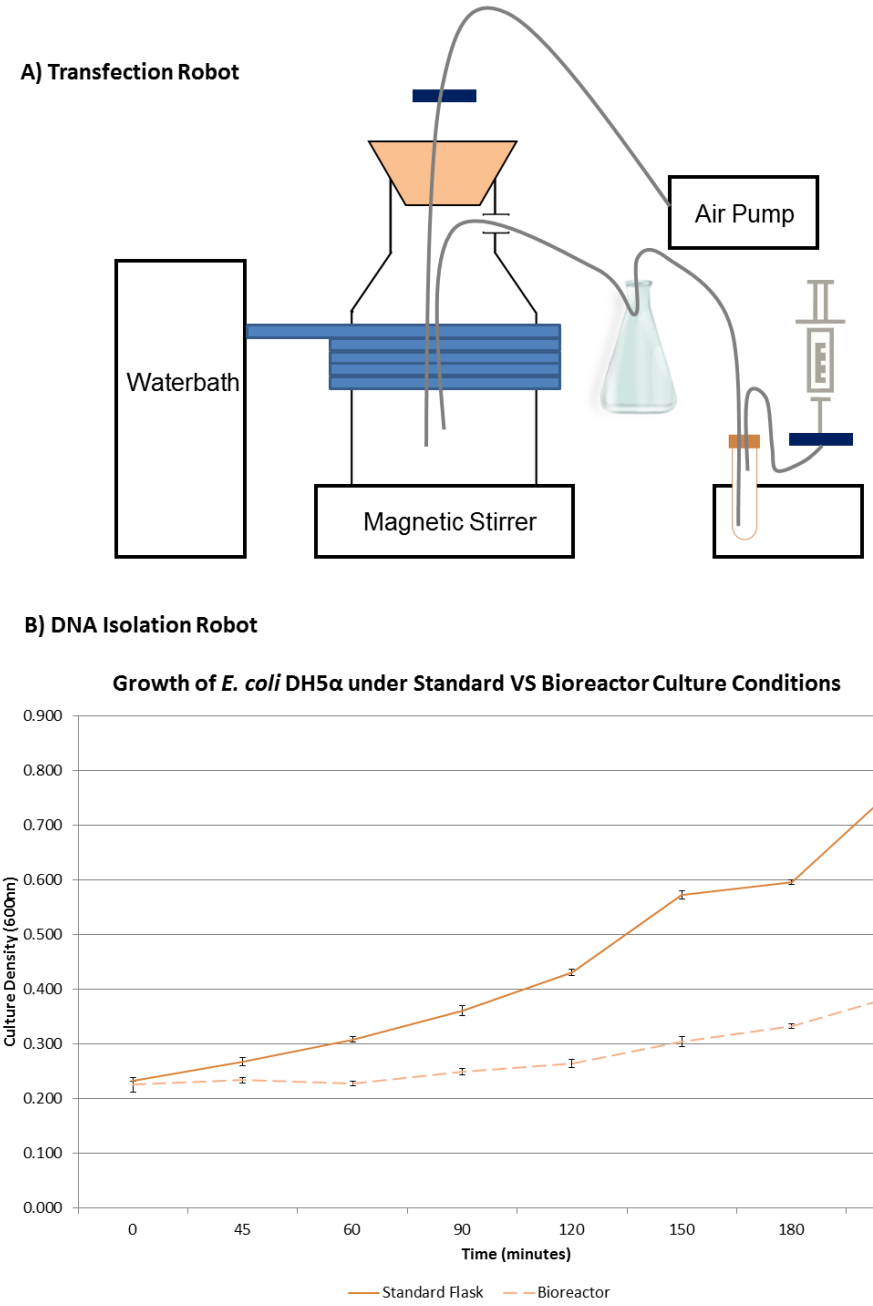


Figure 4: Design and Assembly of a Bioreactor. A) Schematic illustrating the setup of the Bioreactor during large-scale culture of *E. coli* DH5α . The 5L culture was maintained at 37°C through induction from the coils supplied by the waterbath, while filter-sterile air was supplied by a dedicated air pump and the culture constantly mixed using a magnetic stirrer. The entire setup is closed throughout the culture and samples were extracted using a sampling tube to keep the content sterile. B) The growth rate of the bacterial cells within the bioreactor was reduced by a factor of 3.5, likely due to inadequate aeration. However the lower growth rate does not impede on the quality of competent cells and a larger volume of starter culture was introduced to generate the appropriate culture density within the same timeframe. N = 3, error bars represent standard deviation of sample size.

Large-scale Preparation of Competent Cells

A simplified and scaled up version of the protocol described by Chung et al. [291] was used in the large-scale preparation of competent cells. *E. coli* DH5α Library Efficiency competent cells (Invitrogen) were used as the inoculant and host for the NITE Library. Following the bioreactor setup, the DH5α cells were left to multiply until the culture absorbance (A600) reached 0.5. The competent cells were prepared as described in the Materials and Methods (Preparation of Competent).

As the *E. coli* were cultured in antibiotic resistance free condition, quality control steps were introduced to ensure that the competent cells were sterile and not contaminated with bacteria containing antibiotic resistance. Inoculant for the starter culture was aliquot into sterile 1.5mL eppendorf tubes and a single aliquot was used for each large-scale culture and discarded. Samples were retrieved from various steps from the large-scale culture and competent cells preparation process including the starter culture, bioreactor culture and prepared competent cells. The samples were subjected to ampicillin treatment in both liquid LB^{+} and agar plate culture, which allowed for tracing of any potential source of contamination. Only batches of competent cells which were negative for antibiotic resistance and positive for plasmid transformation capacity were used for downstream high-throughput transformation of the NITE Library.

Samples from successful batches were also subjected to heat shock transformation with pGreenLantern-1 and the DNA plasmids isolated and subjected to restriction digest. In one selected control experiment, twenty colonies were randomly picked after the pGreenLantern-1 transformation and the clones digested with BamHI, a single cutter of pGreenLantern-1. All colonies resulted in plasmids with the same size as the pGreeLantern-1, indicating that the plasmids could be replicated in the in-house generated competent cells and that no major sequence recombination was observable (Figure 5).

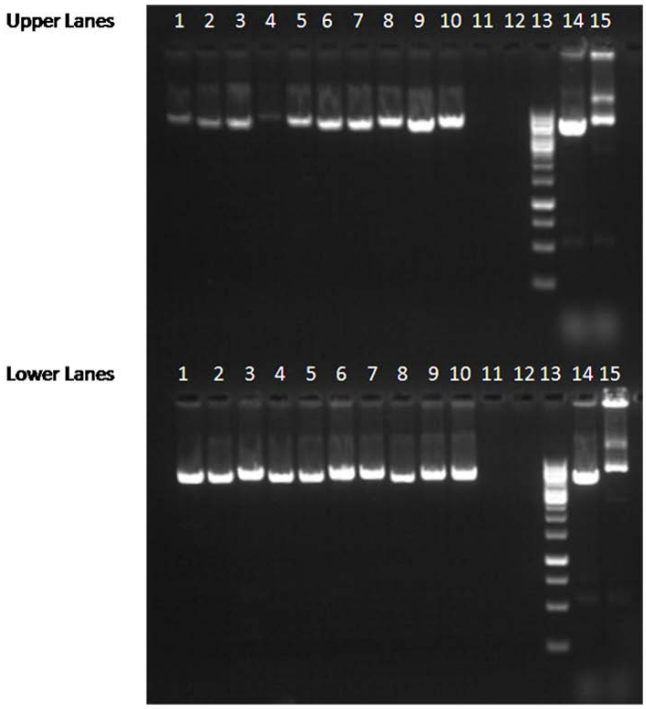


Figure 5: Recovery of pGreenLantern-1 from competent DH5α. pGreenLantern-1 plasmids were transformed into prepared competent E. coli DH5α using the standard heatshock method, 20 colonies were then picked and cultured overnight at 37°C, then miniprep was performed to extract the plasmids. 2.5 µl of the resulting plasmids were then subjected to restriction digest by BamHI, a single cutter of pGreen Lantern-1 before gel electrophoresis on 1% agarose gel. In both upper and lower lanes, lanes 1 – 10 were BamHI digested plasmids extracted from selected colonies, 11 – 12 were empty, 13 – 15 contained Fermentas 1kb O'Gene Ruler, BamHI-digested and BamHI-undigested pGreenLantern-1. All colonies matched the pGreenLantern-1 size, indicating that the plasmid can be transformed into the prepared competent cells and extracted without major mutations or recombinant events.

Adaption of the DNA Isolation Platform for Bacterial Transformation

A new workflow for the high-throughput transformation was designed and implemented on the BASY robot using newly developed software. The software design is separately discussed in Chapter 5.

The wash cycle count was increased to six cycles from three in the DNA isolation protocol, and an ethanol wash step was introduced after each pipetting step which involves bacteria cells. The ethanol wash step was a combination of rinse cycles with water, 30-second immersion in ethanol and finished with further rinse cycles with water. The 4°C refrigerated platform was used as a holding area for the competent cells. The -20°C platform (with an increased temperature) was used for the incubation step after addition of the plasmids to the competent cells as the platform could be closed with its lid to avoid aerial contamination. All positioning routes of the main pipetting arm were programmed such as to minimise and avoid movement across plates containing competent cells after pipetting steps, with the intention of reducing contamination from aerial droplets. Furthermore, direct contact with bacterial cells was avoided where possible such as during addition of plasmid through the use of air pressure to eject the solution. These steps were used to avoid cross contamination occurring between pipetting steps.

The ability of the system to perform liquid handling functions with live biological agents was also tested to determine the level of contamination. A 96-well starter plate containing 150ul of *E. coli* DH5α:pGreenLantern-1 dispensed in a specific pattern of wells (A1, A9, B2, B8, B10, C3, C11, D4, D12, E5, E11, F6, F10, G7, G9 and H8) with the remaining wells containing only the same volume of LB^+. The system was used to inoculate two DWP with 1ml of LB^+ with 50ug/ul ampicillin. The first was performed after system initialisation and three water wash cycle inoculated DWP 1 with 10ul from the starter plate. The pipette tips were next sterilised with the ethanol wash cycle and immersed into the media-containing DWP 2 to determine the levels of contamination. Both DWPs were sealed with the gas permeable membranes and incubated overnight at 37°C and 260rpm. A third plate (DWP 3) was seeded with bacteria in alternate wells and sealed with the Abgene membrane to test cross contamination between wells resulting from shaking. Optical density was measured at 600 nm (A600) after overnight incubation.

The absorbance of each well was illustrated in a 96-well layout in Figure 6. 16/16 (100%) of the targeted wells were successfully inoculated with 4/80 (5%) of the remaining wells of DWP 1 also resulting in bacterial growth. The ethanol step was able to effectively sterilise the pipette tips which were in contact with bacterial cells with all targeted positions of DWP 2 showing no bacterial growth after the ethanol wash. However, 1/96 (1%) well did lead to bacterial growth. DWP 3 which was

processed in ambient condition similar to the BASY platform and sealed with the gas permeable seal (ABgene) reproduced exactly the pattern of inoculation. This indicated that processing of live cultures in open air was not a significant source of contamination, in combination with the use of antibiotic resistance in the culture media. Further, the seal could tightly close the content within each well, preventing cross-contamination which was expected from orbital shaking of the DWP overnight. The gas permeable seal also allowed for gaseous exchange to occur while preventing bacteria from the ambient incubator atmosphere to enter and grow in the plate.

Unintended growth of bacteria in non-inoculated wells could be attributed to the movement of the robotic pipetting arm over the plates. While the software routes were programmed to minimise such movement across other plates, the layout of the 96-well plate format physically impose the need for some tips to move across other positions as the robot arrive in its destination coordinates. Such movement cannot be resolved and remains a limitation of these high-throughput formats. However, minute contamination resulting for small number of bacterial cells would only be visible in wells that were not inoculated. Inoculation of wells with high density starter cultures or transformation mix would contain millions of cells, which would be able to exponentially out-grow any trace contamination. The expectation would be that aerial contamination resulting from robotic movements would only form a small proportional of any resulting culture, allowing the intended plasmids to be replicated.

Overall, the control experiments indicated that open-air high-throughput processing the transformation process using the BASY DNA Isolation platform does not impede on the quality of the library when the adapted procedures such as ethanol wash, use of DWP sealing membranes, strict molecular biology techniques and quality control initiatives are integrated.

A) BASY Platform is able to handle live culture

DWP 1: Successful inoculation with low contamination

	1	2	3	4	5	6	7	8	9	10	11	12
A	1.440	-0.021	1.010	-0.014	-0.014	-0.009	-0.008	-0.005	1.459	0.583	-0.001	0.008
B	0.003	1.527	0.022	0.009	0.018	0.014	0.027	1.451	0.043	1.446	0.042	0.094
C	0.000	-0.010	1.581	0.018	1.015	0.004	0.015	0.016	0.030	0.043	1.559	0.025
D	-0.005	0.003	0.016	1.490	0.016	0.027	0.030	0.029	0.039	0.055	0.038	1.502
E	0.002	-0.001	0.002	-0.003	1.530	0.008	0.011	0.017	0.028	0.035	1.550	0.036
F	0.016	0.021	0.117	0.029	0.037	1.440	0.043	0.035	0.054	1.516	0.048	0.081
G	0.009	0.003	0.015	0.008	0.021	0.013	1.444	0.026	1.442	0.048	0.040	0.048
H	0.007	0.013	0.031	0.019	0.028	0.034	0.049	1.480	0.046	0.061	0.043	0.040

DWP 2: The new ethanol wash function effectively sterilises the bacteria-contacted tips

	1	2	3	4	5	6	7	8	9	10	11	12
A	-0.001	-0.015	-0.014	-0.006	0.004	-0.001	-0.006	0.001	0.005	0.018	0.009	0.014
B	0.004	0.018	0.025	0.010	0.020	0.027	0.039	0.019	0.044	0.054	0.054	0.028
C	0.007	-0.009	0.001	-0.007	0.007	0.005	0.019	0.023	0.029	0.044	0.030	0.032
D	0.002	-0.009	0.013	0.016	0.021	0.035	0.035	0.037	0.046	0.067	0.048	0.032
E	0.010	-0.013	0.008	-0.002	0.016	1.634	0.029	0.024	0.031	0.042	0.045	0.048
F	0.020	0.018	0.031	0.033	0.041	0.046	0.048	0.046	0.059	0.068	0.059	0.063
G	0.009	-0.012	0.007	0.005	0.021	0.028	0.022	0.027	0.032	0.047	0.046	0.052
H	0.010	0.019	0.031	0.016	0.036	0.035	0.050	0.051	0.051	0.060	0.051	0.041

B) Contamination risks from ambient exposure

DWP 3: Gas permeable seals prevents aerial and cross contamination between wells

	1	2	3	4	5	6	7	8	9	10	11	12
A	1.539	0.018	1.537	0.023	1.542	0.021	1.543	0.020	1.527	0.046	1.508	0.037
B	0.032	0.040	0.050	0.042	0.044	0.055	0.066	0.045	0.065	0.078	0.075	0.053
C	1.558	0.022	1.593	0.026	1.579	0.040	1.584	0.052	1.569	0.059	1.516	0.060
D	0.039	0.038	0.049	0.044	0.051	0.063	0.070	0.062	0.076	0.086	0.075	0.066
E	1.568	0.034	1.592	0.049	1.591	0.051	1.603	0.051	1.570	0.067	1.571	0.068
F	0.052	0.060	0.062	0.057	0.067	0.069	0.077	0.067	0.081	0.089	0.080	0.079
G	1.698	0.039	1.643	0.047	1.506	0.057	1.666	0.058	1.549	0.083	1.574	0.085
H	0.043	0.056	0.062	0.052	0.067	0.047	0.080	0.077	0.063	0.094	0.081	0.078

Figure 6: The adapted DNA Isolation Platform can handle live cultures with minimal cross-contamination. Culture density at 600nm absorbance was displayed in a 96-well plate layout for each experiment condition. A) The BASY platform may be adapted to perform liquid handling functions involving live cultures with low contamination. All targeted wells (green) were successful inoculated and 5% of the non-targeted wells resulted in bacterial growth (red) as shown in DWP 1. DWP 2 indicated that the ethanol step introduced could effectively sterilise pipette tips in contact with bacteria and prevent well-to-well contamination across plates. The unintended growth observed in both DWPs was likely to be due to aerial micro-droplets generated as the robotic arm moves across the plates. B) Exposure to ambient conditions in similar to those of the BASY platform was not a major source of contamination. The gas permeable seal (ABgene) was able to tightly close the DWP content, preventing cross-contamination between wells as the DWP is subjected to orbital shaking. It allows gaseous exchange while preventing contamination from the general incubator atmosphere.

Implementation of HT Transformation

The traditional heat shock protocol [292] was adapted to 96-well format for processing using the robotic BASY DNA Isolation platform. All pipetting steps were automated by the robot, enabling the simultaneous processing of 96 samples per pipetting cycle. Briefly, 50ul of competent cells were dispensed into DWP and eight plates were processed during each robotic transformation run. 15ul of plasmids from the original library acquired from the NITE Institute was added into each well and incubated for at least 10 minutes at 4°C in the enclosed -20°C platform. Afterwards, each DWP was retrieved by the robot and manually heat shocked at 42°C in a waterbath before returning for further incubation at the -20°C platform approximately 5 minutes. 400ul of LB^+ lacking ampicillin was introduced to each well and the DWP incubated for 60 minutes at 37°C, after which 500ul of LB^+ containing doubled concentration of ampicillin was added and the cells allowed to incubate for a further 3 hours at 37°C for the first selection. 5ul of the ampicillin-selected transformation mix was used to inoculate 1mL of fresh LB^+ media with ampicillin and incubated overnight (20 – 24 hours) at 37°C with shaking at 260rpm. Finally, 250 µl of bacterial glycerol stocks were created in 30% glycerol.

The protocol was processed on the robot platform in five phases: 1) Dispensing of competent cells into deep well plates, 2) Addition of plasmid DNA, incubation and heat shock, 3) First selection with ampicillin, 4) Dispensing of LB+ with ampicillin into deep well plates and inoculation of bacterial cells, and 5) Preparation of glycerol stocks. Quality control steps were integrated after each medium dispense step in Phase 2 – 4, and these were subjected to ampicillin selection.

The primary round of transformation was completed with a success rate of 73.68% consisting of 224/304 plates. Unsuccessful plates were due to various types of contamination occurring during the intensive transformation process including the acquisition of ampicillin resistance, all of which were identified with the quality control procedures (Figure 7A). Clones which failed to become transformed into the bacterial host were subjected to two further rounds of heat shock transformation, both of which were manually performed since the robot was unable to manage individual pipetting and cherry picking functions. Clones which continued to fail the transformation were no longer pursued and excluded from the subsequent screens.

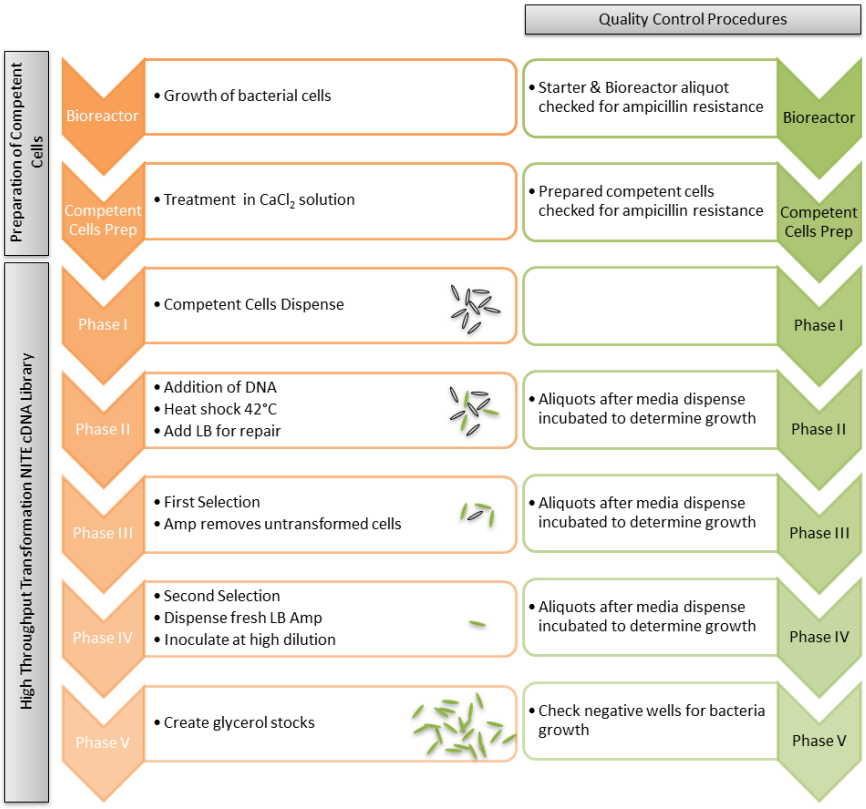
Validation of NITE Library Integrity

In addition to the validation of the transformation capacity of the prepared competent cells, the high-throughput transformation process was trialled using eight plates (Plate ID. AE0001 – AE0008) selected from the NITE Library. Then, 160 clones were randomly selected and the plasmid DNA was extracted using standard alkaline lysis. The resulting plasmids were subjected to EcoRI and NotI

restriction digest and the pattern matched with those predicted by software based on the positions of the EcoRI and NotI within the published sequences of the clones. The workings of the computer script for this purpose will be separately discussed in Chapter5: Software Development. Clones were also sequenced to verify the sequence accuracy.

143 out of 160 clones screened matched the calculated restriction pattern (Figure 7B). In the 17 wells which did not match the predicted pattern, all digested plasmids also did not contain the vector band. These 17 clones were subsequently determined to be part of a minority cDNA sequences which were cloned into the pUC19FL3 instead of the pME18SFL3 vector. Only a total of 322 cDNA sequences were cloned into the pUC19FL3, with all remaining sequences in the NITE library cloned into pME18SFL3. Hence, all clones screened generated restriction patterns matching the software predicted patterns expected from the physical location within the library. DNA sequencing data also confirmed the identity of the sequences as those generated by the NITE Institute. This confirmed the quality of the competent cells and the newly developed high-throughput transformation workflow in maintaining the integrity of the cDNA library. Furthermore, the results indicated that there is no positional cross contamination and that all clones tested are indeed those stipulated by the database location. This allowed for the identification of positive candidates downstream based on available database information without extensive and costly experimental validation.

A) Workflow of the High-throughput Transformation



B) cDNA transformed can be recovered with major plasmid reorganisation

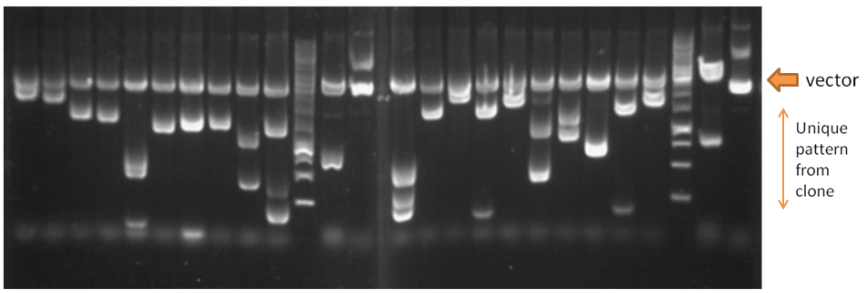


Figure 7: Implementation of high-throughput transformation. A) Workflow detailing the process from competent cells preparation to the transformation of the NITE cDNA Library and its subsequent creation of glycerol stocks for long-term storage. Quality control samples and checks were implemented to trace any potential contamination. B) Validation of the NITE cDNA Library after transformation. All plasmids tested matched the predicted restriction pattern, indicating that no major plasmid reorganisation or contamination occurred.

Transfection Optimisation

Calcium phosphate optimisation

The initial calcium phosphate protocol obtained from the transfection robot programme prepared the DNA-calcium phosphate transfection mix by adding 25 µl of calcium chloride (0.25 M) to 15 µl of DNA solution generated by the DNA isolation robot, followed by 30 µl of Hepes buffer saline (HBS, pH 7.1) and a 13 minutes first incubation at room temperature. Then, 10 µl of chloroquine (2mM) were added and the second incubation at room temperature began for 3 minutes, after which 10 µl of the transfection mix were introduced into 293T cells in 20% FCS-DMEM (Sigma) for 5 hours before the medium was replaced with 5% FCS-DMEM. However, this protocol was found to be inefficient, and modification to the above protocol by using 15ul volume for each component was found to significantly improve transfection efficiency.

In order to further optimise the protocol, a calibration curve of GFP fluorescence signal from the FLUOstar plate reader against the transfection efficiency derived from flow cytometry was prepared to allow for estimation of the transfection efficiency base on GFP fluorescence. 0 – 2.2ug/ul of pGreenLantern-1 was transfected using the protocol using 15ul of each transfection component, with eight replicates for each DNA concentration in a 96-well plate. 24 hours post transfection, the fluorescence was measured using the Optima FLUOstar plate reader, four replicates of each condition were then pooled and subjected to flow cytometry. Correlation coefficient for the dataset was 0.827, indicating good correlation between GFP fluorescence and transfection efficiency (%) (Figure 8A). From the calibration curve, the transfection efficiency may be estimated from GFP fluorescence using Equation 4:

Equation 4: Transfection efficiency estimated from GFP fluorescence

$$Transfection\ Efficiency\ (\%) = 0.0015(GFP\ Fluorescence) + 32.045$$

The optimal volume of each component was then fine-tuned by individually varying the volume of each component between 0 – 35ul while keeping the other components constant as its previously determined optimum. The components were individually investigated in this order: HBS, calcium chloride, chloroquine and final transfection mix volume introduced to the cells. Variation of the HBS volume contributed most to the improvement in transfection efficiency (Figure 8B). The optimised calcium phosphate protocol uses 15 µl of 0.25M calcium chloride mixed with 15 µl of plasmids, then 10 µl of HBS and 15 µl of 2mM chloroquine were added immediately, and 15 µl of the transfection mix was added to cells without any incubation. Incubation was found to be suitable for higher DNA

concentration hence a more acidic transfection mix conditions but incubation at low DNA concentration reduces transfection efficiency. The optimal conditions were also tested on a gradient of cell density from 5,000 – 340,000 cells and it was found that about 60% confluency or between 14,000 – 20,000 cells per well was ideal. The optimised conditions resulted in a net fluorescence of 21737.81 ± 4603.07 or an average of 65%.

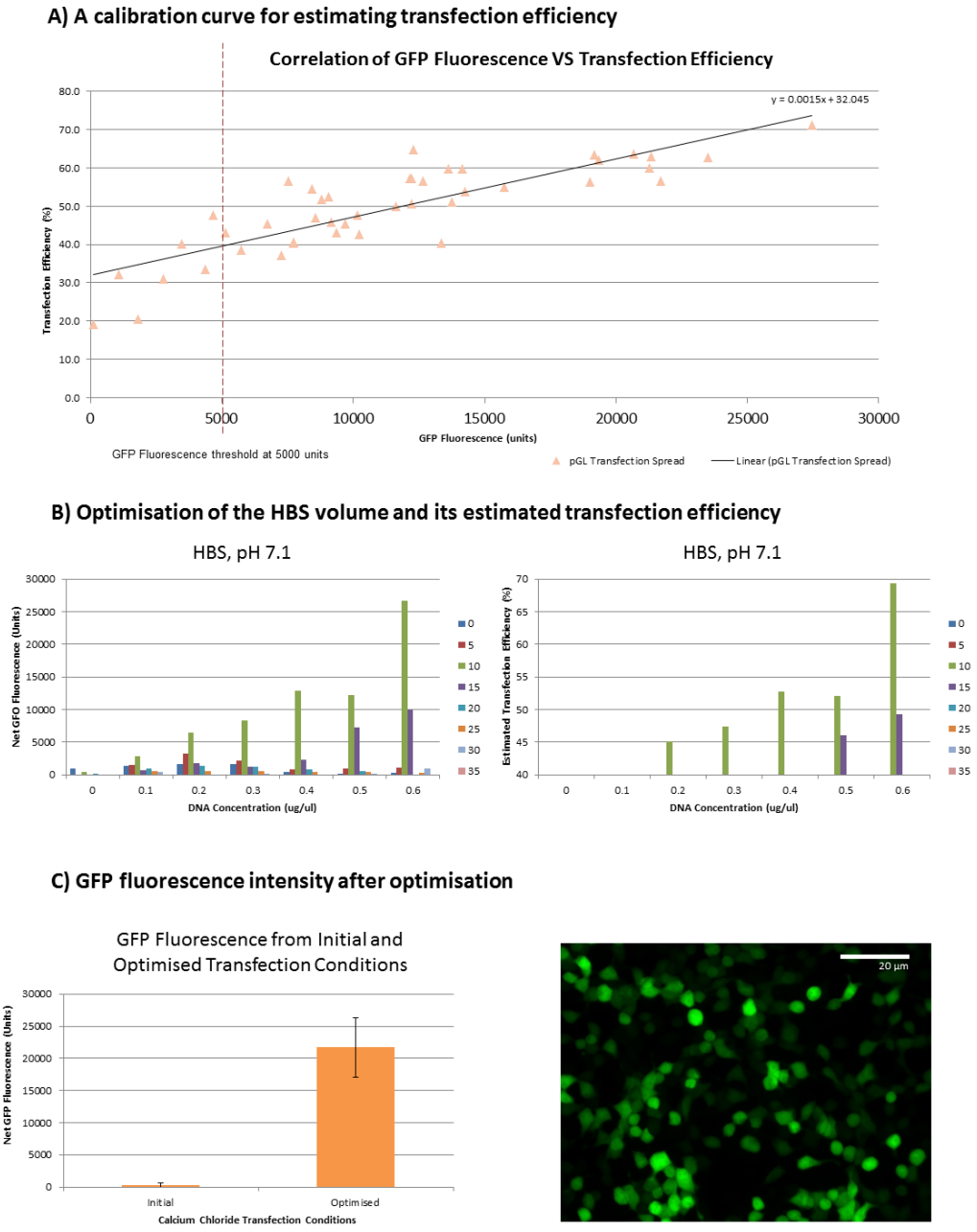


Figure 8: Optimisation of the Calcium Phosphate Transfection in HEK293T cells. A) A calibration curve between fluorescence signals generated by the FLUOstar plate reader and flow cytometry derived proportion of transfected cells was prepared for the purpose of directly estimating transfection efficiency from the plate reader data. B) Optimisation of the HBS component volume (left) and its estimated transfection efficiency (right) resulted in significant improvement to transfection. Optimal HBS volume is 10ul. C) Fluorescence intensity of GFP transfected cells as a result of the optimised calcium phosphate protocol (left) and a fluorescence microscopy image of the transfected cells. The optimised protocol resulted in an average transfection efficiency of 65%. N = 8, Magnification = 200X, scale bar = 20µm, error bars represent standard deviation of sample size.

Polyethylenimine Transfection Optimisation

Polyethylenimine (PEI) polymers of various sizes and structures were trialled in the process of setting up a protocol which uses PEI as the transfection reagent. PEI molecules with molecular weights of 2kDa (Sigma-Aldrich 408700, branched), 25kDa (Sigma-Aldrich 408727, branched) and 60kDa (Sigma-Aldrich P3143) and 25kDa (PolySciences 23966-2, linear) were tested. A stock of each polymer type was prepared in water and neutralised to pH7.0.

Transfection was optimised for 24-well and scaled down for 96-well high-throughput transfection. 1ug of pME18SFL3-GFP was diluted in 50ul of 150mM sodium chloride. Working stock with dilutions between 1 – 2,000 fold (100-fold increment) of the stock was prepared in 150mM sodium chloride. 50ul of the working stock was added to the DNA solution, vortexed and incubated for 30 minutes prior to incubation with cells. The media was then changed after 5 hours and the proportion of transfected cells and fluorescence intensity quantified with flow cytometry 24 hours post transfection. Polymer size around 25kDa was found to be most appropriate for transfection use and subjected to more precise dilutions with 10-fold increment to determine the dilution of the stock which resulted in maximum transfection efficiency and minimal toxicity.

Optimal dilution of 25kDa (Sigma-Aldrich 408727) branched PEI was 135X while the (PolySciences 23966-2) linear PEI was 300X of their respective stock, with the final concentration being 7.2ng/mL and 16.67ng/mL respectively. These final concentrations were found to give the maximum transfection efficiency with little or no observable cytotoxicity after transfection. Cytotoxicity was quantified using flow cytometry using the forward- and side-scatter (FSC-H and SSC-H) [293] together with transfection efficiency using the FL1 channel. Dead cells have lower forward-scatter and higher side-scatter. Branched PEI was found to be more toxic than linear PEI by a factor of two, as approximately twice the amount of linear polymer may be introduced to the cells without toxicity. This limitation may have contributed to the maximal transfection efficiency in the branched polymers at 59.30% ± 8.21 compared with linear polymers which consistently reaches 94.00% ± 0.56. The mean fluorescence intensity between branched and linear PEI was relatively similar at 2412.26 ± 524.93 and 2978.18 ± 89.85 respectively.

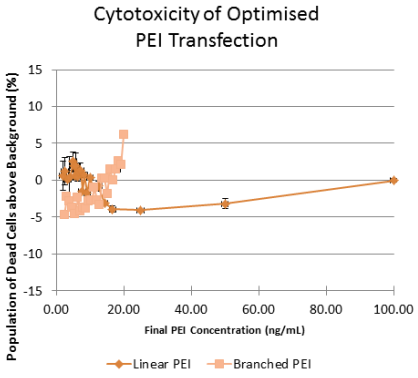
Using the optimal protocol for the branched polymer, the DNA concentration tolerance was also investigated by transfecting between 0 – 2.5ug of pME18SFL3-GFP. It was found that the amount of DNA may be reduced by a factor of two from 1.0ug to 0.5ug with the transfection efficiency reducing slightly from 59.30% ± 8.21 to 52.43% ± 0.72. Increasing the amount of DNA from 1.0ug to 1.5ug caused the transfection efficiency to decrease by two-fold to 32.05% ± 0.81, with each further

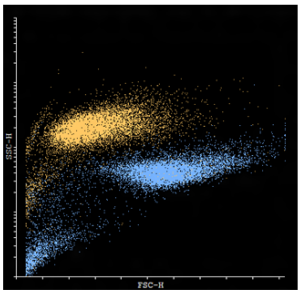
increment of 0.5ug reducing the transfection efficiency by one third. This wide tolerance of DNA concentration is crucial for high-throughput screens since the amount of plasmid DNA varies.

This stability was also observed when the pH of the linear PEI stock solution was adjusted to range from 6.8 to 7.2 with increment of 0.1. The PEI reagent resulted in similar transfection efficiency across the whole pH range. Hence pH was not a significant factor in determining the transfection efficiency. Filtering of the stock solution using 0.22um filters was also found to be important in generating higher transfection efficiency of >70%. It has been experimentally observed that this was due to the filter size removing the excessively large PEI molecules which would otherwise bind the DNA but prevents the complex uptake by the cells due to its size.

In addition to the GFP control, both optimised transfection protocols were used to transfect the positive controls RIP and caspase-2, and the reporter enzyme β-galactosidase. Apoptosis induction with the strong apoptosis inducer RIP and weak inducer caspase-2 was found to be positive, while β-galactosidase did not result in toxicity. Further characterisation of these positive controls using the optimised PEI protocols was performed using the CPRG assay.

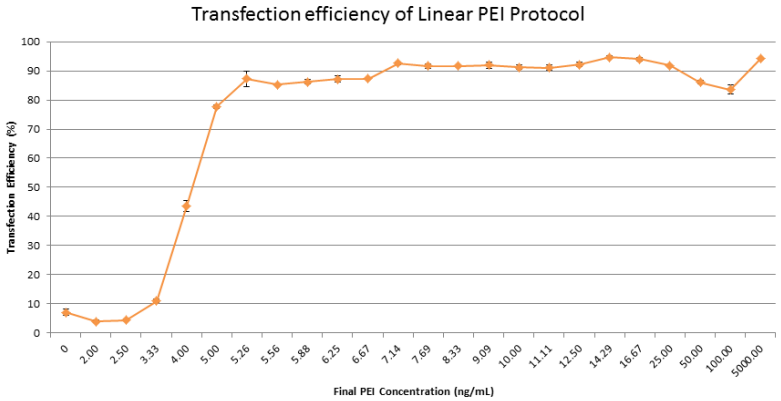
A) Low cytotoxicity of Optimised PEI transfection protocol





FSC-H and SSC-H overlay of cells transfected using 5,000ng/mL of linear PEI (orange) and untreated cells (blue).

B) PEI protocols are stable across a wide range



C) Images of pME18SFL3-based controls

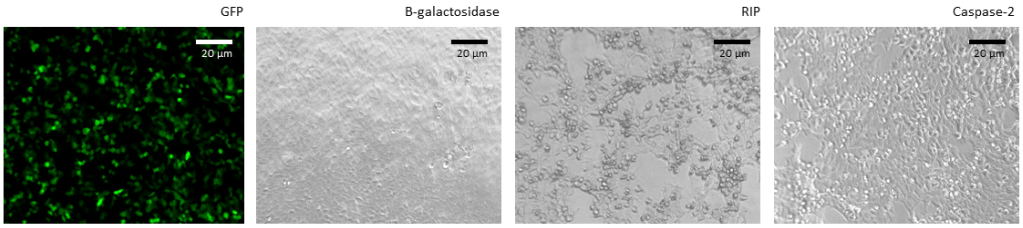


Figure 9: Optimisation of the PEI Transfection in HEK293T cells. A) Low cytotoxicity of the optimised protocols. At the concentrations investigated, both branched and linear 25kDa PEI polymers did not induce more than 5% toxicity above background (Right). Branched PEIs were found to be more toxic since they began to induce cell death above 40ng/mL. This was in contrast with the linear PEIs which only exhibited toxicity at excessive high concentration such as 5,000ng/mL (Left). N = 3, error bars represent standard deviation of sample size. B) PEI polymers exhibit an extended tolerance range for DNA complex formation, here linear PEI was shown to generate stable transfection across a wide concentration range of the polymer. N = 3. C) Images showing cells transfected with controls cloned into the pME18SFL3 library vector. Magnification = 100X, scale bar = 20µm, error bars represent standard deviation of sample size.

CPRG Assay

Characterisation of CPRG Assay Kinetics

The performance of the CPRG assay was explored using the optimised branched PEI transfection protocol. HEK293T cells were manually transfected with apoptosis positive controls RIP and Caspase-2, and negative control GFP. Between 0 – 35ul of each control plasmids derived using the BASY platform was co-transfected with 150ng of β-galactosidase reporter plasmid and the CPRG assay performed 48 hours post transfection. The kinetics of the enzyme-substrate generated signal was quantified every minute before and after lysis with triton for 60 minutes each.

Increasing the amount of plasmids co-transfected with the reporter had the effect of enhancing apoptosis activity (Figure 10A). This increased activity was quantifiable with the CPRG assay, where the strong inducer RIP reached signal equilibrium in less than 20 minutes when 25ul or more volume of plasmid was used. Increased volume of plasmids did not result in a degradation of transfection efficiency, and the Triton signals generated after lysis (including those from volumes less than 20ul) were above the threshold of 0.7.

At lower volumes of 5 and 10ul where total plasmids quantity was reduced, the signal reached equilibrium at a slower pace, and the absolute signal intensity was reduced approximately by 18.5% with every 5ul reduction in plasmids volume. In the absence of co-transfected plasmids, the concentration of reporter plasmids did not result in transfection since the Triton signal did not increase above the CPRG blank measurements. The assay reaches equilibrium within 60 minutes, after which the signal becomes stabilised.

The CPRG Ratio for each control was calculated 60 minutes after either CPRG or Triton addition (Figure 10B). When 15ul or more RIP plasmids were used, the CPRG ratio stabilised above 0.9, averaging at 0.938 ± 0.023; the CPRG did not deviate more than 2.47% from the mean. Below 15ul of plasmids, RIP CPRG ratio was reduced to 0.769 ± 0.17 and 0.686 ± 0.233 for 10 and 5ul respectively. The reduction in volume increased the deviation from the mean by up to 33.95% at 5ul. However, all CPRG ratios of RIP were above the apoptosis selection threshold of 0.55. The efficiency of caspase-2 in stimulating apoptosis were increased approximately 16.36% with every 5ul increment. 15ul of caspase-2 was the minimum volume required to initiate apoptosis with a CPRG ratio of 0.648 ± 0.173, and further reduction in volume resulted in the weaker positive control failing the selection criteria for apoptosis induction. The negative control, GFP, was stable with a mean CPRG ratio of 0.345 ± 0.061 across the whole volume range tested.

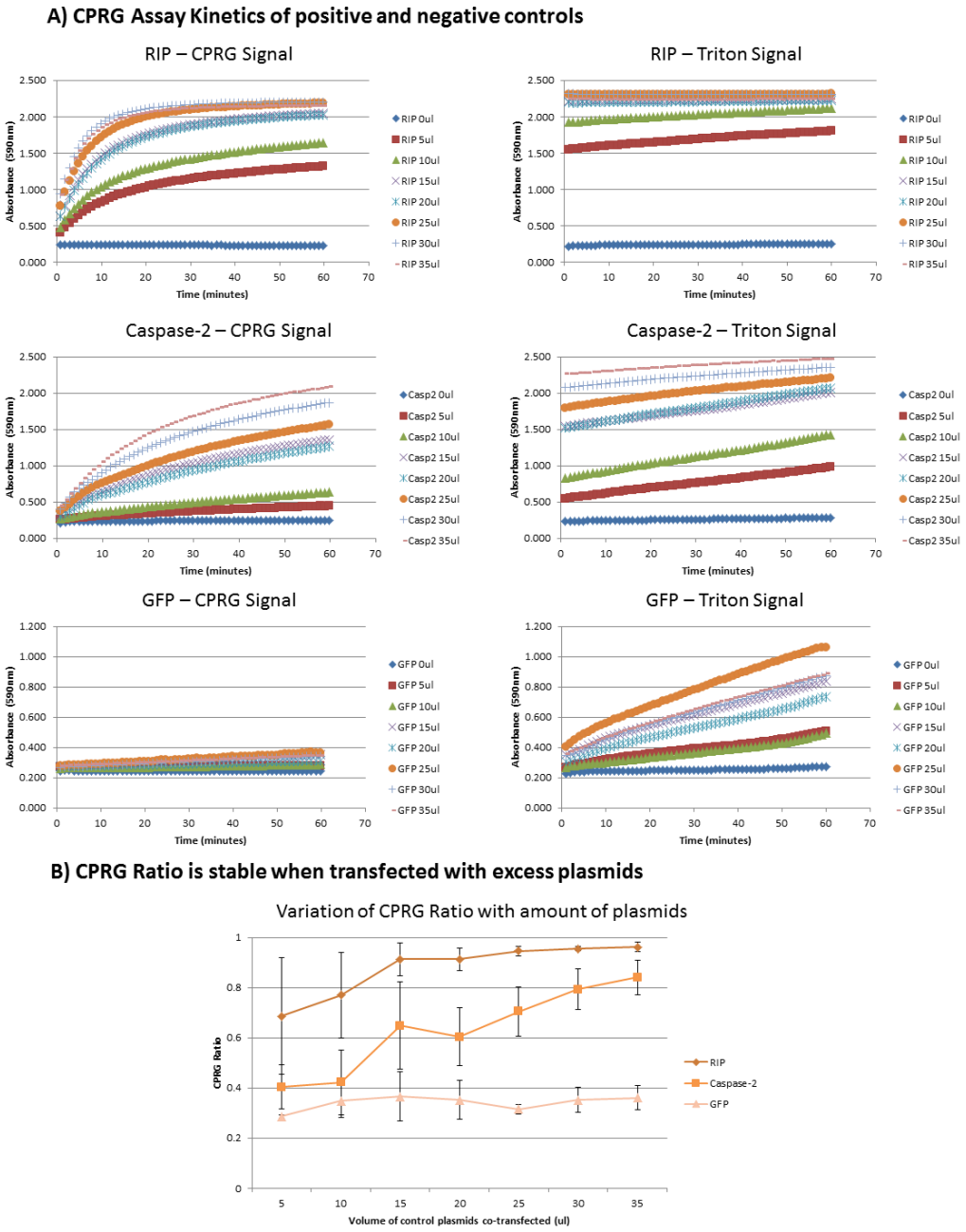


Figure 10: Characteristics of CPRG Assay Kinetics in HEK293T cells. A) Time course kinetics of absorbance 590nm signal generated before and after triton lysis of the controls. B) CPRG Ratio calculated from CPRG and Triton signals taken 60 minutes after addition of CPRG substrate or triton respectively. N = 12, error bars represent standard deviation of sample size.

Methods of calculations of Cell Death Indicators

The CPRG ratio was calculated according to Equation 1. Alternative methods of calculation (Equation 5 and Equation 6) were explored to determine if other more sensitive methods of calculations were available.

Equation 5: Ratio of dead to live cell population A590 signals

$$Cell\ Death\ Indicator\ 2 = \frac{CPRG\ Signal\ "Dead\ Population"}{(Triton\ Signal - CPRG\ Signal)\ "Live\ Population"}$$

Equation 6: Ratio of dead to live cell population based on rate of change of A590nm signals

$$Cell\ Death\ Indicator\ 3 = \frac{Rate\ of\ Change\ of\ CPRG\ Signal\ per\ unit\ time\ "Dead\ Population"}{Rate\ of\ Change\ of\ Triton\ Signal\ per\ unit\ time\ "Dead\ Population"}$$

In contrast with the CPRG ratio, which is a proportional measure of dead cell population to the total transfected population, both indicator 2 and 3 measure the amount of dead cells to live cells. Hence the CRPG ratio calculates apoptosis intensity as a percentage while indicator 2 and 3 generate the same data as fold-change of dead population relative to the live ones.

Indicator 2 uses the absolute value of the 590nm signal before the lysis as the measure of dead cell population and the net increase in 590nm signal generate by the lysis as contribution from the live cell population, combining the two measures generates a fold-change of dead population to live population. Indicator 3 generates the same fold-change ratio using the rate of change between 590nm signals. A higher rate of change before lysis indicates increased amount of cell death, while higher rates after lysis indicates a greater proportion of live cells. Under apoptotic conditions, the rate of change will be higher before the lysis, and decreases to an infinitely small value depending on the intensity of apoptosis. The reverse will be true for apoptosis negative condition. Since the assay's enzymatic reaction begins immediately after CPRG substrate addition, the rate of change (per minute) of the CPRG signal is calculated by the signal at 60 minutes less the aggregate blank measurements of CPRG substrate only, while the rate of change of the triton signal (per minute) is the difference between the Triton and CPRG after 60 minutes incubation.

The averaged CPRG ratio across all plasmids volume for RIP, caspase-2 and GFP is 0.88 ± 0.15, 0.69 ± 0.19 and 0.34 ± 0.06. The mean indicator 2 and 3 for the respective controls are 11.13 ± 42.18, 2.88 ± 2.47 and -3.03 ± 19.90, and 1.15 ± 0.34, 0.61 ±0.28 and 0.13 ± 3.24.

Whereas the CPRG ratio was able to distinguish between the positive and negative controls with little or no overlap of the standard deviation (Figure 10B), indicator 2 appears to not be able to distinguish the moderate caspase-2 control from the negative control GFP. Taking 20ul plasmid volume as the optimal volume, indicator 2 also indicated massive standard deviation at the low and high volumes. The indicator appears to be stabilised at ranges near the optimal volume where the standard deviation does not overlap between each controls. The large standard deviation also caused the mean of the strong positive control RIP to dip below the negative control when the highest volume of plasmids was used. Indicator 3 also has extensive standard deviation which overlaps across the controls at the lower range of 5 – 15ul, but remains stable after 15ul to sufficiently distinguish between each control.

The standard deviation of each ratio was calculated as a proportion of the mean to determine the extent of fluctuation between each indicator (Figure 11B) using the manually transfected CPRG assay kinetics dataset. Each indicator was calculated from a set of 12 independently generated data points. The fluctuation of the CPRG ratio tends to dissipate when the amount of positive control plasmids (hence apoptosis intensity) was increased. For example with RIP, the CPRG ratio deviated by 33.95% when only 5ul of the plasmid was used, which decreased and stabilised at under 10% when 15ul or more of plasmids was used. The positive controls RIP and caspase-2 have a similar pattern of fluctuation which tends to be higher when less of the plasmids were available. The negative control tends to have a rather stable level of fluctuation across the range. CPRG ratios of RIP, caspase-2 and GFP deviated within the range of 0.92 – 33.95%, with a mean fluctuation of 10.43% ± 12.69, 18.67% ± 8.35 and 14.85% ± 8.81 respectively.

The extent of fluctuation when the same dataset was processed into indicator 2 and 3 was much greater. For indicator 2, the three controls ranged between -9,670.05% - 101.86%, with a mean fluctuation of -163.33% ± 591.81, 59.91% ± 13.82 and -1,446.96% ± 3,632.57 for RIP, caspase-2 and GFP respectively. Indicator 2 thus has an extensive range of error for both the strong positive control RIP and the negative control GFP, with the error range of caspase-2 stabilising at around 59.91%.

Indicator 3 ranged between -8,188.37% - 3,439.82%, with the RIP, caspase-2 and GFP mean fluctuation being 20.42% ± 16.00, 31.57% ± 18.88 and -547.49% ± 3,590.61 respectively. Indicator 3 thus has a smaller fluctuation in data for the positive controls, but is inherently unstable for the negative control.

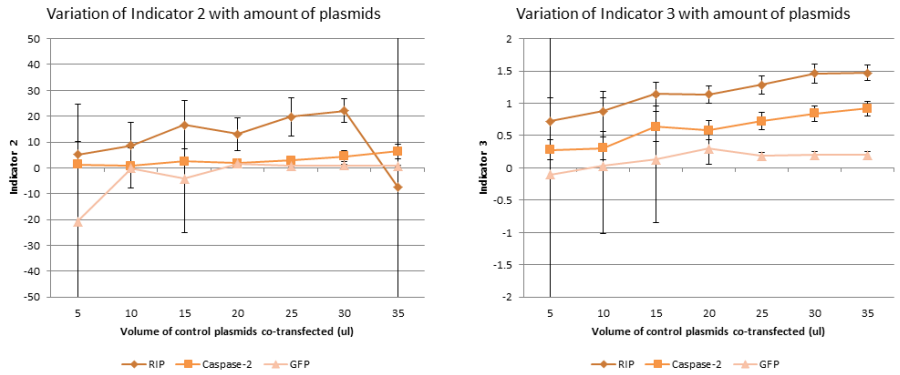
There is a tendency for both indicators 2 and 3 to fluctuate more than 1,000% or 10-fold for a single control, with indicator 3 potentially fluctuating by more than 10,000% or a 100-fold. This deeply

contrasts with the CPRG ratio which has a maximum fluctuation range of approximately 35% across the controls.

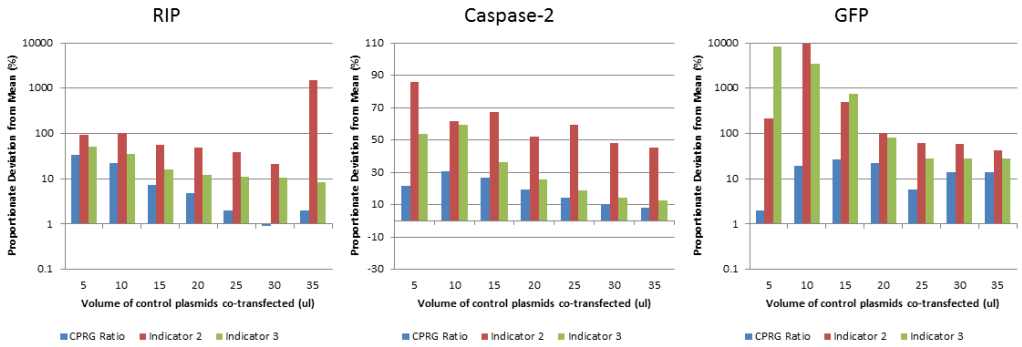
A key measure of the usefulness of the apoptosis indicator is its ability to distinguish the apoptosis positive samples from the negative samples. The t-test was employed to compare the degree of segregation of the positive controls RIP and caspase-2 from the negative control GFP, with the resulting p-value used as a gauge of the significance of each indicator (Figure 11C). The CPRG ratio is able to distinguish the positive samples from the negatives in almost all scenarios with a lower degree of error (lower p-value). Notable exceptions are denoted by asterisks in Figure 11C. For example, at the optimal 20ul volume of plasmids, all three indicators do not differ in the p-value when identifying the strong positive control RIP, while the CPRG ratio has significantly lower p-values when identifying the moderate control caspase-2. Also, Indicator 3 appears to be able to identify RIP at volumes higher than 20ul where apoptosis signal is expected to be stronger.

The 0ul volume or untransfected dataset was used as a false positive dataset since no positive controls were expected to be identified. For the false positive dataset, the CPRG ratio of false positives was calculated to be 0.91 ± 0.12, while indicator 2 and 3 were calculated as being 84.19 ± 454.90 and 11.76 ± 62.16 respectively. The calculated indicators of false positives were generally higher than the mean of RIP, since there is little change in the values before and after lysis. The CPRG ratio however presents the false positives as extremely high values, and more often exceed 1.0 which makes it convenient to eliminate in combination with the 0.7 threshold of the triton signal. False positives calculated using the other two indicators were distributed across a wide range, and often intersect the range of the positives, hence do not adequately separate them.

A) Variation of Indicator 2 & 3 with amount of plasmids co-transfected



B) Proportion deviation of the calculated indicators from mean



C) CPRG ratio provides more statistical power than the other calculated indicators

		RIP			Casp-2		
		CPRG Ratio	Indicator 2	Indicator 3	CPRG Ratio	Indicator 2	Indicator 3
Volume of control plasmids co-transfected (ul)	5	4.66E-05	0.008403	0.375629	0.005857	0.05986	0.440926
	10	0.00179	0.003334	0.009538	0.156853	0.344567	0.197405
	15	1.09E-08	2.24E-05	0.002273	0.000104	0.144809	0.0552
	20	1.01E-05	1.25E-06 (*)	5.77E-09 (*)	0.000141	0.350128	0.001907
	25	4.63E-24	2.62E-09	2.47E-13	7.49E-09	0.000456	2.08E-09 (*)
	30	1.22E-09	0.40205	1.06E-13 (*)	5.53E-12	4.56E-05	1.19E-11
	35	2.06E-10	0.40205	5.44E-16 (*)	3.4E-13	1.55E-05	9.41E-13

Figure 11: Comparison between different calculations of indicators. A) Calculated indicators 2 and 3 using the HEK293T kinetics dataset from the previous section. Both indicators measure the dead : live ratio as opposed to the dead : total by the CPRG ratio. N = 12, error bars represent standard deviation of sample size. B) The proportion of fluctuation of each indicator calculated from the same dataset. The CPRG ratio shows the lowest amount of fluctuation in all scenarios investigated. Charts for RIP and GFP were expressed in the Log_{10} scale. C) Statistical power of the CPRG ratio in resolving positive controls from negative controls. The CPRG ratio has the lowest t-test statistical probability of random error in all except five cases. These exceptions are denoted by asterisks (*).

Reproducibility of the CPRG Assay on the Robotic Platforms

A fully automated DNA isolation and transfection protocol was performed and the reproducibility of the CPRG assay assessed in a similar manner as described in the previous section Methods of calculations of Cell Death Indicators. On a 96-well plate setup, which was performed in duplicates by the robots, 91/96 (94.79%) of the wells did not generate CPRG ratio that deviated more than 10% from the mean.

In a separate experiment, bacterial cells containing the control plasmids of RIP and luciferase were inoculated in alternate wells of a DWP, and subjected to the automated DNA isolation and transfection process. Here, 45/48 (93.75%) of the positive control RIP were correctly identified using a CPRG threshold of 0.65, while 2/48 (4.17%) of the luciferase were incorrectly identified as positive.

The fully automated DNA isolation and transfection processes were hence able to generate reproducibility similar or better than manually performed protocols.

Modifications of the DNA Isolation Platform

Technical improvement to robots

General maintenance of the robots included the routine exchange of pipette tips and air-channelling tubes, sterilisation of the wash station and secondary water tank and general cleaning and sterilisation of the platform with ethanol.

With each exchange of the pipette tips, the exact pipetting height at each destination position was altered. This required readjustment and testing of the heights after each exchange for which the software was adapted for increased efficiency. A spacer was also incorporated to align each pipette tips to the same height, increasing the pipetting efficiency.

The water for cleansing the pipette tips was distributed between the two robots through a central 10L tank, which is manually refilled using deionised water from a dedicated tap. The water pressure drives the water into the tank, located approximately a metre above the robots, and distributes to the secondary tanks of both robots using gravity. The small tank size requires frequent refills and hinders the automated process with “low water” warning. This setup was replaced with an automated pump supplied with double-distilled water from the Millipore RiOs^{TM} 5 and an accompanying 30L reservoir.

The aging computers of both robots were also upgraded from those running Microsoft Windows Millennium Edition to Microsoft Windows XP powered by Intel® Core2 processors, and adjustments

to the software in line with the system upgrades. The robotic centrifuge was also replaced with a similar centrifuge (Sigma 4K15).

Engineering support and maintenance on the electrical circuitries was routinely provided by Dr Volker Kachel, and a team of electrical engineers based at the Imperial College Hammersmith Campus.

Changes to the automated DNA Isolation protocol

LB^+ was used as the bacterial culture media instead of standard LB media, which resulted in increased culture density and plasmid yield. The use of the gas permeable seal to close the DWP for incubation allowed the culture volume to be increased from 1.0 to 1.4mL while improving aeration and preventing contamination.

The volumes of the reagents used in the DNA isolation protocol was increased approximately 47%, which enhanced the supernatant transfer between steps and reducing the loss of supernatant volume to "dead volumes" which cannot be pipetted by a robotic system.

A semi-automated protocol was developed such that the robot manages major pipetting steps while repetitive steps such as individual transfer of DWP between platform and centrifuge, and cyclic removal of waste media or supernatant were manually performed. This reduced the total duration of a single run by 1.5 hours, enabling two to three runs to be scheduled daily. This in turn translated to a two- or three-fold increase in daily throughput.

Discussion

The NITE cDNA library was provided as plasmid DNA, which required the transformation into a bacterial host for replication and production of the plasmids. Scaling an experiment towards high-throughput levels create various sorts of technical and logistical hurdles, even for a protocol as simple as heat-shock transformation.

E. coli DH5α strain (Invitrogen) was selected as the primary bacterial host as it is a strain commonly used for cDNA library cloning, with advantageous mutations in genes like endA1 and recA which reduce non-specific endonuclease activity and DNA recombination events, hence offer an overall higher plasmid yield. However, with 30,000 different transformation reactions, 2.5 to 3.0L of competent cells would be required, the use of commercially available library-efficient E. coli

DH5αTM (Invitrogen) would be cost prohibitive at GBP 180,000. Consumables costs such as pipette tips would also be staggering.

The only possible solution to culturing the required 30L of bacterial cells was the use of a bioreactor, since the amount of flasks required would be difficult to manage and process due to the lack of incubator space. The self-assembled bioreactor was designed with a number of key features, including a constant air supply, continuous suspension of the culture with a magnetic stirrer, stable incubation temperature, outlets for waste media sterilisation and sampling. Sterility is the primary concern since no antibiotics could be used. The bioreactor was designed to remain closed from the external environment throughout incubation, with regular sampling achieved through a sampling tube.

It was discovered that the rate of growth in the bioreactor was 0.0445 units/hour and 0.1567 units/hour for the standard flask conditions. Hence the growth rate in the bioreactor was reduced by a factor of 3.5 compared to the normal culture conditions. This is likely to be the result of inadequate aeration and mixing of the 5L culture volume, which was supplied through a dedicated pump at 0.1 bar pressure. The tubing supplying the air was optimally perforated for improved efficiency of aeration with smaller and higher counts of bubbles generated. Further perforation would decrease air flow due to insufficient pressure while increasing the pressure leads to turbulence within the culture content which disrupts the mixing and excessive overflow. Similarly, the mixing was operated at maximal revolution as further increase would have caused the magnetic stirrer to slip, halting the mixing.

However, competent cells derived from either culture conditions could both be efficiently transformed and resulted in similar colony counts. This is expected since both cultures were harvested at similar optical density indicative of similar bacterial cell count, therefore, the rate of growth is not a significant indicator on the quality of competent cells as the culture density. Twice the amount of starter culture volume (100mL) was used to inoculate the bioreactor such that it reached the appropriate culture density within the same timeframe.

Competent cells were prepared using a single reagent, calcium chloride, which greatly reduced the complexity of the preparation process. Each batch was subjected to strict control procedures where samples were tested for antibiotics resistance. The prepared competent cells were shown to be able to produce high plasmid yields with no major sequence recombination events or contamination observed.

Automation was required for the pipetting steps to handle repetitive functions and reduce process time and consumable costs. The DNA isolation robot was selected as the platform of choice to avoid bacterial contamination on the transfection robot. The implementation of the new transformation protocol required the development of new software, since the BASY robot was hardcoded to only perform the DNA isolation protocol.

Contamination was again the main concern since the platform was not designed to handle live cultures in a sterile environment. The control experiments performed indicated that the use of the new ethanol rinse function and increased wash cycles could sufficiently prevent cross-contamination between plates. Some unintended growth of bacterial were observed, which could be attributed to the movement of the pipetting arm as it moves over the plates for inoculation. This source of contamination is expected to result from a small number of cells, which would be a negligible background as the inoculant is a dense starter culture. The denser culture is expected to exponentially outgrow any background and form the bulk of the culture. Prior to implementing the high-throughput transformation, the workflow was tested on eight plates of cDNA. 143/160 clones were matched with the predicted restriction pattern of pME18SFL3-based clones, while the remaining 17 were matched to those cloned into pUC19FL3 which forms a minority (<1%) of the library. This indicated that contamination across the platform and process was sufficiently low to allow for sequential processing of bacterial plates without cross-contamination.

The primary round of transformation was completed within six weeks with a success rate of 73.68%. The remaining clones were manually transformed and completed within two weeks. These clones failed the primary round due to various reasons, most often due to the lack of contact between the DNA and cells since the solution was ejected from a distant. Contaminated batches were also identified and rejected using the quality control procedures, and were likely to be due to the intensive operation schedule or blocked pipette tips.

Two types of transfection protocols were investigated and optimised for use on the transfection platform. The first, using calcium phosphate, was a common method of transfection employed for HEK293T cells transfection. The initial conditions used in a previous screen were found to be ineffective for the plasmid concentration generated by the robot. Previously, this was not a concern since the previous screen setup uses a normalised and non-sequenced cDNA library which was actively cloned and cultured at high dilution prior to DNA isolation and screening; any untransfected cDNA were basically excluded. With the NITE library, each well contains a unique plasmid which is variably replicated by the host, resulting in a range of DNA concentration. This DNA concentration range introduced instability in the transfection protocol.

During one experiment, 15 µl of each component as part of the transfection mix resulted in a massive shift in the detected fluorescence from insignificant background autofluorescence to a detectable increase of the GFP reporter. This volume ratio formed the basis for further optimisation.

A calibration curve for the estimation of transfection efficiency from fluorescence was prepared in order to avoid the need to optimise in 24-well before scaling down, since this was not possible with the initial conditions. The correlation was linear for fluorescence signals above 5,000 units but rapidly slopes off to zero, so this region was not suitable for use to estimate transfection efficiency. Generally, higher fluorescence indicates a better transfection.

Incubation periods appeared to be suitable only for higher DNA concentrations of more than 1 µg/µl (data not shown), so the incubation was eliminated from the optimisation protocol since the estimated plasmid yield was 0.3 µg/15µl.

Variations of individual component volumes showed that a further 5 µl reduction in HBS was able to improve the transfection efficiency, after which changing the volume of calcium chloride and chloroquine did not produce further significant improvements. The optimised protocol for calcium phosphate transfection was set as 0.3µg plasmids in 15 µl water, 15 µl calcium chloride, 10 µl HBS and 15 µl chloroquine. 60% cell confluency with 14,000 – 20,000 HEK293T cells per well was also found to give the best transfection. The optimised protocol resulted in significantly higher levels of transfection as determined by the GFP fluorescence intensity, and is proportionally scalable to 24-well and 6-well plate formats and the quantification of GFP with flow cytometry had consistently indicated transfection efficiency between 70 – 80% (data not shown).

However, it was impossible to implement the optimised calcium phosphate protocol to the robotic platform. The reason has yet to be determined, but it was theorised that the residual water droplets remaining on the pipette tips after each wash cycle eventually dilutes the transfection mix and reduced transfection efficiency.

The alternative polyethylamine (PEI) based protocol was selected for its cost-efficiency and simplicity in transfection. The protocol uses a single PEI reagent compared to three components in the calcium phosphate, making it faster and simpler to automate. Furthermore, it was found during the optimisation that both the branched and linear PEI protocols were very stable to DNA concentration variation. This is advantageous since the apoptosis modulating capacity of a potential candidate gene may become reduced but remains detectable. In contrast, the calcium phosphate protocol would result in no transfection once the DNA concentration is beyond the precise optimal point.

The branched PEI polymer (Sigma-Aldrich 408727) was initially found to be the best commercially available polymer for high-throughput transfection. It reached a maximum transfection of 60%, but the transfected cells generated strong protein expression indicated by the high fluorescence intensity of 2,412 units using flow cytometry. As the CPRG assay was normalised to include only signal from transfected cells, it was not necessary for the transfection to reach 100% efficiency. Moreover, the stability of the protocol was evident as the branched PEI protocol was adapted onto the transfection platform and used during the primary screen.

It has since become widely established that PEI-based protocols to be excellent for high-throughput applications and even in vivo transfection [294, 295]. The linear form of PEI polymer has since become available commercially. The linear PEI was more efficient and routinely reaches near complete transfection of cells with strong fluorescence intensity of 2,978 units. It was also less toxic than the branched form as the cells could tolerate twice the active concentration of the linear polymer. This higher tolerance may explain why the linear polymer could reach near complete transfection in HEK293T cells. The linear PEI was employed in the secondary screen.

The stability of the PEI transfection was confirmed in the kinetics studies of the CPRG assay, where a range of plasmid volumes was used with a single set of transfection conditions. Reducing the active plasmid component caused the apoptosis intensity to be reduced, but for a strong inducer like RIP, this did not cause the clone to fail the selection criteria test. However, weaker inducers such as caspase-2 could have its effects reduced beyond the positive selection threshold when lesser plasmids were used. The assay reached equilibrium 60 minutes after introduction of either the substrate or Triton, providing a stable signal, which could be measured.

The use of the CPRG ratio as the basis for assessing apoptosis activity was also optimal as this ratio could sufficiently distinguish the negative and positive controls. When compared to fold-change indicators using either the signal intensity or rate of change of these signals to quantify the dead : live ratio, the CPRG ratio offered greater statistical power in distinguishing the controls. This is because the CPRG ratio represents apoptosis as a proportion of dead cells to the total transfected population, while the other indicators present the same information as magnitudes of change. Fold-change indicators give a higher value for a positive control, but tend to fluctuate more.

Furthermore, the false positives as measured by the CPRG ratio always have a ratio which is extremely near or above 1.0. This is the consequence of the signals not changing due to the lack of transfection, and the slight dilution in Triton signal after the addition of the detergent which increased the total assay volume by approximately 10%. As a result, the false positives are presented

within a small range by the CPRG assay. Both fold-change indicators present the same false positives across a wide range, which frequently overlap with the candidate selection range. False positives may generally be excluded from the analysis through the application of a 0.7 threshold value for the Triton signal (Discussed in Chapter7: Implementation of the Screen).

The deviation of the CPRG ratio increases when lower volumes of plasmids were co-transfected. This reduced amount of plasmids resulted in lower protein expression, which causes the gene to miss the threshold for apoptosis activation. The effect was particularly acute for the weaker controls, which could miss the threshold by up to 50%, leading to the wide standard deviation. This effect of increasing error range is a familiar feature of high-throughput assays [296].

The error range of the assay when manually performed is approximately 33.95%, which reduced to less than 20% when automated on the robotic platforms. The test trial using the optimal DNA isolation, transfection and assay conditions on the robotic platforms correctly identified 93.75% of the positive controls, with a 4.17% error rate where the negative control was identified as positive.

Since the effects of the over-expressed genes were expected to improve with higher DNA concentration, 20ul of the plasmids was selected at the volume of plasmids to be co-transfected with the reporter during a screen. The CPRG ratio was statistically capable of resolving the controls at this volume, while retaining sufficient volumes for up to four replicates to be performed from a single plate of DNA plasmids.

The DNA isolation robot comprises various components including a centrifuge, reagent reservoirs, freezer positions, acetone dispensing chamber and a tip washing module. The complexity of the robot was initially preventing a fully automated process for the isolation of DNA, since the lack of position sensors occasionally resulted in a system crash when a DWP was positioned out of alignment after centrifugation. Such simple problems were often interrupting the process, and in serious cases caused the movable pipetting arm to physically crash into the equipment.

With the development of new software, a semi-automated protocol was developed where the robot manages the liquid handling functions while time exhaustive cycles such as the removal of media or transfer of DWP between centrifuge and platform positions were performed manually. This modification reduced the time spent on repetitive cycles, cutting the total process time from 4.5 hours to less than 3 hours. This represents a 30% increase in efficiency, allowing two DNA isolation runs to be comfortably performed daily, while an additional run could be performed on demand.

Additional problems included low pressure developing in the main water tank, creating a “vacuum” effect that prevented water from refilling the secondary tank. While an isolation valve was initially introduced to manually correct the pressure difference, the manual setup for tank refill was replaced with the Millipore RiOsTM 5 with uses an electric pump to fill the tank from an accompanying 30L reservoir. This resolved the scenario where the system halts the process as it awaits tank refill.

Pipette tips were manually loaded and could be reused after cleaning by a dedicated washing module, to reduce consumable costs. This manual loading of tips, however, was highly inconsistent resulting in downstream pipetting errors. The tips could also dislodge from the pipettor, becoming trapped in the deep well plates on the pipettor, which then crashed into other components if unsupervised. A solution was founded by first assembling the 96 tips over a plastic film and a spacing guide before this tip assembly was loaded onto the robot pipettor. This arrangement allowed all 96 tips to sit on a similar level thus improving liquid handling accuracy while the plastic film prevented rogue tips from dislodging. The incorporation of a new ethanol wash step keeps the system sterile while handling live bacterial culture.

The working coordinates of the robot pipettor on the platform were also important, and incorrect values could introduce pipetting errors, carry over contaminants, or resulted in physical crashes. These coordinates and pipetting height were altered and optimised following each reassembly of the main pipettor.

Reagent volumes were also increased by 47% to minimise the loss of supernatant between transfer steps due to dead volumes. This resolved a critical point in the automated system where the pipette is unable to pipette all the volume from a plate, with increased the final plasmid yields.

After this initial optimisation work, a stable run with consistency in the plasmids yields between each well was achieved on the robot platform. This is important since the DNA concentration is not quantified for the downstream transfection as the robots are unable to pipette individual wells. Concentrations differ slightly as individual pipette position had a minuscule level of error. Average yield per well was determined to be around 0.3 µg/15µl. Overall, 95/96 wells clearly indicated the presence of plasmids and the system was ready for automation of plasmids isolation.

Chapter 5: Software Development

Background Information

Specialised software is an integral part of automation, working in tandem with the robotic hardware to process the workflow. The uniqueness of the RICSI platform and its hardcoded nature necessitate the continual development of its software in order to implement new workflows. The sheer scale of our high-throughput experiments, which resulted in over 120,000 data points after a single screen, further required the need of software to process the data.

The software was designed to be robust and user-friendly, allowing for user involvement to be minimised. This reduction of user involvement in the screen is an important factor in improving the workflow efficiency, and the software for the robots and data handling were designed with a simplified interface, and to setup and execute the systems in a few clicks, and provide live system status such as current process and platform position. The ultimate objective was to minimise user programming with the implementation of each new workflow, making the system also accessible to users with minimal programming knowledge.

The software for the RISCI platform was originally developed in HP VEE 4.0, and has since been upgraded to Agilent VEE 9.0. The VEE (Visual Engineering Environment) was a proprietary package developed by Hewlett-Packard for automated measurements, data acquisition, reporting, instrument control and virtual instrumentation, and became the product line of Agilent Technologies when the company was spun-off. VEE is a graphical programming language which offers an integrated software development environment. In contrast with development languages such as C or Java which describe the functions as text, VEE represents the operation flows and software architecture graphically. VEE objects and functions are represented as boxes with various types of sequence, execution and dataflow pins. Operation sequence flows vertically from top to bottom through a single sequence pin, while data flows horizontally from left to right through one or more input/output pins. This graphical interface along with the dynamic IO library helped to speed up the development for the RISCI software and simplified the development of the software graphical user interface (GUI).

Automation of data processing was performed primarily using Microsoft Visual Basics, since the data was captured into a Microsoft Excel® spreadsheet by the FLUOstar plate reader. Integration with the

Microsoft Office Suite® also provided the various charting and presentation functions, while its extensive library of mathematical and statistical functions aided in the calculations.

Results and Discussion

RISCI Software Development

Early stable version and minor adaptations

Since the BASY DNA Isolation Platform was originally developed for a single purpose of processing the ultra-pure DNA isolation protocol, its software was designed to be rigid and initiated with a single activation button. The innate structure was designed to avoid user involvement on the programming aspect, restricting user involvement to load and run functions where the reagents and plates were setup at pre-defined locations after which the robot fully automates all downstream steps. This single protocol was extensive with a runtime of 4.5 hours and comprised of more than 100 individual steps by simultaneously processing two interlaced workflows with four plates each. It also featured a GUI panel providing status updates as the robot processed the protocol. The entire protocol was hardcoded into a single program, and its complexity with hundreds of sequence and data flow lines linking various functions and objects was a major impediment for further optimisation or minor changes.

However, minor changes to platform coordinates and pipetting height were routinely required since changing to a new set of pipette tips may affect previously optimal coordinates. Implementing such routine updates was difficult and tedious as each set of coordinates were coded within subroutines of individual steps. Optimisation of the DNA isolation protocol also identified certain limiting steps, which were reducing the final plasmid yield; for example, the small reagent volume used resulted in a higher loss of supernatant between each transfer steps. Potential improvements were hence difficult to implement without disassembling the hardcoded workflow, while any change needed to be followed up with relevant changes to the status update GUI which was also hardcoded.

Furthermore, a major shortcoming of the BASY platform was identified during optimisation runs, in that the robot has a higher probability of suffering a physical system crash compared with the TFA transfection platform, a consequence of its ability to transport plates between positions and the centrifuge using the robotic gripper. The plate transfer function was built on the assumption that each plate does not move at its position prior to transfer, an assumption which is not completely accurate since the capacity of the shaking platform to homogenise the solution and centrifugation

steps may cause the plates to move slightly. These slight shifts in position may cause the gripper to place the plates in an inclined position, which cannot be detected by the robot due to a lack of positioning sensors, and downstream attempts by the robot pipette to pipette from the incorrectly positioned plate would result in a serious crash. Such crashes can be dangerous since the system cannot automatically sense the damage and continue to process its workflow in the absence of the user. The hardcoded nature of the software further prevented the continuation of the workflow when the system was temporarily stopped to correct errors or major crashes, potentially wasting the entire run. This inherited software design was appropriate for the previous screen design, which did not require every well position to be transfected due to the randomised nature of the cDNA library.

For a structured and sequenced collection like the NITE library, the position of each well becomes important and necessitates the optimisation of conditions. The changes made for the implementation of the optimised experiment settings were minimised to maintain the software architecture integrity. Changes such as the increase in volumes and changes in the protocol were made by modifying the software and similarly hardcoded. A number of new functions were created to facilitate the routine coordinate optimisation. This include a height increment function, whereby a loop structure was used to repeatedly increase the pipetting height in minute steps until the optimal height is obtained, after which it returns the numerical value. A similar function was made for optimising the x-y coordinates for each plate position. The function for performing the pipetting steps was also improved by introducing a two-step height change function. The robotic pipettor was programmed to first stop above the plate allowing for the ends of each pipette tip to be positioned in each well, followed by the second height change to the final pipetting height for liquid handling. This change allowed the plates to be “repositioned” by the robot and slows the approach speed, both of which greatly reduced the chance of the robot crashing into an incorrectly aligned plate.

To enable the process to restart from the point of termination, each step of the protocol was reorganised into an individual function and sequentially labelled. These functions were then placed under the control of a for-next loop structure, which the initiation variable value being the step to restart from. This allowed runs with minor errors to be salvaged.

The optimised conditions for the DNA isolation protocol automated by the BASY platform was validated to produce a stable and equal quantity of plasmid DNA in each well, and the introduction of new functions enhanced the ease of using the platform.

Flexible and user-customisable version

The need to harness the BASY platform for the automation of high-throughput bacterial transformation protocol, coupled with the first major upgrade of the development environment from HP VEE4.0 to Agilent VEE 9.0 dictated the development of new platform software architecture. The objective was to open the robotic platform and allowed for automation of a variety of user-generated protocols, with the software version codenamed as “ProjectFlex”.

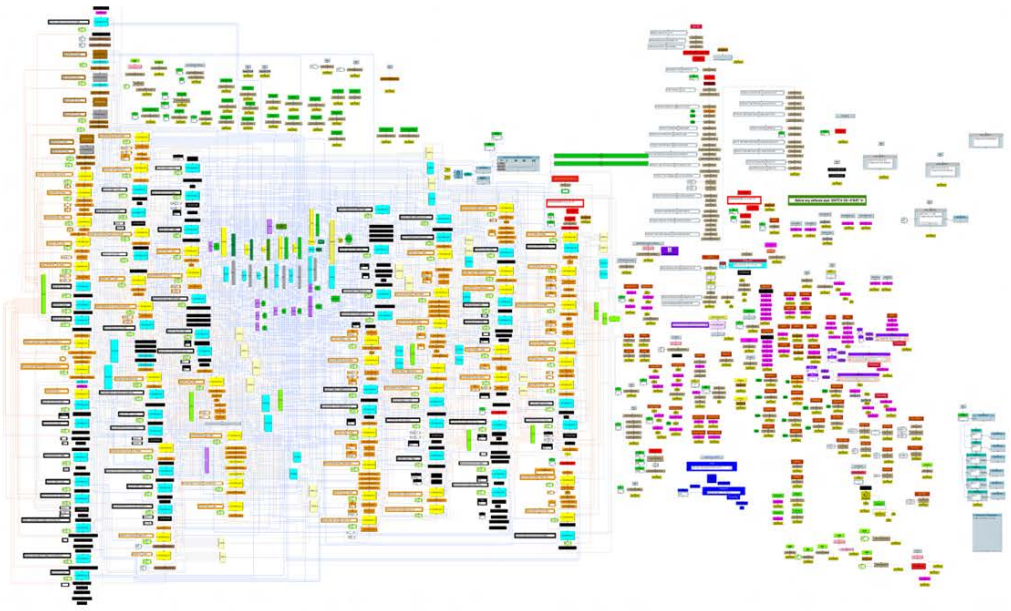
Taking a simplistic view, the robots are controlled by electrical bits in a binary fashion, each of which regulate a specific action such as the opening of a valve, and a set of bits was orchestrated for the robot to perform a particular action such as pipetting or washing. The robotic centrifuge was controlled via a serial interface by sending ASCII commands. The paradigm shift to a fluid structure required all hardcoded functions to be replaced by variable-controlled independent functions, and the organisation of each set of electrical bits as a robotic action rather than a step of a particular protocol.

These independent functions were developed according to five main classifications: administrative, basic, actions, creation and error handling. Administrative-type functions control the background process which generally do not translate into a physical motion, such as the variable declaration, position offsetting and retrieval of status information for display. Basic functions were the minimal possible organisation of the bits to perform a robotic motion. These include robot actions for pipetting in and out solutions, activation of gripper, wash station pumps and vessel lid motions. Actions-type functions defined routinely used robotic processes including dispensing of solution, transferring plates between positions, cleaning of the pipette tips and pipette mixing motion to homogenise solutions. Creation-type functions manage the user interface for assisting the user in building and testing of the custom workflows. Actions and Creation type functions are high level functions built from the Basic functions. For example, the “Transfer DWP” Actions function consists of sequential implementation of the following Basic functions: “Go To Platform”, “Grip DWP”, “Go To Platform” and “Release DWP”. Error-handling functions check for system error such as low pressure, low water levels, disengaged platform doors, incorrect motions for detecting jammed lids and shaking platforms and other safety features. Each of these functions required a specific set of variables to work, for example the pipetting functions required input for volume and speed. The use of variables enabled settings to be actively defined.

Under the ProjectFlex development, the program operating the robots required two user-defined text files for coordinates and protocol. The coordinates file defined the name and x-y-z coordinates of all platform positions, which were previously hardcoded into the DNA isolation workflow. This

coordinates file offers many advantages, such as allowing the user to quickly update the coordinates following optimisation and even full-scale reorganisation of the platform. It is defined once and automatically loaded by the software during system initialisation, with each position declared as a variable holding coordinates. The robotic pipettor may be moved to a defined position using the position name, instead of all three coordinates previously.

A) Hardcoded workflow for DNA isolation



B) Program for flexible implementation of user-generated workflows

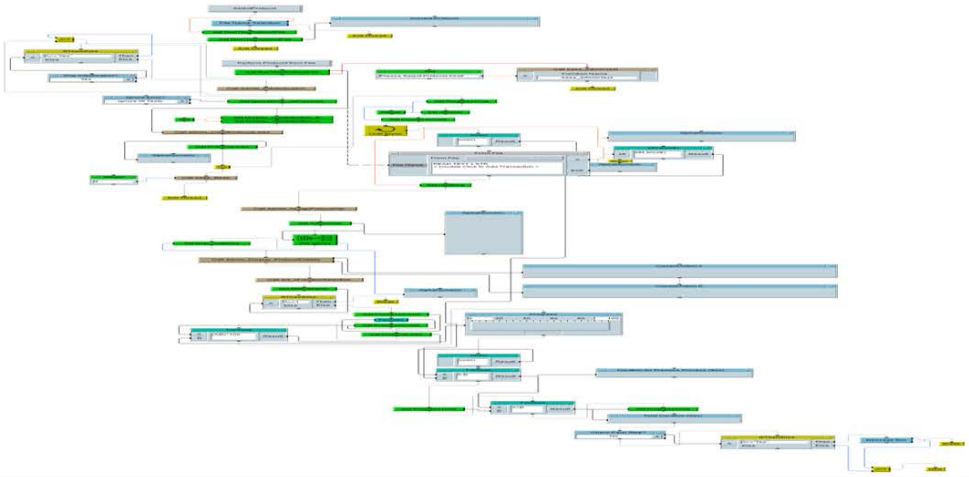


Figure 12: Original BASY and ProjectFlex core codes. A) The original BASY program was hardcoded, resulting in an extensive 4.5 hour workflow comprised of more than 100 steps. This version was incredibly difficult to modify and cannot be used to implement any other protocol. B) ProjectFlex was designed to be flexible, with the core code simply reading off the user-defined protocols and translating the instruction into robotic actions.

The protocol file defines the entire process required by the user in a step-by-step format. The user may define as many protocol files as required and select the appropriate one prior to execution. Each protocol file may be manually written in a text (.txt) file or build using the Creation functions. The Creation functions are organised into a single Panel, and on activation, keep the program in a loop allowing the user to add or delete each step, and even test the protocol, until the desired process is achieved after which the protocol is converted into a permanent text file. Each line within the protocol file refers to an individual step with all its variables delimited by a semicolon ";". The first five variables define general information such as the step count and status description, the sixth variable defines the function name to execute, with all remaining variables attributed to those required by the specific function. An example of a single instruction is given as:

7;A;2;a;Cleaning the pipette tips;act_WashTipsN;3;

Where "7" refers to the seventh step of the protocol, "A" refers to the processing of the first of two interlaced protocol, "2" and "a" are placeholder variables reserved for future use, "Cleaning the pipette tips" describe the step and is reflected in the status information, "act_WashTipsN" refers to the wash function to execute, and "3" refers to the cycle count required by the wash function which in turns execute the function repeatedly thrice.

The core algorithm (summarised in Figure 13) for processing of the protocol file works similarly to the Creation function. The system was first initiated whereby the robot is required to align itself with and contact a reference pin; its height was next incrementally changed to the point where reference signal becomes negated and the z value used to reference and offset any variation in the vertical pipetting height. The system initiation function was inherited from the previous software and placed under the control of a Boolean variable to enable it to be bypassed as required. All other platform bit signals controlling valves, lids and action activation would also be reset to the default settings. Next, the protocol file was processed and the total number of steps counted and used to estimate the protocol progress. The function was then kept active through the use of a "Do until break" loop structure, which repeated its subroutine for processing each protocol step. The time prior to and at the end of each step was recorded and its difference used to gauge the step duration and total process duration. The instruction at each line was then sequentially read and processed if an instruction was available; this instruction will be split using the semicolon delimiter into its component variables, which are then used to update the status information panel and execute the appropriate function. If the protocol has reached the end of the file, known as "End of line", the break function was activated to allow the loop to end.

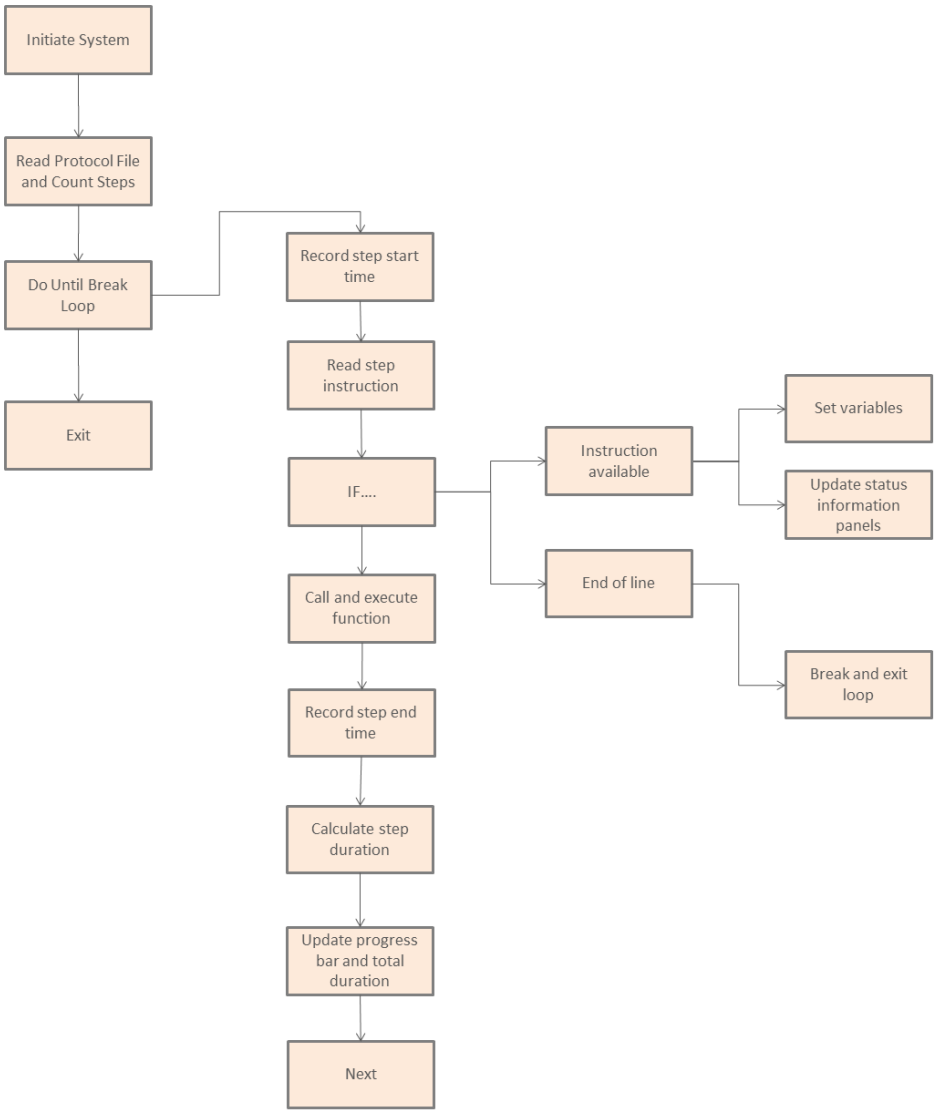


Figure 13: Summary of the core algorithm for flexible processing of user-defined protocol. The BASY platform was opened up for user-defined protocol by organising the robots into action functions, which could then be called and executed in any combination as defined by the user in a protocol file. The core algorithm was a reduced from a complex hardcoded structure to a simple loop structure, which continually reads and executes each instruction until the end of the protocol file.

Due to the flexible nature of ProjectFlex, additional set of checks and safety functions were added to the previous set of error handling functions. The new safety features focused on the platform actions, for example, the “safety height” function, which changes the robotic pipettor to a safe z-coordinate was built into all functions involving the movement of the robot. Also, each position was checked to be correct and the shaking features inactivated before transfer of plate. These safety features ensured that any combination of user-defined protocols could be performed without causing damage to the robot, but it also introduced a greater degree of redundancy which is a drawback of any flexible program. All error-handling functions were compiled into a single management panel, enabling selective activation of each function. As ProjectFlex was developed with the intention for the implementation of the high-throughput bacterial transformation workflow, new functions such as the ethanol wash cycles and media dispensing were also introduced to handle live culture work.

The new program proved to be useful in its deployment for high-throughput transformation. Various pre-processing protocols were defined. For example, one protocol was defined for the robot to slowly pierce the foil-sealed 96-well PCR plates containing the lyophilised plasmid DNA, while another was defined to add sterile water and resuspend the plasmids.

The transformation workflow was implemented in five phases: 1) dispensing of 100ul competent bacterial cells into DWP, 2) introduction of DNA and heat shock, 3) dispensing of selection media after cell recovery, 4) dispensing of 1mL fresh selection media and inoculation of these new plates and 5) preparation of library glycerol stocks. The protocol files were defined for processing of one and eight plates and the entire process of defining these protocol files was completed in less than an hour. Switching between protocols was also extremely quick and easy, with the user simply selecting the desired protocol files and hitting the “Perform Protocol” button.

Its versatility eventually made it the default choice for automation of routine high-throughput liquid handling functions and also for troubleshooting and optimisation when checking for protocol inefficiencies.

Update of the DNA Isolation Process on the BASY Platform

The core design of ProjectFlex was used to implement the DNA isolation workflow. Due to the extensive nature of the DNA isolation protocol and the occasional need to pause and restart the protocol, each step of the workflow was reorganised into a separate function. This allowed for more precise optimisation to remove redundancy and the step count allowed for continuation of the protocol.

The major feature still missing from the hardcoded program was the position status update GUI indicating the physical position of the robot on the platform as it processed the workflow. The sequential programming nature of Agilent VEE was excellent for managing a workflow type program, but makes it difficult to run another function in parallel for GUI-purposes. All indicators had to be positioned together with the active functions on a single panel. For this reason, colour indicators marking each platform position were added to the core structure, all of which are managed by a variable display function.

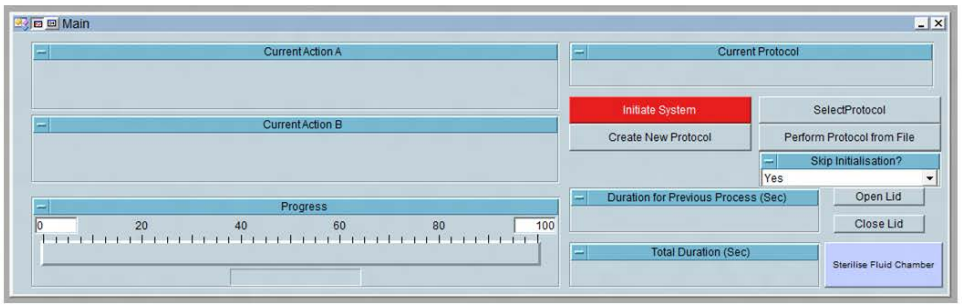
Within each line of instruction in the protocol, the variables were organised so that the first and second variable positions refer to the function name and step description, while the remaining positions three to twenty-five each represented an individual position status value. A value of -1 indicated that the position was inactive, while value of 1 or 2 represented the first or second of the interlaced process. This allowed the user to manually define each position where the step would be active. A more sophisticated approach was attempted by incorporating the indicator variables into each function, which then mark the indicator when activated, but the linear nature of the programming language imposed on the effectiveness of the implementation. Furthermore, such approach would only mark the latest position, and does not reflect the status accurately in steps involving multiple positions, such as the transfer of solution between two positions.

The DNA isolation workflow was optimised into fifty individual steps for the processing of eight plates, and classified under the “BASY” function class. The core algorithm from ProjectFlex was then used to call on these functions sequentially, with the additional variable display function updating the position colour indications of each step. This new GUI allowed the user to track the process in real time, without having to figure out the progress by studying the platform activity.

The upgrading of the water management system to an automated pump supply resolved previously frequent error message of “low water”, but intensified minor problems such as the overflow of the waste tanks, which were easily missed during busy workflow. The software was updated to include the water levels management by calibrating the flow rate of the waste water into the tank with the wash cycles, and imposing a “Waste tank full” error handling function when the limit was reached.

This update allowed the ProjectFlex core codes to be used for DNA isolation purposes.

A) GUI during development of ProjectFlex



B) GUI when ProjectFlex was used to implement the automated DNA isolation run

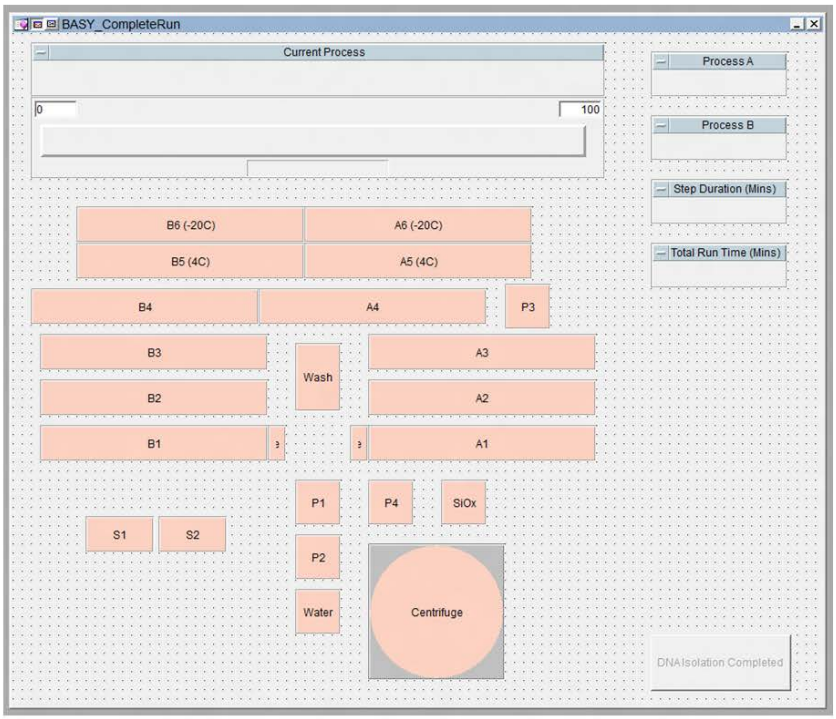


Figure 14: Improvement to the graphical user interface. A) During the development of ProjectFlex, and its implementation of the high-throughput transformation workflow, status updates provided were limited to text-based panels. B) The extensive nature of the DNA isolation protocol made it difficult to track using text updates, hence a graphical status display was created to report on the platform activity.

Implementation of ProjectFlex on the TFA Platform

ProjectFlex was also adapted for use on the TFA transfection platform. The TFA robot was a basic liquid handling robot equipped with a robotic pipettor similar to the BASY. The platform has sixty default positions defined in six rows of ten plates. Depending on the configuration of the workflow, the TFA is able to process up to sixty plates in a single run for simple protocols involving dispensing of solution, cell culture media or treatment. For example, during the validation screen, the TFA was used to process the PEI transfection workflow in triplicate sets of nine plates to prepare 27 96-well plates of transfection mix in a single run of about 65 minutes.

The ProjectFlex codes were in a stable version after extensive improvements from two major implementations for high-throughput transformation and DNA isolation. Only minor adaptations were introduced, with the core design remaining constant.

Due to the repetitive workload of the TFA platform, execution of the instruction line in the protocol file was placed under a nested loop structure. Repetition is possible in two forms: 1) repetition of protocol for unique plates and 2) the number of times each unique plate is repeated within a set of replicates. Optimal repetition of experiments may be achieved by looping these two types of repetition across the x- and y-axis. Using the validation screen example above, the nine unique plates of plasmid DNA would be subjected to the same experiment protocol on the y-axis, while three plates of the same transfection mixture was prepared for each unique plate. The distinction between the two forms of repetition greatly reduced the time dedicated to washing of the tips since repeating protocols for the same plates allowed the pipette tips to be reused.

To achieve this two axes repetitive effect, the function processing and executing the instruction line was placed with a “For” loop, which repeats the plate X times. This loop was nested within another “For” loop, which repeats the experiment protocol Y times. Functions were created to process the additional command-type instruction denoted by double square brackets “[[Command…]]”, which are used to indicate the start and end of loops so that a whole range of instructions may be repeated. Additionally, functions were also created to map the loop index position to actual physical coordinates. For example, if the experiment loop starts at the loop index of 1,1 which has a platform position of A1, the next plate to be processed will have a loop index of 1,2 which has to be translated to position A2. The loop index allows the process to be mathematically controlled, while the new position translation function retrieves the physical coordinates for the robot to process. In the absence of the loop commands, the software assumes the standard count of 1 for the respective “For” loops, which returns to non-replicate processing mode. This update allows for any combination of unique plates and replicates to be performed, limited only by the maximum plate

load at each axis. It also simplifies the protocol file through the use of command lines to mark region of instruction lines for repetition, eliminating the need to manually repeat the definition for each plate.

Other problems unique to the TFA platform includes its susceptibility to extended protocol, where unlike the BASY, it would terminates after processing certain number of instruction lines. This usually occurred around a similar region of instruction lines and was initially attributed to the lack of processing power and memory on the previous generation computer. While the upgrade of the computer did prolong the length of the protocol approximately proportional to the amount of memory (RAM), it did not completely resolve the problem. Extensive referencing of the TFA source code with those of the BASY platform and discussion with the systems engineer, Dr. Volker Kachel, identified the cause as the depletion of a serial interface memory buffer. A function was developed for the BASY platform, which routinely cleared this memory buffer, enabling the system to sustain the workflow. This function had not been applied to the TFA platform previously and a similar function was developed for the TFA to resolve the premature termination error.

Furthermore, the TFA had a smaller volume range of 10 – 100ul through the use of a 200ul tip compared with the BASY which has a range of 10 – 200ul using a 300ul pipette tip. The use of the 200ul tip increased the precision when handling smaller volumes necessary for transfection protocols, but created additional problems for the tip cleansing steps. Its maximum volume is half of the BASY robot, requiring the wash cycles to be doubled accordingly in the software. The smaller pipette tip also has the tendency for liquid droplets to form at the end, which is difficult to clear with the air ejection system. This residual droplet is estimated at approximately 5ul per tip, which could quickly add up to dilute the working solutions by over 15%. This is likely to be the reason why the calcium phosphate protocol was extremely difficult to implement on the platform. This was resolved by reprogramming the wash action; instead of a direct clearing of the residual droplet with increased air pressure, the robot was first programmed to touch the flowing water surface to remove the droplets before projecting the bust of air through the tips to clear any remaining residuals. This update together with the use of the robust PEI transfection protocol, dedicated reagent reservoirs for each plate and excess volume of transfection reagent greatly improved the stability of the transfection efficiency.

Data handling and other mini-scripts

Data handling scripts for processing screen data

Processing the huge amount of data generated during the screen was performed with Microsoft Visual Basic based scripts, also known as macros, and the Excel spreadsheet was used to store and aggregate the data and derivative calculations.

A suite of functions was developed to manage the data flow, centred on the Data Import and Candidate Identification functions, which are supported by a range of auxiliary functions such as statistics collection, error logging and ID mapping. These functions were created with a similar “Do While” loop structure, which keeps repeating the function for the next line until an empty space counter reached its limit.

The Data Import script was the most important function and was employed for the translation of the raw data format generated by the Optima FLUOstar plate reader. This raw data was encrypted and required the use of its associated commercial software to process and analyse. A basic license was available on the computer controlling the plate reader, with limited data viewing, export and analysis. Furthermore, the basic analysis can only be performed on the single licensed computer which prevented the use of the equipment. Each set of data comprised of 96 measurements of each microplate at 590nm wavelength. These sets were manually extracted into individual worksheet in Excel before automated processing.

The Data Import script was coded in sections: initiation, temporal data import, validation & transfer to final worksheet and housekeeping. Initiation comprises the variable declarations for the various worksheet names such as “Plate View”, “Error Logs”, “Attention Required”, “Dataset List” and other temporal worksheets, and the string, integer, real and Boolean variables. A dialogue box was then created for the user to select the data file workbook to import, and the file checked to have the correct file extension (.xls or .xlsx) prior to opening and import. The temporal data import section employs a “For” loop to sequentially process all worksheets in the selected file. Within each worksheet, the space delimited description is retrieved from the cell F2 and split to retrieve the date, lysis status, plate name and treatment type. Next, the 96 data points were retrieved into the temporal worksheet, and the step repeated until all worksheets were processed. This was followed by the validation and transfer section, whereby the CPRG datasets were identified and cut into the final worksheet “Plate View”. The remaining datasets were searched for the Triton lysis status, and matched to the corresponding CPRG dataset with the plate name. Any remaining datasets, along with datasets which were unsuccessfully imported, were flagged in the “Attention Required” while error messages were recorded under “Error Logs”. Once the import was completed, the script

proceeds to the housekeeping sections where the temporal worksheets are deleted and the source workbook closed. The Data Import script then calls for the Inducer, Inhibitor and Sensitizer candidate selection functions (Figure 15).

The candidate selection functions performed the implementation of the experimentally determined thresholds. For all selection, only pairs with the Triton signal greater than the 0.7 threshold were analysed. Inducers were selected with CPRG ratio greater than threshold of 0.55, and excluded from the Inhibitor-Sensitizer selection. Inhibitors and sensitizers were selected by ranking the TNTD between the treated and non-treated datasets, and the selection proportion for each candidate types calculated with Equation 11 to pick the appropriate candidate counts from each tail. The selected candidates were marked in the "Plate View" and shortlisted into three mutually exclusive candidate lists.

Other supporting functions were created to generate reports on data statistics such as the frequency and distribution of candidates, calculation of the distance-weighted estimator (DWE), and ID mapping. The ID mapping function uses the input from one list to search the collection in another, and is extremely useful for retrieving clone information using its accession numbers or plate position, mapping reindexed clones to the original library information and the incorporation annotation information provided by the Bioinfomatics Service or generated with online databases.

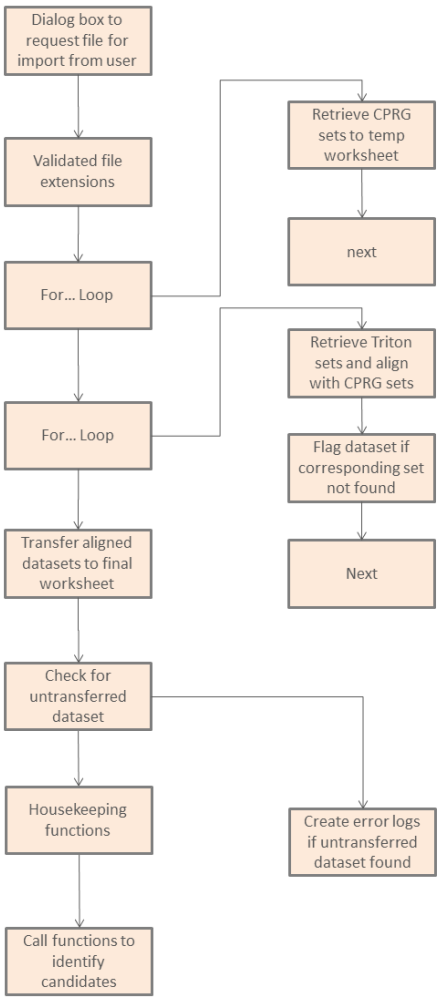


Figure 15: Data import algorithm for processing the FLUOstar raw data. After the manual extraction of the FLUOstar plate reader generated raw data, the data import macro automates the extraction, alignment and assembly of all data points into a single Excel spreadsheet. This allowed for downstream data handling and analysis with further macros or built-in Excel functions.

Mini Scripts

A host of macros were developed to assist with other aspects of the scaled up experiment settings. During the validation of the transformation efficiency and cDNA library integrity after the implementation of the high-throughput transformation, more than a hundred clones were randomly subjected to EcoRI and NotI double restriction digest analysis. A macro was created to search the DNA sequence information for the presence of each restriction site and calculate the potential restriction pattern. The predicted pattern could then be used as the reference for comparison with the experimentally generated pattern to validate each clone.

Another macro was created for the management of the sequencing data and selection of internal primers for further sequencing. This macro comprises of two parts for data input and primer selection. On activation, a panel was presented for the sequencing data to be entered after which this data was stripped of whitespaces and line breaks to generate a continuous string. The sequence length was counted, and if the sequence was from the BGH reverse primers or other primers sequencing the alternate strand, a standard reverse and complement function was performed. This processed string was submitted for primer selection. 80% of the sequencing data from the 5' end was retained and this marked the end point for primer selection. The remaining 20% was not used for selection and reserved for the purpose of complementary and sequence assembly. 100 primer candidates were selected before the 80% mark for each DNA sequence, with each candidate 20 nucleotide in length.

Each of the candidates were analysed for its GC content and predicted annealing temperature. The GC content needs to be approximately balanced with the AT count, with the predicted annealing temperature between ranges of 40°C - 45°C. Annealing temperature was predicted using Equation 7:

Equation 7: Estimation of primer annealing temperature

$$Annealing\ temperature = 69.3 + (0.41\ \times\ GC\ Count) - (\frac{650}{Primer\ Length})$$

Next, preference was given to candidates ending with a guanine or cytosine due to the higher binding affinity, and sequences with long repeats of five or more of the same nucleotide were excluded. Each successful criterion increased the candidate scores, and only those that were positive for every defined criterion were selected. The candidate nearest to the 80% mark was selected as the primary choice, with the positions of the remaining successful candidates listed as alternatives.

Over 300 primers were designed with this macro with approximately 95% success rates. The sequencing experiments were discussed in greater details in Chapter 7: Implementation of the Screen (Sequencing and Alignment).

This early version of the loop structure and scoring system for candidate selection was eventually optimised and developed into the structure deployed in the data handling set of functions central to the screen analysis.

Concluding Statement

High-throughput approaches are exceedingly dependent on the ability to automate the experimental process, and this extends beyond the bench work to include the post-experiment data processing and analysis phase. The updating of the RISIC platform software greatly contributed to the ease of use when implementing new protocols on the robotic platforms. Relinquishing the management of error handling and other tracking procedures such as water levels to the robot software also greatly relieved the workflow, allowing attention to be directed towards more significant tasks. Furthermore, the move away from hardcoding towards more dynamic software enabled new alternatives for simultaneously integrating various workflows, improving the overall time efficiency.

The use of supplementary codes such as the Excel macros to handle the data processing removed the need for otherwise laborious manual processing. The current coding structure which heavily relied on various types of loops to manage the data processing works well due to the limited nature of the data points, which meant that data processing duration is generally within minutes for batches and generally completed under an hour when processing the entire screen. However, the pursue for higher throughput which is still theoretically possible for the current hardware may necessitate the development of more dynamic structures. Such structures may for example require the datasets to be split and processed in parallel to drive down the overall processing time since loop structures have a tendency to slow with cycle increase.

Nonetheless, the software and data processing software presented have contributed to improving the automation process and the ease of use for the systems.

Chapter 6: Characterisation of the Nanoparticles

Background Information

The diversity of nanoparticles and their increasing industrial and medical applications created exciting opportunities and necessity to explore any toxicity and its mode of action. While excessive concentration of nanoparticles was known to result in necrosis, apoptosis regulated via various signalling pathways could be differentially associated with each type and chemistry of nanoparticles. The RISCI platform offers a systematic approach to study the regulation of these signalling pathways when the cellular system becomes exposed to nanoparticles induced toxicity. Integration of nanoparticles treatment is a novelty with numerous advantages. It would enable the RISCI screening approach to identify two additional classes of regulators, namely the inhibitors and sensitizers, while the candidate lists could potentially indicate enrichment of specific signalling pathways or novel regulatory mechanisms unique to nanoparticles induced toxicity. This, in turns, could lead to the prospect of generating new hypotheses, contribution to the regulatory knowledgebase crucial to implementation of health and safety guidelines and the discovery of biomarkers.

This chapter explores the toxicity associated with nanoparticles, the characterisation of their chemistry and structures, and the attempts to incorporation this new treatment approach into the RISCI screen.

Results

Toxicity of Nanoparticles

A549 or HEK293T cells were exposed to a variety of nanoparticles under a dose-response experimental setup and screened for toxicity using light microscopy. Nanoparticles tested included commercially available ~157nm silver dispersion citrate-stabilised in water (Sigma Aldrich 675318-5ML), <150nm titanium(IV) oxide dispersion in water (Sigma Aldrich 700347-25G), LUDOX® TM-40 ~22nm colloidal silica suspension in water (420786-1L), as well as 10nm and 100nm silica dispersion in water specially synthesised by Dr Agnieszka Dybowska at the Natural History Museum (NHM).

The silver nanoparticles tested did not result in observable toxicity even when the undiluted solution with the final concentration of 0.0269ug/mL was used in HEK293T (Figure 16A). The titanium(IV)

oxide nanoparticles was not soluble in the cell culture media, increasing the opacity of the media. Increasing its concentration resulted in visible precipitate forming, which eventually generated huge aggregates covering the cells at high concentrations such as 150ug/mL. While toxicity was not observed at lower does such as 5ug/mL, use of concentration of 50ug/mL or more led to visible cell death (Figure 16B). At 50ug/mL, the cells appeared to die by a variety of cell death including apoptosis, but this remained inconclusive. However, apoptosis may be excluded as the form of cell death at 150ug/mL since the apoptotic phenotype was definitely absent. The presence of precipitate and the resulting cloudiness of the media greatly interfered with methods for quantification of cell death including the LDH and CPRG assay, as well as flow cytometry.

Silica nanoparticles in contrast were easy to work with, their stability in water allowed for very high concentrations of up to 40% w/v to be achievable. Nano-silica displayed consistent toxicity, with the apoptotic phenotype of rounded cells clear present between the dose ranges tested (Figure 16C). The NHM-synthesised 15nm particles were observed to induce 23.92% cell death at 125ug/mL in A549 cells, which rapidly increases to 74.61% at 150ug/mL. 100nm particles displayed a delay in toxicity, with only 42.39% cell death at 150ug/mL but increasing to 84.34% at 175ug/mL which is similar to the 15nm particles in A549 cells (Figure 17A). The commercially available LUDOX® was more toxic to the cells, with 27.17% cell death at only 40ug/mL, but displayed similar characteristics to the NHM-synthesised particles whereby the toxicity increased sharply towards the maximal within a small increment (Figure 17A). Quantification of cell death was performed using DiOC6-PI flow cytometry.

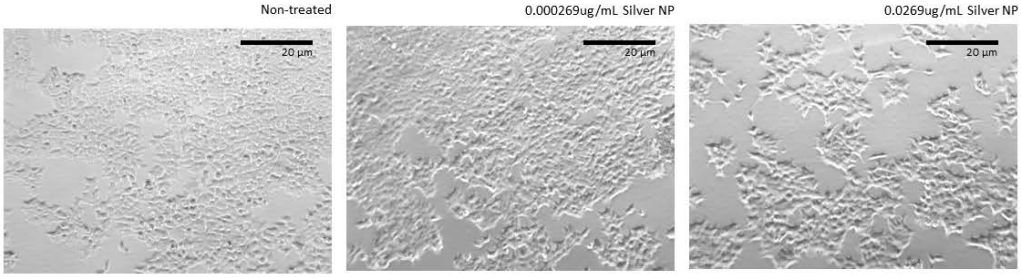
Silica nanoparticles from both sources were subjected to the CPRG assay, during which the nanoparticles were applied to HEK293T cells transfected with GFP. The CPRG assay could detect the increase in toxicity with greater degree of sensitivity. Here, the 15nm silica nanoparticles (NHM) induced cell death was quantified (Figure 17B), with an increase in mean observable at 75ug/mL or 40% improvement in the lower detection limit. Furthermore, at 100ug/mL the mean CPRG ratio was extremely similar to the apoptosis criteria of 0.55. However, at 125ug/mL the treated sample mean showed a decrease and the silica treated samples generally had greater standard deviation. This is likely to be due to the narrow transit range of the nanoparticles in combination with pipetting errors.

HEK293T cells treated with LUDOX® silica nanoparticles were also subjected to the LDH assay (Figure 17C), which measures the amount of lactate dehydrogenase (LDH) released from the cells due to cell lysis. The LDH assay may be a suitable indicator for estimation of necrosis. Across the range of 0 – 120ug/mL treatment concentration, none generated significantly LDH release against the

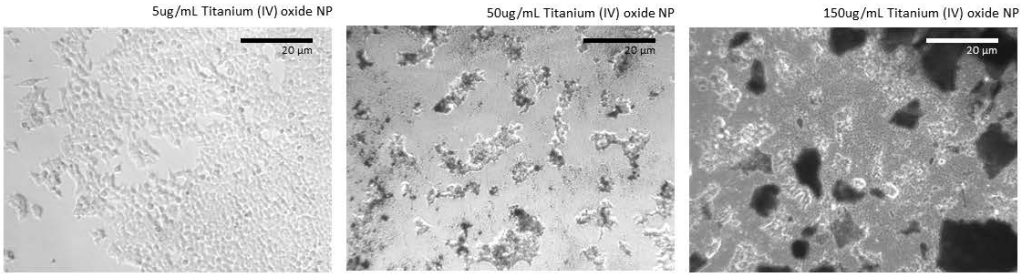
background, whereas the signal control which were lysed in parallel resulted in a strong LDH signal. This indicated that uncontrolled lysis of the cells was not triggered by the LUDOX® silica nanoparticles; in contrast the signal generated by the CPRG ratio is likely to be attributed to apoptosis.

Finally, a range of concentration of LUDOX® silica nanoparticles were applied to HEK293T cells to study the dose response against time. HEK293T cells were tolerant to concentration of the silica nanoparticles less than 20ug/mL, with no observable differences between the treated samples and non-treated controls. At 40ug/mL, moderate level of apoptosis was first observed at the 24hours time point. Increasing the concentration further causes the cells to become susceptible to the silica nanoparticles at an earlier time point (Figure 18). 80ug/mL induces apoptosis six hours after treatment, 100ug/mL at 4 hours while 150ug/mL results in an almost immediate activation of apoptosis. The increase in concentration also causes the cells to undergo apoptosis with a greater magnitude, with nearly the entire population becoming apoptotic at high concentrations (>60ug/mL).

A) Silver nanoparticles (Sigma Aldrich 675318-5ML)



B) Titanium(IV) oxide nanoparticles (Sigma Aldrich 700347-25G)



C) LUDOX® TM-40 silica nanoparticles (420786-1L)

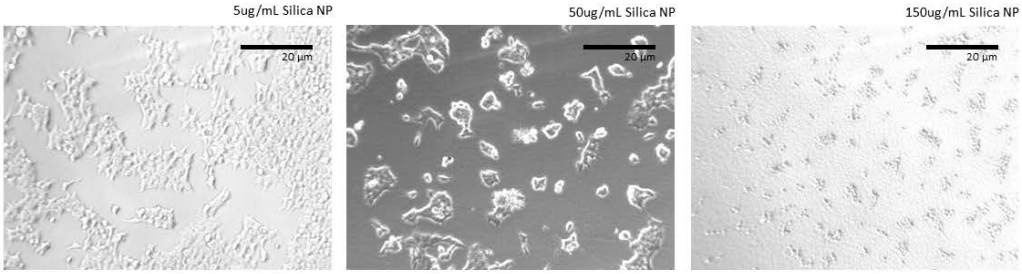


Figure 16: Dose response of various nanoparticles in HEK293T cells. Magnification = 200X, scale bar = 20µm. A) The silver nanoparticles tested did not induced observable cell death under the chosen conditions, and is likely due to a result of insufficient quantity of nanoparticles present. B) Titanium (IV) oxide nanoparticles did not induce cell death at low concentration such as 5ug/mL, but exhibit toxicity at concentration equal to or greater than 50ug/mL. C) Silica nanoparticles also induce toxicity at 50ug/mL and above, and continue to display morphology of apoptotic cells at higher concentrations.

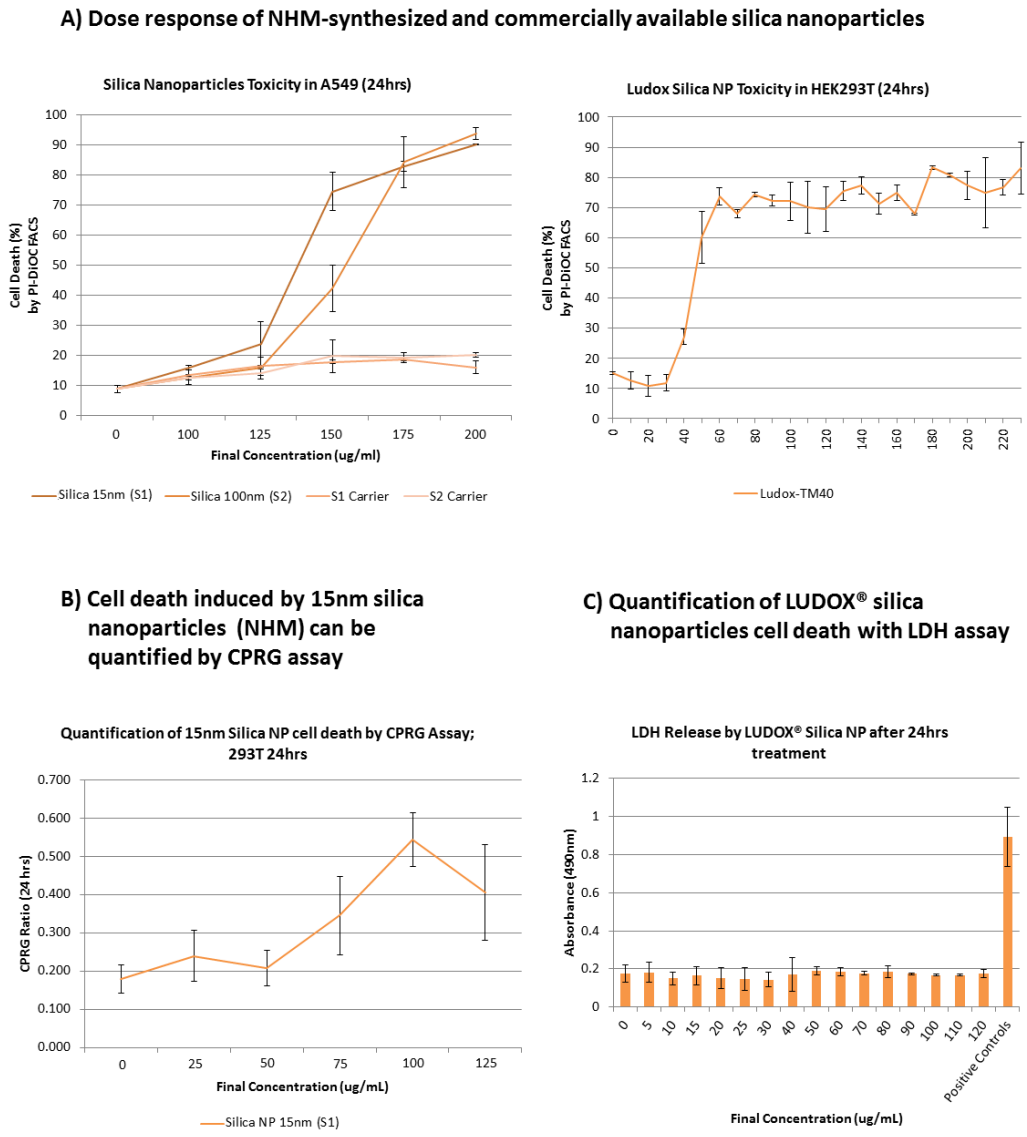


Figure 17: Quantification of the dose response of cells to silica nanoparticles. A) NHM-synthesised provided by our collaborators and commercially available silica particles were used in the toxicity studies and quantified with DiOC6-PI flow cytometry. The NHM-synthesised nanoparticles were tested in A549, while commercially available nanoparticles were tested in HEK293T cells. N = 3, error bars represent standard deviation of sample size. B) The CPRG assay could detect cell death induced by 15nm silica nanoparticles (NHM) in HEK293T cells. N = 3. C) Increased cell death was also detected in HEK293T cells treated with the LUDOX silica nanoparticles using the CPRG assay (not shown) but not in the LDH release, suggesting that the cell death comprised predominantly of apoptosis. N = 3, error bars represent standard deviation of sample size.

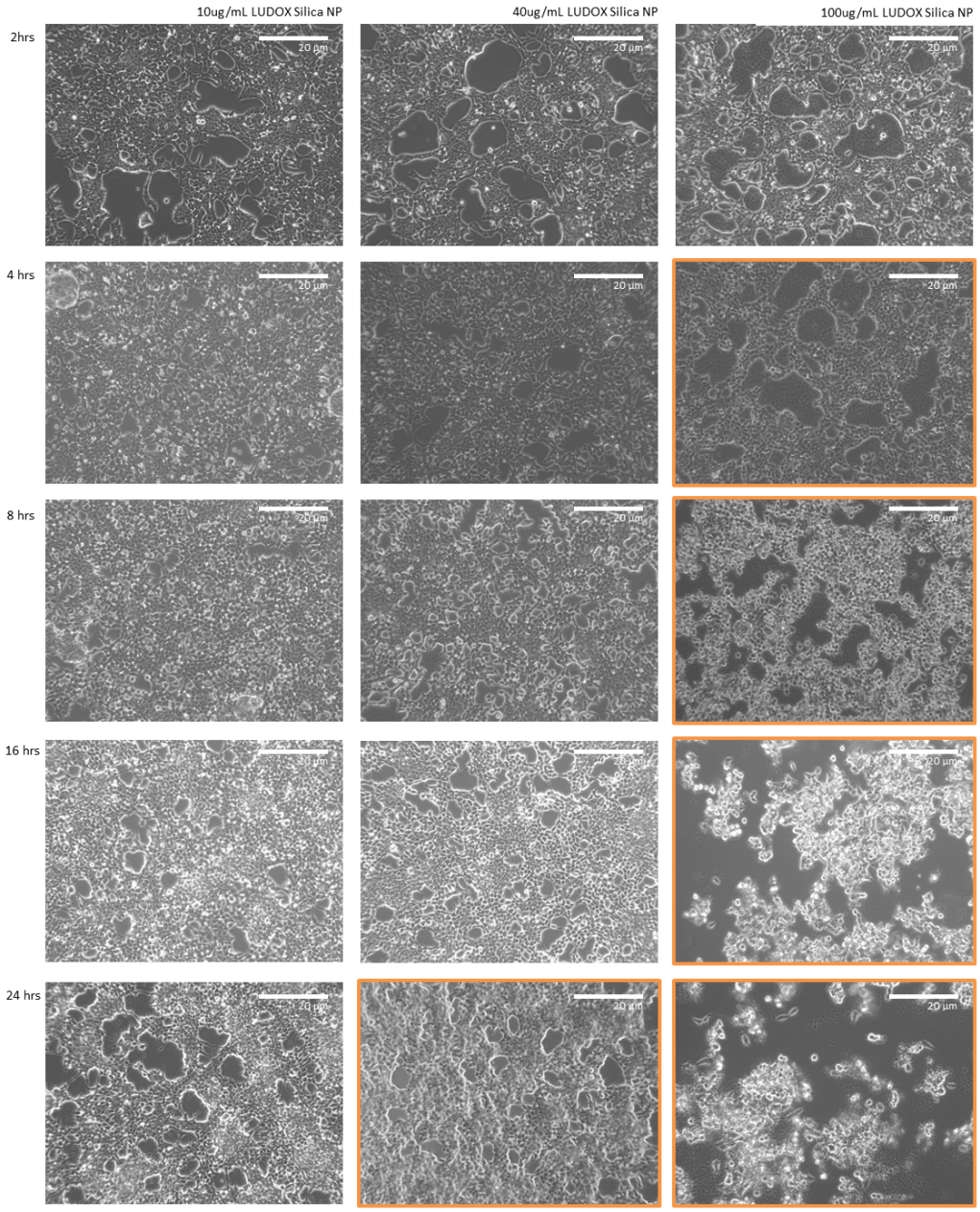


Figure 18: Dose response of the LUDOX® silica nanoparticles with time. HEK293T cells appear to be tolerant to low concentrations of the LDUOX® silica nanoparticles (<20ug/mL), with toxicity observed at 40ug/mL after 24 hours. Increasing the concentration causes the cells to become susceptible to the nanoparticles at an earlier time point and eventual cause cell death on a greater magnitude. The time points where apoptosis were observed were labelled with an orange outline. Magnification = 200X, scale bar = 20µm.

Characterisation of the LUDOX® Silica Nanoparticles

Chemical characterisation of the LUDOX® silica nanoparticles was performed by Dr Agnieszka Dybowska from the Earth Sciences Department, Natural History Museum (NHM).

The LUDOX® nanoparticles were characterised using x-ray diffraction (XDR) to study crystallinity, the BET method to study the surface area, transmission electron microscopy (TEM) to analyse the size, Fourier transform infrared (FTIR) spectroscopy for purity and surface characteristics, and dynamic light scattering (DLS) for size, surface charges and stability of the nanoparticles in the cell culture media.

The samples were dried at 60°C to obtain a silica powder for the XDR, BET and FTIR analysis, while the original colloid was used for the DLS and TEM.

The silica powder was scanned using Enraf-Nonius diffractometer coupled to INEL CPS 120 positive-sensitive detector with Co K_{α} radiation. The XDR results indicated that the LUDOX® sample consists of only amorphous silica (Figure 19A). Traces of crystalline phase of silica or other impurities were not found in the sample.

For FTIR analysis, the powder was mixed with potassium bromide to form a pellet, which was then analysed by PerkinElmer Spectrum 1 FRIR. Three absorption peaks were detected at 477cm-1, 802cm-1 and 1105cm-1 which are characteristic of silica (Figure 19B). The 477cm-1 peak corresponds to the bending vibration of the Si-O-Si bond, while the 802cm-1 and 1105cm-1 peaks correspond to the symmetric and anti-symmetric stretching vibration adsorption of Si-O bond respectively. Another peak at 3433cm-1 corresponds to the Si-OH and stretching vibration absorption of the O-H bond of physically adsorbed water and the 1641cm-1 peak represents the bending vibration absorption of physically adsorbed water. The surface of the silica nanoparticles was not functionalised since only absorption peaks characteristics of silica and water were detected.

For TEM imaging, the colloidal silica nanoparticles were diluted with water then deposited on a copper grid and left to dry at room temperature overnight prior to imaging. Images captured using the Hitachi 7100 TEM with accelerating voltage of 100kV show the silica nanoparticles to be spherical with an estimated size of 28nm (Figure 19C).

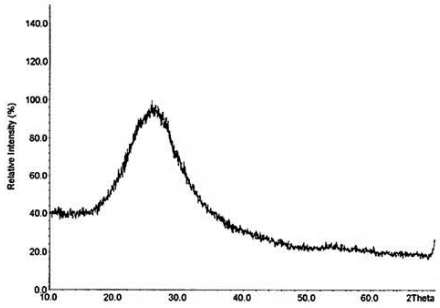
The zeta potential of the silica nanoparticles was measured with the Autotitrator on Malvern Zetasizer Nano, and the pH was titrated from pH 9.0 to pH 1.0. The zero point of charge was determined to be pH 1.96 (Figure 19D). The zeta potential measures the degree of repulsion between particles with the same charge. A high zeta potential confers stability to small particles or molecules since the dispersion will resist aggregation. The colloid is electrically stabilised when the

zeta potential is positively or negative high, and a decrease in the zeta potential indicates attraction forces increasingly exceeding repulsion, causing the dispersion to destabilise, flocculate and eventually precipitate. While the stability of the nanoparticles is reduced by under the near neutral pH of the cell culture environment leading to a shift towards larger particle size, the nanoparticles remained fairly well dispersed in the culture media.

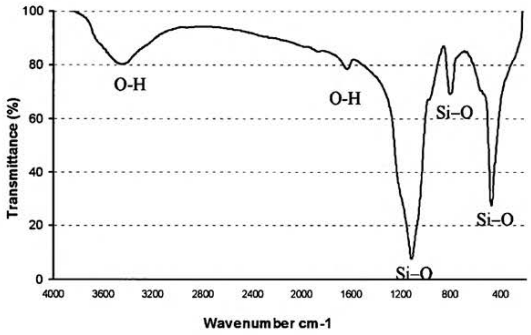
DLS was performed using the Malvern Zetasizer Nano instrument equipped with a He-Ne 633 nm laser. 40ug/mL of the silica nanoparticles were diluted into the media, vortexed for one minute and incubated for five minutes at 37°C before transfer into size cuvettes for measurements. In the absence of serum, the particle size in cell culture media (DMEM) was 26.3nm, and follows an approximately linear scale as the serum (FBS) concentration was increased up to a maximum of 173nm at 10% FBS. The particle size of 101nm at 1.25% appears to be an experimental outlier, and estimation from the correlation function would give a particle size of approximately 50nm (Figure 20A). The quoted average particle size refers to peak position in DLS graph (Figure 20A) after excluding peak contribution generated by the media and FBS.

Figure 20B shows the results from a 24-hour incubation period during which the particle size in media without FBS or supplemented with 1.25% FBS was monitored using DLS. In the absence of the serum, particle size appeared to continuously increase over time in the DMEM, while the addition of 1.25% FBS stabilises the particle size within the first six hours of incubation. This may suggest that the serum protein plays an important role in the formation of the corona (a shell of biomolecules or proteins) around the nanoparticles, which could aid the passage of the nanoparticles across the plasma membrane.

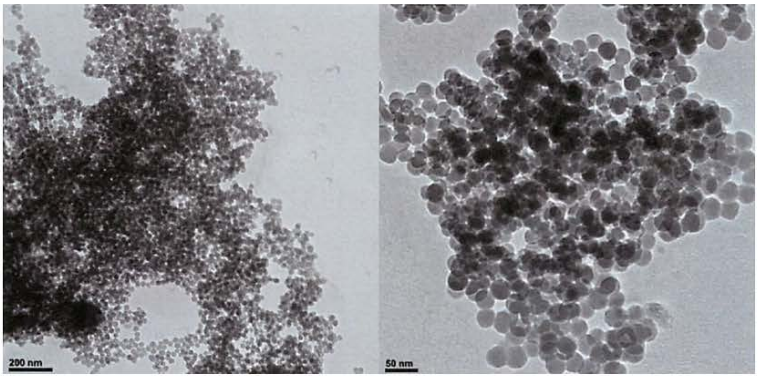
A) XDR pattern



B) FTIR absorption pattern



C) TEM images of the LUDOX® silica nanoparticles (colloid)



D) The point of zero charge

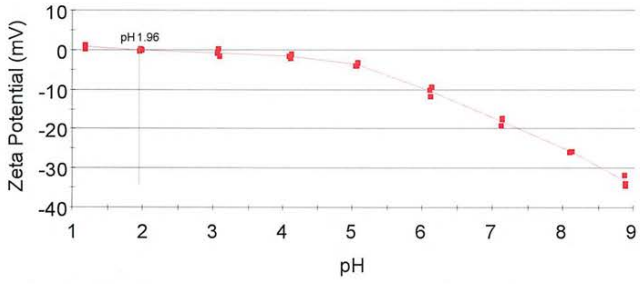
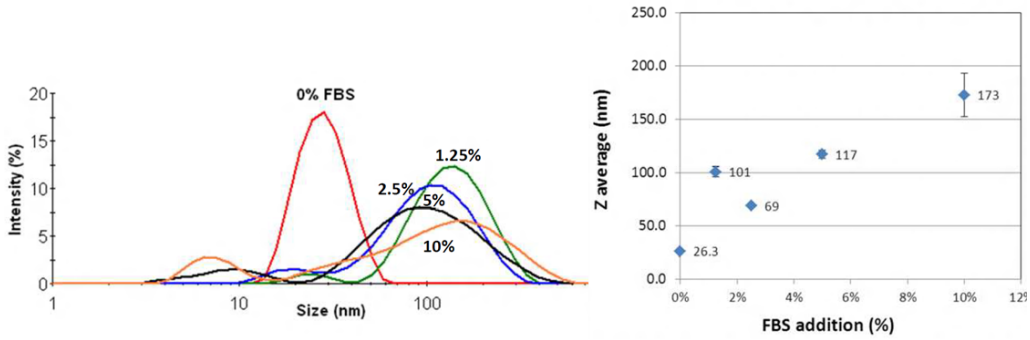
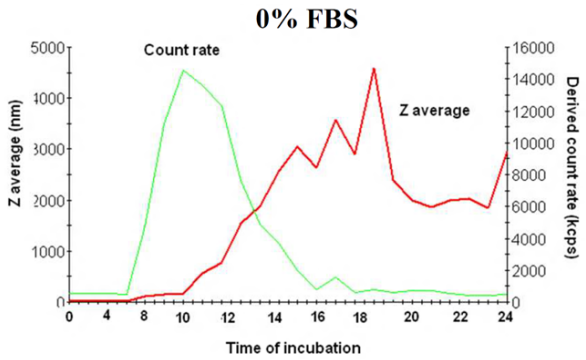


Figure 19: Characterisation of the LUDOX® silica nanoparticles. A) X-ray diffraction indicates that the silica nanoparticles exist as a single amorphous phase, with no other crystalline phases or impurities detected. B) Fourier Transform Infrared spectroscopy results show only absorption peaks characteristic of silica and confirm that the nanoparticle surface is not functionalised. C) Transmission electron microscopy images show the spherical form of the silica nanoparticles, with an estimated diameter of 28nm. D) The point of zero charge estimated at pH 1.96 from the titration experiment. The decreased in zeta potential at neutral pH of the cell culture media destabilises the nanoparticles, increasing its size. However the nanoparticles remain well dispersed in the media. Figure prepared by Dr Agnieszka Dybowska.

A) Average particle size as a function of serum addition



B) The presence of serum stabilises particle size in solution over long incubation periods



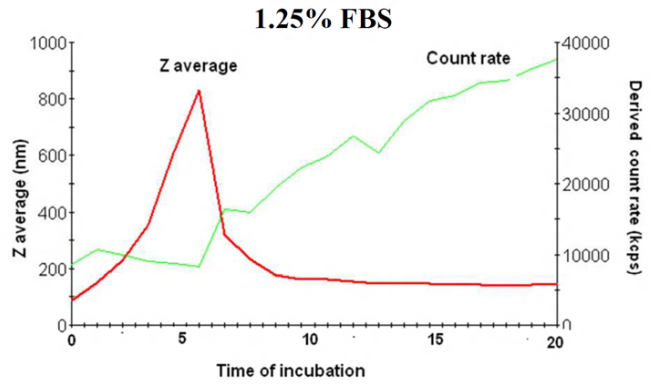


Figure 20: Effects of cell culture serum on the LUDOX® silica nanoparticles. A) Increasing the amount of serum (FBS) in the cell culture media (DMEM) led to increase in particle size, with a maximum size of 173nm at 10% serum concentration. B) The particle size of the LUDOX® silica nanoparticles was tracked using DLS over a 24 hour incubation period in DMEM with or without the supplement of 1.25% FBS. Addition of 1.25% FBS caused the particle size to stabilise within the first six hours of incubation. Figure prepared by Dr Agnieszka Dybowska.

Inhibition of Silica Nanoparticles Induced Cell Death

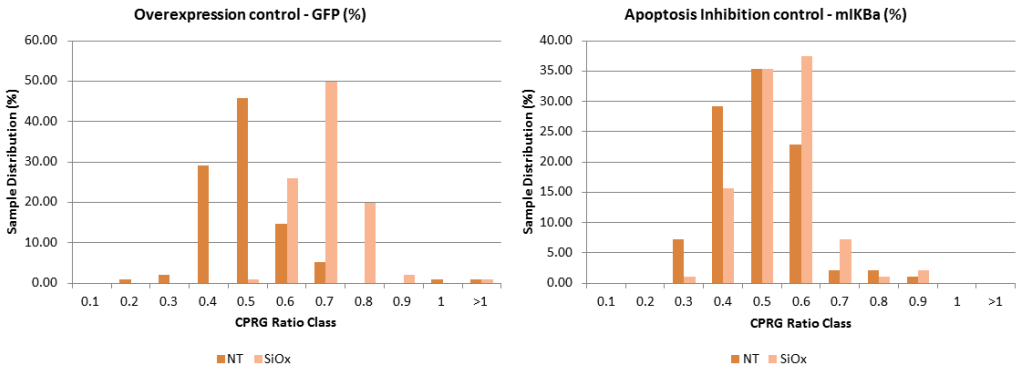
The toxicity studies on nanoparticles revealed that the CPRG assay could be used to quantify the cell death induced by the silica nanoparticles with a greater sensitivity than flow cytometry. The converse needs to be demonstrated for the CPRG assay and RISCI screen confirming the inhibition of the CPRG ratio after a treatment condition which induced cell death.

A construct of IκBα targeted to the mitochondria outer membrane (mIκBa) generated by Evangelos Pazarentzos, a PhD student in our group, was demonstrated to show strong apoptosis inhibition effect. This construct was used as the apoptosis inhibition control while the GFP was used as an overexpression control. Both controls were co-transfected with the β-galactosidase reporter enzyme into HEK293T cells, and 24 hours post-transfection the cells were subjected to 0 – 70ug/mL of LUDOX silica nanoparticles treatment for a further 24 hours. At 40ug/mL, mIκBa was observed to prevent increase of CPRG ratio compared with the GFP control. This was further investigated with 96 samples transfected by the optimised robotic platform. The distribution of the CPRG ratios for non-treated cells and in the presence of 40ug/mL silica nanoparticles was presented in Figure 21A. For the overexpression control GFP, the treatment caused the CPRG ratio to shift towards higher values compared with the non-treated baseline. In contrast, mIκBa was able to prevent this shift towards higher values.

This inhibition effect was also observed by comparing just the CPRG ratios from the treated set without accounting for the baseline (Figure 21B). This comparison method formed the initial selection criteria for identifying inhibitor candidates, but was eventually changed to the Treated-Non-Treated Difference (TNTD) which accounted for the baseline and provided greater statistically significance (See Chapter 7: Implementation of the Screen). The TNTD threshold was calculated from these readings to be 0.2 for GFP and 0.05 for mIκBa.

Established inhibitors of apoptosis such as Bcl-2 and Bcl_{XL} were also tested using similar experimental setup, but results were inconclusive with both controls showing excessive toxicity on transfection. The plasmids were eventually sequenced and found to be damaged, and was not pursued further since the apoptosis inhibition was readily observed using mIκBa.

A) Distribution of CPRG ratio values after treatment with silica nanoparticles



B) mIKBa inhibited increase in CPRG ratio after treatment with silica nanoparticles

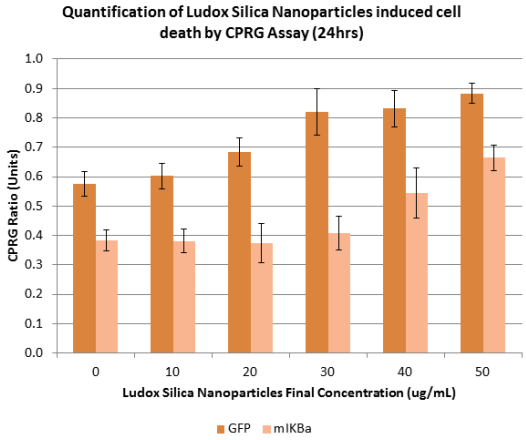


Figure 21: Inhibition of LUDOX® silica nanoparticles induced cell death in HEK293T cells by mIκBa. A) Distribution of the CPRG ratios in the overexpression (GFP) and apoptosis inhibition (mIκBa) controls indicated that the apoptosis-inhibitory effect may be observed using the CPRG assay. N = 96. B) Comparison of the CPRG ratios between GFP and mIκBa subjected to silica nanoparticles treatment. N = 12, error bars represent standard deviation of sample size.

Discussion

The diversity of nanoparticles and available surface functionalization make it likely that each type of nanoparticles is likely to trigger toxicity, if any, through a unique mode of action imposed by its chemistry. This diversity also makes studying the toxicity of every nanoparticle types and modification impossible and the project focused on nanoparticles commonly used in industrial and consumer products, including silver commonly employed for its anti-microbial activity [297], titanium oxide which is used in sun screen and paints [298] and silica which has extensive applications including in the biomedical field as drug delivery carriers [299, 300]. At the start of this project, it has been commonly recognised that application of high concentration of nanoparticles to cells often results in necrosis, but information regarding the type of toxic response and associated regulation mechanisms remained rare.

The nano-titanium induced toxicity resulted in cell morphology which did not correspond to apoptosis. It was also difficult to work with, in particular due to increased opacity of the media and the precipitation of the particles onto the cells. The opacity caused by the nano-titanium interfered with traditional 96-well assay format such as MTT, LDH and CPRG assays, while the insoluble precipitates interfered with flow cytometry techniques, all of which made quantification of cell death difficult. Nano-titanium is likely to cause necrosis at high concentration such as 150ug/mL, where the presence of large amount of cell debris was observed in the absence of typically rounded apoptotic cells. This necrotic effect may help explain reported allergy to sunscreens containing titanium dioxide [301], since the uncontrolled cell lysis is likely to result in an inflammatory response. It was excluded from further optimisation and integration for the RISCI screen due to its interference with the CPRG assay.

The silver nanoparticles were not observed to induce toxicity at the concentrations tested. This lack of toxicity could be the result of insufficient nanoparticles concentration, although this could not be verified further in our experiment settings. Due to the lack of sufficient quantity of the nano-silver and the high acquisition costs, these particles were not tested further. The anti-bacterial effect of silver nanoparticles had been well established but the current literature remained vague in the toxicity of these nanoparticles to biological systems. Recent publications indicated that silver nanoparticles may have an advert effect on human macrophages [302, 303] and could induce oxidative stress, genotoxicity and even apoptosis in cultured cells and animal tissues [304, 305].

Silica nanoparticles however displayed consistent toxicity with in-house synthesised nanoparticles provided by our collaborators and commercially sourced nanoparticles inducing cell death with cell morphology similar to apoptosis. The size of the nanoparticles appears to be an important factor to the toxic response, with smaller diameter nanoparticles displaying increased toxicity. The interaction of silica nanoparticles with macromolecules from its biological environment is likely to mediate its final size, and hence play a critical role in determining the toxicity for this type of nanoparticles.

The LUDOX® silica nanoparticles were selected as the primary toxic agent in the RISCI screen due to its availability and costs. The nanoparticles were extensively characterised by Dr Agnieszka Dybowska to confirm that the silica remained stable and in the nano-scale under the experimental setup. XRD and FTIR analysis confirms the purity of the nanoparticle sample and that the nanoparticle surface was not functionalised. TEM imaging confirmed the spherical form of the nanoparticles and provided an estimate to its diameter at 28nm. DLS and zeta potential measurements indicated that the silica nanoparticles remained in the nano-scale in the cell culture media, and while the zeta potential is reduced in this environment causing the particles to destabilise and increase in size, the nanoparticles remained evenly dispersed in the media. Supplementing the media with serum increased the estimated particle size to a maximum of 178nm at 10% FBS, and this follow an approximately linear function. The presence of 1.25% serum also seemed to stabilise the size of the nanoparticles within a few hours of interaction. This observation confirms the importance of the corona, which is the clustering of proteins and other biomolecules in the immediate vicinity around the nanoparticles, in mediating the interaction and toxicity of the nanoparticles with the cells. It may be concluded that under the experiment conditions of 1.25% and 40ug/mL of the LUDOX® sample, the silica is pure and remained as nanoparticles which are evenly dispersed.

The cell death induced by the LUDOX® nanoparticles was also characterised by the CPRG assay (Figure 21A) and the LDH assay (Figure 17C). The LDH assay results indicated that the nanoparticles were not causing non-specific necrotic lysis of the cells during the induction of cell death, suggesting that the membrane lysis detected by the CPRG assay comprised mostly apoptosis. Furthermore, this form of cell death could also be inhibited by a known apoptosis inhibitor, mIκBa, as measured using the CPRG assay, providing further supporting evidence that apoptosis was induced by the LUDOX® silica nanoparticles.

Dialysis experiments by the postdoctoral researcher Dr Odu Okoturo on the LUDOX® silica nanoparticles further confirmed that toxicity results from the nanoparticles and not the carrier solution. Membranes with various pore sizes were used to retain the nanoparticles during dialysis.

Application of the retained content to A549 cells was shown to induce cell death (data not presented).

The effectiveness of the CPRG assay in quantifying this inhibitory effect by the mIκBa construct confirmed the possibility of using the CPRG assay to identify apoptosis inhibitors and potentially sensitizers.

Chapter 7: Implementation of the Screen

Background Information

The implementation of the screen followed on the extensive setup and optimisation phase which looked into the various aspects of the screening workflow such as process efficiency, costs, practicality of the implementation and potential bottlenecks and pitfalls.

The process, especially the validation assays, was first trialled on mouse cDNA clones identified as apoptosis inducing candidates through a previous screen to accustom to the high-throughput workflow. This tested various processes such as the potential validation assays which included the nucleosome ELISA and PARP western blotting to confirm the apoptotic activity of the candidates. Bioinformatics analysis approaches such as DAVID were also trialled for the first time on the candidate genes isolated from the RISCI platform.

The primary screen of the NITE library cDNA collection was then scaled up in phases, beginning with control runs, followed by limited number of plates from four per day. As these initial results of the screen were generated and familiarity with the workflow improved, the throughput was scaled upwards to the standard eight plates and further doubled as the pace reached its maximal practicality.

Based on the data generated in the primary screen and the selection criteria determined experimentally, candidates for apoptosis inducing, inhibiting and sensitizing activity were then isolated into a hit library for further validation.

Results

Validation of the Mouse cDNA Apoptosis Candidates

Sequencing and Alignment

118 *Mus musculus* derived cDNA sequences cloned into pcDNA3.1 which were identified as pro-apoptosis gene candidates in a previous screen were sequenced to verify the identity of the sequences. Sequencing was performed first in the forward and reverse directions using the T7 promoter and BGH polyadenylation signal sequences flanking the cDNA. Next, internal primers were

designed from these initial sequences in the 5'-3' direction using a computer script to continue the sequencing. DNA sequencing was performed using the standard chain-termination sequencing by the MRC DNA Core based at the Imperial College London Hammersmith Campus.

Each cDNA was repeatedly sequenced using the internal primers until all sequences may be aligned into a single contig. In total 340 sequencing primers were designed and used, from which 321 primers generated good quality reads (94.41%). Approximately 281,442 bases were derived for the clones, giving fold coverage of 1.67. 117/118 (99.15%) clones were successfully sequenced and aligned, with only a single clone not generating any sequences despite both manual and automated primer designs.

The sequencing reads were assembled into a continuous sequence using the online CAP3 Sequence Assembly Program provided by the Université Claude Bernard Lyon 1, France [306]. 22 clones (18.80%) required the use of an additional published sequence homology from the NCBI database to completely align the sequencing reads due to small gaps which failed to be closed after repeated sequencing.

Validation of Apoptosis Activity[2]

The mouse pro-apoptotic candidates were independently validated for other markers of apoptosis. The late-stage degradation of chromosomal DNA characteristic of apoptosis was investigated using the Cell Death Detection ELISA^{PLUS} (Roche) kit. The kit uses the absolute signal at 405 nm referenced to the negative control to calculate the fold increase in apoptosis. GFP was used as the overexpression negative control (the reference) while RIP was used as the apoptosis positive control. The calculated value of GFP was always 1 since it was the reference, while RIP signal was quantified at 3.393. 79 of the non-redundant candidates were positive for increased nucleosome release into the cytosol and significantly ($p < 0.1$) increased from the GFP negative control (Figure 22A). This reaffirms the capacity of the CPRG assay to identify pro-apoptotic genes.

Accumulation of cleaved PARP was also quantified using western blot for each candidate. This assay is an indirect measure of caspase-3 activity, an established marker of apoptosis, and validated 75 candidates as positive (Figure 22B). The general tendency between the ELISA and western blot generated validation results appears to be similar, with the apoptosis intensity increasing with the candidate genes clone number, indicative of the accurate ranking achieved with the primary CPRG ratio (Compare Figure 22A and B).

[2] Results from this section has been published in the following article: Lin B, Huntley D, AbuAli G, Langley SR, Sindelar G, et al. (2011) Determining Signalling Nodes for Apoptosis by a Genetic High-Throughput Screen. PLoS ONE 6(9): e25023. doi:10.1371/journal.pone.0025023

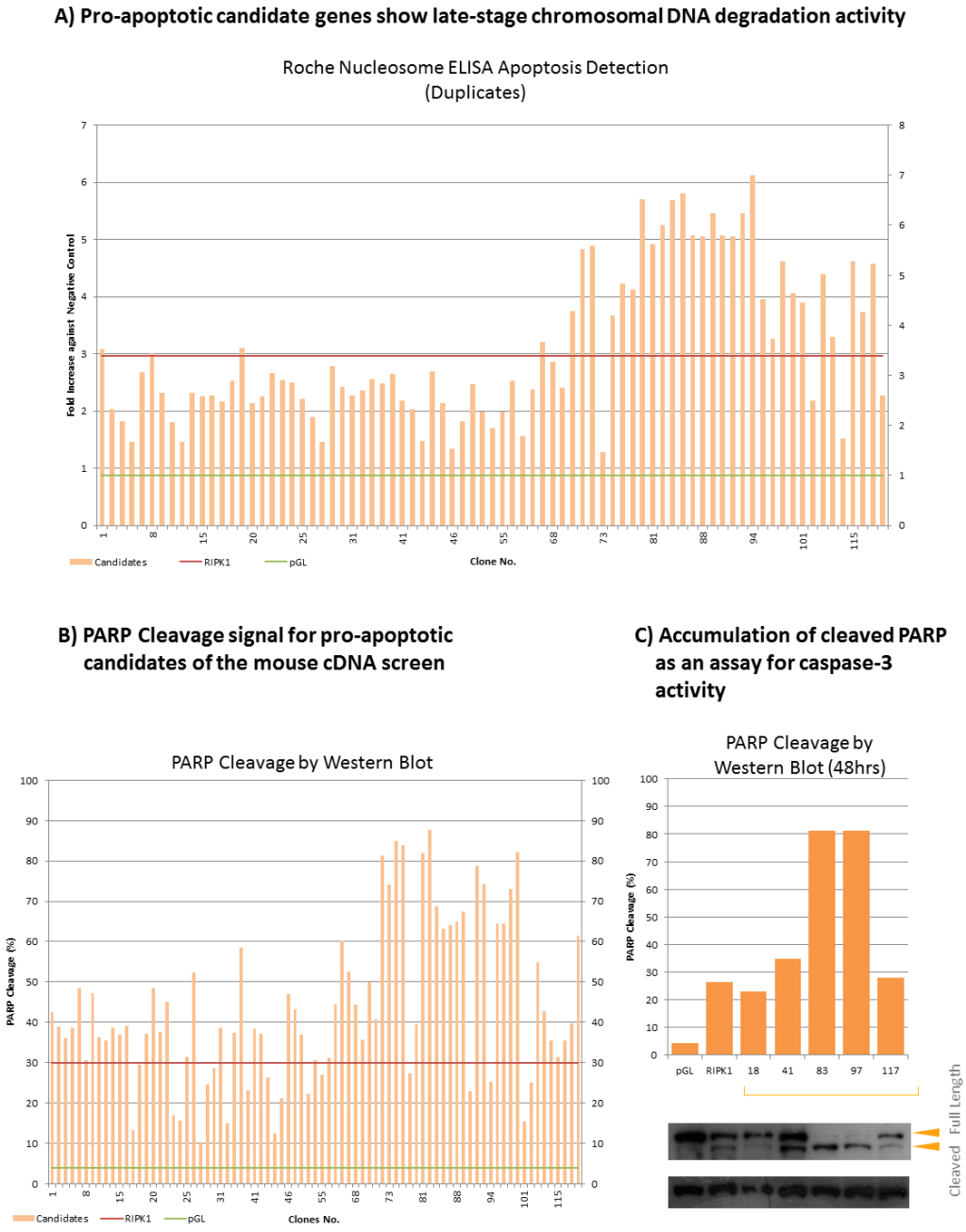


Figure 22: Validation of the mouse pro-apoptotic candidates using alternative apoptosis assays in HEK293T cells. A) The Cell Death Detection ELISA^{PLUS} Kit (Roche) was used to probe the pro-apoptotic candidates for nucleosome release into the cytoplasm, a marker for late-stage DNA degradation characteristic of apoptosis B) PARP cleavage quantified using western blot band intensity determined the level of caspase-3 activation for each candidate. C) Control experiment performed in duplicates showing caspase-3 activation assayed with the accumulation of its cleaved substrate, PARP.

Bioinformatics Analysis[2]

Bioinformatics analysis was performed in collaboration with Dr. Derek Huntley of the Bioinformatics Support Service, Imperial College London. Comparison of the mouse candidate list with disease and cancer associated microarray expression profiles were performed by Sarah Langley and Dr Enrico Petretto of the MRC Clinical Sciences Centre, Imperial College London.

The Ensembl gene IDs for each of the assembled sequences were retrieved from the Ensembl peptide database (NCBIM37.58) using BLAST [307, 308]. Further data-mining was performed using the derived gene ID, retrieving information such as the gene names, UniProt protein identifiers, InterProScan motif/domain [309] and the GO annotations [310] including the cellular localisation, molecular functions and biological processes. Information on the signalling pathways was retrieved from the PANTHER website (v7.0) [311, 312].

21 of the sequences were found to be redundant. A further 6 clones were identified as potential candidate dependence receptors, having fulfilled the criteria of apoptosis induction, presence of predicated caspase recognition sequences and presence of a DART motif common to dependence receptors [181]. Caspase substrate and dependence receptor predictions were performed using the CASVM Server 1.0 [313] and Dependence Receptor Database [181] respectively. The potential dependence receptor clones were excluded from the analysis to investigate them in collaboration with Professor Patrick Mehlen, National Scientific Research Center (CNRS), Lyon, France.

Transporter and catalytic function clusters were the most significant association between the validated pro-apoptotic genes, with 43 genes identified within these clusters. This was followed by previously established genes involved in apoptosis, signalling factors and genes without known functional annotation.

The established pro-apoptosis genes formed approximately 9% of the validated candidates, with seven unique genes: Mitochondrial carrier homolog 1 (Mtch1), cathepsin L, glutathione peroxidase-1, Rhob, Itm2b, Fam82a2/PTPIP51 and Fis1. Mtch1 is a transporter localised to the mitochondria and is involved in apoptosis by presenilin-1 in a Bax/Bak dependent manner [314-316]. Itm2b has been described extensively as a BH3-only protein and promotes apoptosis activation in a p53-independent manner [317, 318]. Overexpression of Fam82a2 was previously published to induce apoptosis [319]. Fam82a2 is also a substrate to a host of other enzymes include protein kinase A, Src, PTP1B/TCPTP protein tyrosine phosphatases and is known to interact with Nuf2, 14-3-3 proteins, diacylglycerol kinase alpha and CGI-99 protein. This implicates Fam82a2 in various signalling cascades including apoptosis. However, the precise function of Fam82a2 has yet to be discovered, but it may involve

the calcium release coordinated with VAPB [320, 321]. The involvement of cathepsin L, glutathione peroxidase-1, Rhob and Fis1 in apoptosis has also been previously established [322-325].

The proteins of the isolated genes were found to be localised across all major cellular compartments, a finding which was in line with studies proposing that sensors were present in each of these compartments to activate apoptosis resulting from excessive damage or stress [326, 327]. The mitochondrion, a major apoptosis signalling hub, was also found to be a major component where the gene products localised to. 20 proteins were localised to the mitochondria, 16 of which were novel genes previously not associated with apoptosis. This indicates that signalling via the mitochondria may be more complex than the major established pathways. The endoplasmic reticulum (ER), another key organelle involved in apoptosis regulation primary via ER stress and calcium release, was another prominent cellular localisation with 13 genes. It was hypothesised that transporters found to localised to the ER such as Napa, Slc37a4, Surf4 and Yipf5 may facilitate the release of calcium or other pro-apoptotic factors directly or indirectly. It is interesting to note that only five of the genes were localised to the nucleus, which may hint that the nuclear control of apoptosis may be vested with a few master regulators such as p53, with most other apoptosis signals initiated at the protein level elsewhere which may offer a faster response.

Transmembrane proteins were also found to be a significant portion of the positive genes, many of which are transporters that may indirectly affect apoptosis signalling by altering the cellular homeostasis, or being involved in complex formation or protein sequestration such as the proposed involvement of Mtch1 in the permeability transition pore [314].

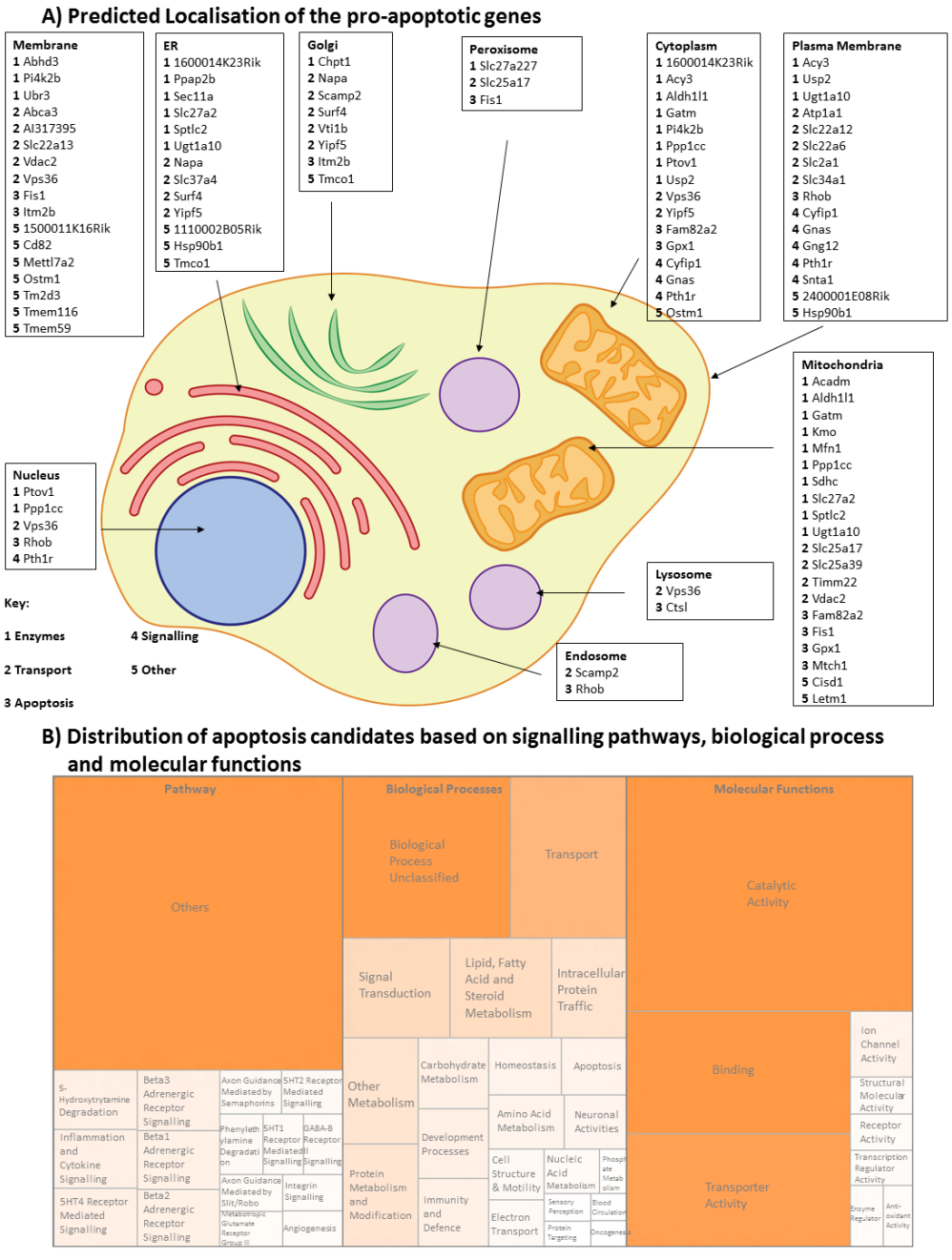


Figure 23: Bioinformatics analysis of the mouse pro-apoptotic candidates. A) The localisation of each of the pro-apoptotic genes. B) The signalling pathway, biological process and molecular function clusters derived using the PANTHER Classification System were presented as a tree map. A significant proportion of the genes were found to possess catalytic and transporter activities.

Determining signal thresholds and candidate selection criteria

Triton Signal Threshold

The CPRG ratio is calculated by Equation 1 using two absolute signals at equilibrium. The CPRG signal (before lysis) estimates the level of apoptosis, while the Triton signal (after lysis) gauges the total level of transfection.

A plate of the substrate CPRG in the absence of the reporter enzyme was measured at 590nm as the blank values. The CPRG ratio of each blank well was then calculated using randomly computer generated Triton signals between the ranges of 0.1 to 2.0. The sample size of each signal range was 96 and the mean CPRG blank value is 0.238 ± 0.012. The CPRG ratio at 0.0 - 0.1 Triton signal range was 6.89 ± 6.11, decreasing to 1.66 ± 0.38 at 0.1 - 0.2 and 0.527 ± 0.040 at 0.4 - 0.5, before stabilising under 0.3 at Triton signals of 0.9 or above.

A Triton signal of 0.5 represents a 100% increase in signal against the blank and indicates a well is transfected. However it generates a CPRG ratio above 0.5, which still classes the false positive well as a positive candidate with approximately 67.71% probability of this error occurring. This false positive rate increases to 100% below Triton signal of 0.4, since the diminishing transfection reduces the enzymatic activity to zero causing the Triton signal to be equal to the blank. Above a Triton signal of 0.6, the false positive rate resulting from the assay design reduces to zero. Since the estimated fluctuation in dataset automated by the robots is between 10 – 15%, a Triton signal of 0.7 was selected as the minimal threshold a data point has to achieve prior to analysis. Any values below 0.7 are excluded as the transfection was too weak to generate a signal that sufficiently removes the false positives.

This threshold was further emphasised in a separated automated control experiment with 192 sample size for the luciferase and RIP controls. Below the threshold of 0.7, the negative and positive controls become indistinguishable (Figure 24A). Within the same control, the two sets of 96-well plates resulted in similar range of CPRG and Triton signals. Furthermore, Figure 24A indicated that only 4 / 384 (1.04%) samples transfected resulted a Triton signal lower than 0.7, reasserting that the threshold is sufficient to identify transfected samples.

In general, the level of transfection correlates with the intensity of the Triton signal. The higher the Triton signal, the better the transfection and consequently improves the quality of the data point.

CPRG ratio for pro-apoptotic candidates

The luciferase, caspase-2 and RIP controls were isolated and transfected using the robots under the optimal conditions. The negative control luciferase has a CPRG ratio of 0.35 ± 0.079, while the positive controls caspase-2 and RIP are 0.60 ± 0.12 and 0.91 ± 0.04 respectively. Due to this linear separation of the controls, a fixed threshold may be employed for the selection of pro-apoptotic genes.

Caspase-2 was used as the weak control with its mean CPRG ratio between 0.6 – 0.7. By reducing the mean by approximately 10%, the minimal threshold was determined at 0.55. This threshold allows selection of the lower range of the caspase-2 signals, while allowing for slight increases in toxicity as result of ectopic expression.

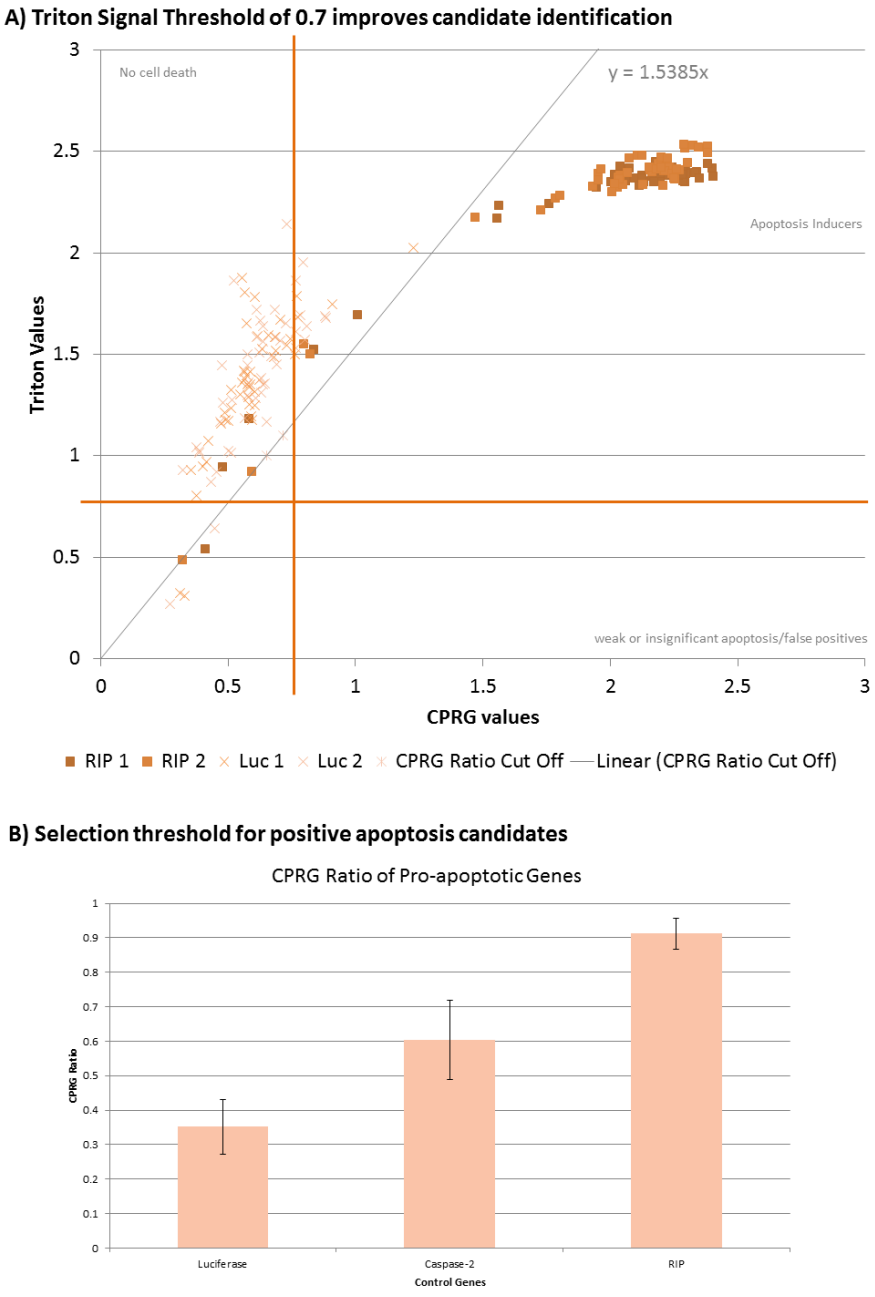


Figure 24: Thresholds for Transfection and Apoptosis determined using datasets derived from HEK293T cells. A) Distribution of the CPRG and Triton signals from the positive and negative controls. The gradient between the two signals is the CPRG ratio, and the pro-apoptotic CPRG ratio threshold of 0.55 is represented by the grey line. The orange horizontal line represents the 0.7 threshold for the Triton signal. N = 192 per control type. B) The CPRG ratios of luciferase, caspase-2 and RIP. Caspase-2, the weak apoptosis control, is used as the minimal threshold for apoptosis. N = 96, error bars represent standard deviation of sample size.

Indicators for selecting inhibitors and sensitizers of apoptosis

The attempt to identify inhibitors and sensitizers of apoptosis in addition to the pro-apoptotic inducers is a novel aspect of this screen. Inhibitors are defined as genes which supress cell death based on a reduction or maintenance of the CPRG ratio after treatment. Sensitizers are genes which do not induce apoptosis upon overexpression, instead predisposing the cellular system to an external trigger. In this screen, this external trigger is the treatment with the silica nanoparticles.

Unlike the CPRG selection criteria for the apoptosis inducers which is on a linear scale using a single CPRG ratio, selection of inhibitors and sensitizers may only be possible when non-treated and treated sets of data are analysed relative to each other. Such relationship causes the assay to assume a quadratic distribution, where the negatives are located within the central range while extremely negative and positive values indicate the presence of inhibitors and sensitizers respectively. This makes the direct comparison of the CPRG ratio of an external negative control such as luciferase or GFP with the candidates sub-optimal, since the overexpression of each cDNA is expected to modify the cellular system differently leading to different baselines. This dictates the use of relative selection criteria to determine the inhibitor and sensitizer candidates, rather than the fixed threshold used for inducer selection.

Two indicators were examined to quantify this relationship. First, the Treated-Non-Treated (TNT) ratio is calculated by putting the CPRG ratios of the treated and non-treated conditions into relationship as in Equation 8:

Equation 8: TNT Ratio

$$TNT\ Ratio = \frac{CPRG\ Ratio\ of\ Treated\ Condition}{CPRG\ Ratio\ of\ Non-Treated\ Condition}$$

The TNT ratio uses the rate of change between the treatment condition and the non-treated condition to measure the effects of apoptosis inhibition or sensitization. The faster the CPRG ratio is changed under treatment relative the non-treated baseline, the more likely that the gene is sensitizing the cells to apoptosis. Conversely, if there is little or no change in the CPRG ratio after treatment, then the gene is likely to be inhibiting or delaying cell death.

GFP was used as the overexpression control while mIκBa was used as the inhibitor control. 96-well plates of each control were transfected, and subjected to silica nanoparticles treatment previously optimised at 40ug/mL for 24 hours. The calculated TNT ratio for GFP was 1.512 ± 0.319, with mIκBa at 1.157 ± 0.290. This represents a 23.18% reduction in TNT ratio by the inhibitor control under treatment, with a t-test p-value of $8.06912E^{-14}$.

An alternative method of calculation using the TNT Difference (Equation 9) was proposed by Dr. Enrico Petretto of the MRC Clinical Sciences Centre, Imperial College London.

Equation 9: TNT Difference

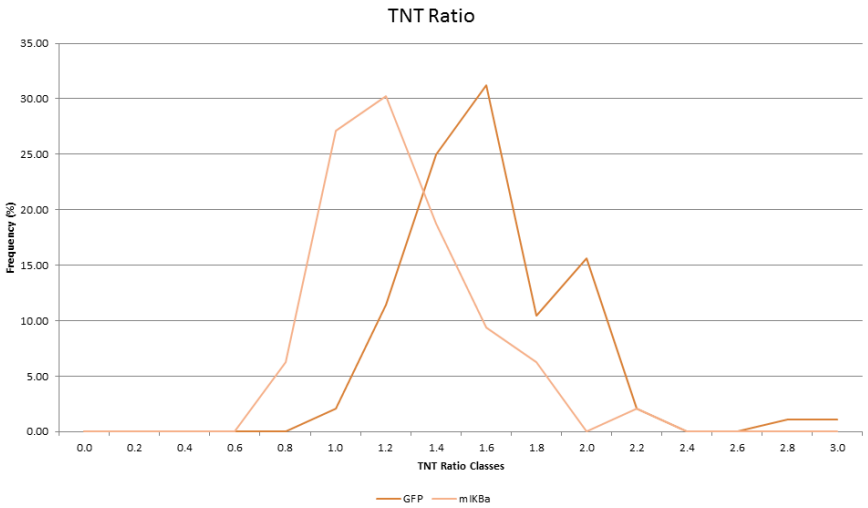
$$TNT\ Difference = CPRG\ Ratio_{Treated} - CPRG\ Ratio_{NonTreated}$$

The TNT Difference (TNTD) quantified the changes in cell death using absolute difference between the two CPRG ratios, rather the rate of change. Using the same experimentally generated dataset, the TNT Difference of GFP and mIκBa were 0.204 ± 0.096 and 0.052 ± 0.107 respectively. The TNT Difference indicated a reduction of 74.51% with a t-test p-value of $3.90973E^{-20}$.

The change in method of calculations to TNT Difference led to a three-fold increase in sensitivity, while the p-value was improved by a factor of 2,063,855.80. Instability of the TNT ratio which may cause the indicator to fluctuate more wildly leading to multiple peaks (Figure 25A) was also improved using the TNT Difference.

Hence, it was determined that the TNT Difference was sufficient to identify inhibitory response from the ectopic protein expression based on the controls. While a control for sensitizer was not available, it could be hypothesised that a similar effect is likely to occur at the opposite spectrum, where a significant increase in TNT Difference relative to the GFP baseline is indicative of a sensitization effect.

A) Distribution of control TNT Ratio



B) Distribution of control TNT Difference

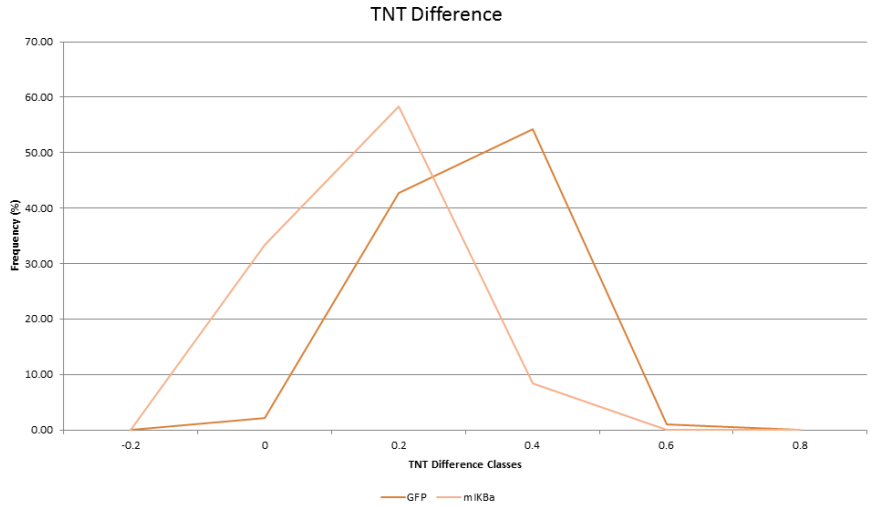


Figure 25: Comparison between the TNT Ratio and TNT Difference as the indicator for changes in cell death using datasets derived from HEK293T cells. A) Frequency (%) of the TNT Ratios between the overexpression and inhibitor controls across classes with 0.2 unit increment. Instability of the indicator is observable with multiple signal peaks at different ratios, and small proportion of extremely high ratios for both controls. N = 96. B) Frequency (%) of the TNT Difference between the overexpression and inhibitor controls across classes with 0.2 unit increment. The TNT Difference improves the assay sensitivity by three-fold, resulting in a smooth distribution of signals across a narrow range. N = 96.

DWE as the measure of central tendency

It was discovered that the dataset generated during the screen fluctuated between different experiments. This could be due to minor changes in final nanoparticles concentration resulting from pipetting errors, or variation in response of the cells among other reasons. Hence, an absolute TNT ratio or difference calculated in one experiment may be of another magnitude in another. For example, in an experiment where the cells are more sensitive to nanoparticles treatment, more sensitizer candidates will be identified using a "fixed" threshold of the control. As this was not anticipated in the setup, insufficient control counts prevented the accurate calculation of a baseline for determining the relative toxicity baseline for normalisation.

However, the central tendency of the whole experiment could offer a solution to the baseline calculation. By taking the reasonable assumption that candidates are randomly distributed across the entire NITE library, and not organised based their apoptosis sensitivity, the experiment population will follow a normal distribution. The true central tendency will form the most appropriate baseline toxicity as a result of the nanoparticles treatment.

Standard calculations such as mean or median are not accurately weighted and may be disproportionally skewed by outliers. Since these outliers are the inhibitor and sensitizer candidates, non-weighted skewing will reduce their significance. The statistical approach of trimming these outliers prior to calculation is also not ideal since it introduces biasness. The most appropriate measure of central tendency for the screen will be one based on where most of the data points are clustering, which could be calculated by proportionally reducing the weights of large outliers. The Distance-Weighted Estimator (DWE) calculated in Equation 10 offers this advantage.

Equation 10: Distance-Weighted Estimator

$$weight, w_i = (n-1).\sum_{j=1}^{n} abs(x_i - x_j)^{-1}$$

$$DWE = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3 + w_n x_n}{w_1 + w_2 + w_3 + w_n}$$

The merits of the DWE such as its robustness in the presence of outliers versus classical measures such as the arithmetic, harmonic and geometric means, and data trimming approaches were discussed by Dodonov *et al.* [328]. The DWE was used to estimate the central tendency of each experimental population of data points and thus the baseline on which the inhibitors and sensitizers was identified.

Selection of Inhibitor and Sensitizer Candidates

Unlike inducer candidates which may be easily selected based on a fixed threshold of the CPRG ratio, inhibitor and sensitizer candidates have to be determined relatively due to the greater amount of experimental fluctuation. Previously, the use of DWE to estimate the baseline toxicity of each experiment set was explained. The TNT Difference (TNTD) threshold for mIκBa, GFP and sensitizer set at 0.05, 0.20 and 0.45 respectively; mIκBa and GFP thresholds were determined experimentally previously while the sensitizer threshold was forward-estimated based on the other two threshold. These thresholds are used as the references for calculation.

Under the optimised screening condition, the set DWE is expected to be near the GFP threshold, allowing for both the inhibitors and sensitizers to be equally selected. In hyper-sensitive conditions where the set DWE approaches the sensitizer threshold, the quality of sensitizer selection is diminished and only the extremely effective sensitizers may generate sufficient signal to qualify selection. However, selection of inhibitors will be improved since genes which prevent cell death under these hyper-sensitive conditions will be likely to possess certain inhibitory functions. The reverse will be true under inhibitory conditions where the cells are not responding sufficiently to treatment-induced cell death. Here, the emphasis will be on sensitizer selection. Under optimal conditions, the 15% of the genes at each tail will be selected as inhibitors and sensitizers, with a higher proportion selected from the opposite end as the dataset become skewed.

A linear estimation was first employed based on the gradients or rate of change of the controls, with the inhibitor and sensitizer gradients at 218.75 and 145.83 times DWE. However, this linear estimation may cause the selection to become bias, since DWE baselines around the GFP threshold becomes over- or under-estimated too early. The change in selection proportion is expected to assume a sigmoidal behaviour, which remains fairly stable around the GFP threshold but rapidly emphasised the opposition proportion as the baseline is skewed.

The weighting factor for determining the sigmoidal proportional selection was determined by trial-and-error, cumulating in Equation 11:

Equation 11: Selection Gate Weighting Factor

$$Gate\ Weighting\ factor = DistGFP\ \times\ |\ DistGFP\ |^{DistGFP}\ \times\ 10$$

$$Inhibitor\ Proportion = \frac{1}{e^{-Gate\ Weighting\ Factor} + 1}\ \times\ 30$$

Where DistGFP is the distance between the DWE baseline and the GFP threshold. The absolute DistGFP determines the degree of deviation from GFP threshold, while the positive or negative signs of DistGFP determine the direction of selection.

A standard sigmoidal fit (Figure 26) will cause no clones to be selected once the DWE baseline past either inhibitor or sensitizer thresholds. However, while the significance of selection may be reduced, it does not necessary mean that the set will contain no sensitizer at all. Hence, an 85% selection limit was built into the fit, allowing some of the insignificant clones to be selected at the opposite tail. Since a total of 30% of clones will be selected from each experiment set as inhibitors or sensitizers, this limit translate to approximately 4.5% selection of insignificant candidates in the worst scenario.

A normalised score for each gene across different experiments may be further calculated using the same principles with Equation 12:

Equation 12: Gene Score Normalisation

$$Normalised\ score = \left(\left(TNTD_{gene} - DWE_{set}\right) \times \left(DWE_{set} - Threshold_{GFP} + 1\right)\right)$$

Under the optimal condition, the normalised gene ranking score will be its TNT Difference less the DWE baseline since scaling factor will be around 1. Negative values indicate inhibitors, while positive values are sensitizers. As the dataset becomes skewed, the gene score is proportionally scaled according to the amount of under- or over-sensitivity. A negative scaling factor indicates inhibitory condition, causing any negative gene score indicative of inhibitors to be scaled towards the central region. The reverse will also be true for the sensitizers. This allows the data points from different experiments to be combined into a single list, which is important for enrichment analysis such as GSEA.

Equations from this section were refined using valuable contributions from Dr. Enrico Petretto and Professor Stefan Grimm. Dr. Petretto also contributed greatly to the validation of Equation 12 which formed the basis of data normalisation and ranking for subsequent analysis.

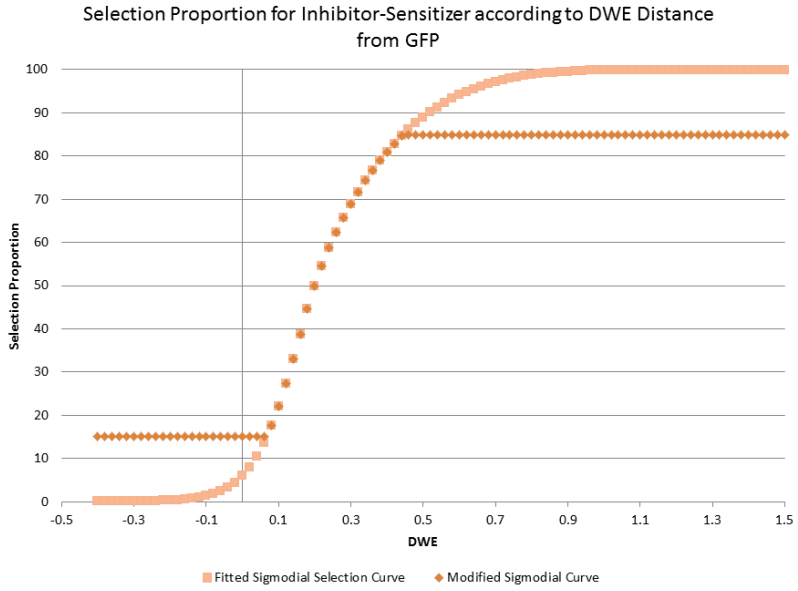


Figure 26: Determining the Inhibitor-Sensitizer selection proportion. A sigmoidal fit was most appropriately for determining the selection proportion for inhibitor and sensitizer. Inhibitors and sensitizers were equally selected under optimised conditions where the baseline centres on the GFP threshold. As the baseline approaches the inhibitor or sensitizer threshold, a greater proportion of candidate will be selected from the opposite tail. Control thresholds were experimentally determined in HEK293T cells.

Primary Screen

Workflow Management

Two runs of DNA isolation was performed daily with a throughput of 1,536 samples while transfections were performed in duplicates thus processing 3,072 samples daily. Both processes were performed simultaneously to maximise workflow efficiency. The entire transfection process was automated and separated into three phases involving the preparation of transfection mix, introduction of transfection mix and change of media after incubation. The reagents and plates were setup on the TFA platform by the user, after which the robot fully managed each process.

This "setup-and-go" performance allowed the simultaneous run of the semi-automated DNA isolation process, which requires user involvement in DWP transfer and media removal steps for increased efficiency. Both processes were tightly interlaced so that the incubation time in one process corresponds to the user involvement steps in the other. DNA isolated were transfected the next day, minimising degradation of the plasmid over time.

One set of the transfected cells was subjected to the silica nanoparticles treatment 24 hours post-transfection, the process of introduction of media with or without the nanoparticles was automated on the TFA platform. The CPRG assay involving first the addition of the CPRG substrate followed by the addition of 1% triton lysis buffer after incubation and measurement was also automated on the TFA platform. The robot manages the liquid handling steps, while measurement at equilibrium was manually accomplished.

The workflow is summarised in Figure 27. Briefly, Day 1 of screen initiation involves mainly the DNA isolation process and transfection if previously isolated plasmids were available. 24 hours post-transfection on Day 2, the treatment of cells with nanoparticles were included into the workflow, while 48 hours post-transfection on Day 3, the transfected and treated plates were subjected to the CPRG assay. The workflow reaches the maximal on Day 3, and further processes were repeated according to the Day 3 workflow until the end of the screen. At the peak, the TFA platform will be actively processing plates in the order of transfection, treatment, assay and media change, using near 100% of its capacity. The BASY platform uses approximately 65% of its daily available runtime.

All chemical reagents were purchased prior to the screen, and working solutions required for DNA isolation and transfection prepared on a weekly basis. Other bulk purchases include plastic consumables such as the microtitre plates, were acquired every fortnight due to the lack of storage space. All supplies for the screen were actively managed such that the required quantity arrives as the previous stocks become exhausted.

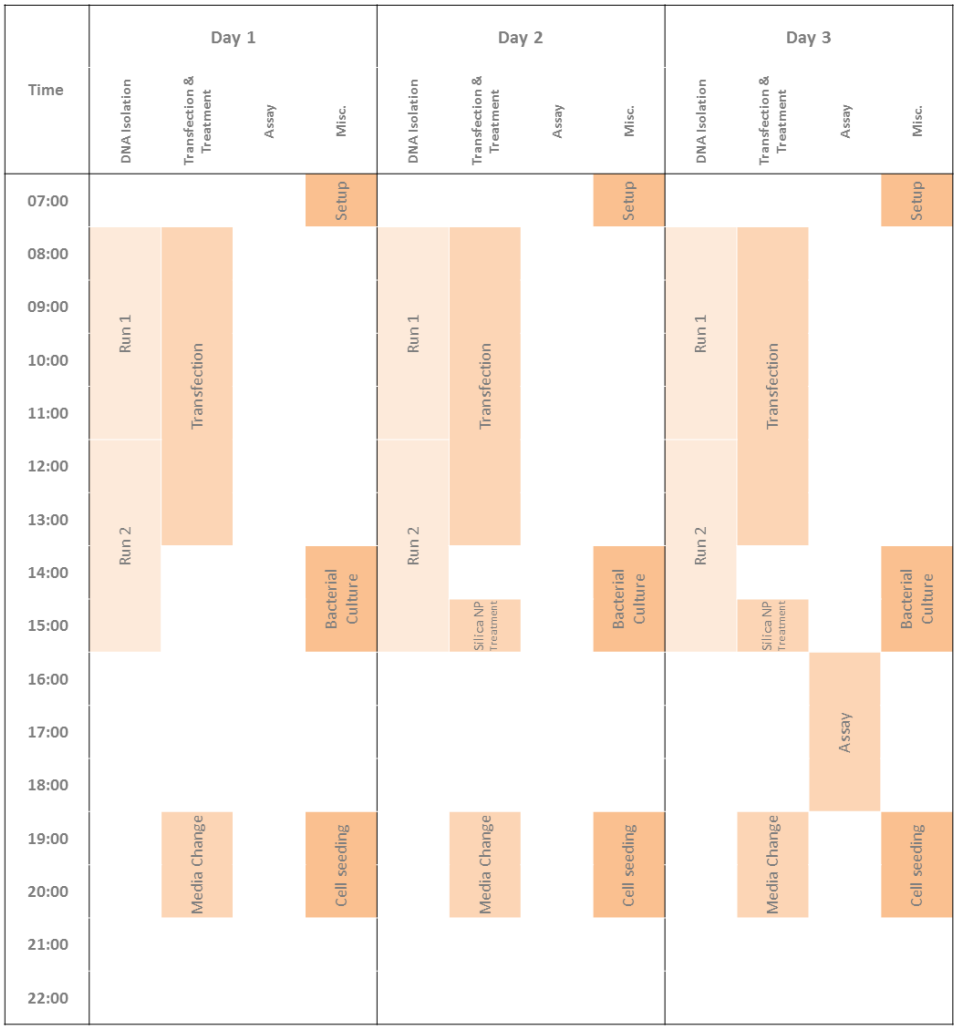


Figure 27: Daily workflow for the NITE Screen. Two DNA Isolation runs and four transfection runs were included as part of the standard screen workflow. Steps involving treatment of cells with nanoparticles and assay were included on the second and third day after screen initiation, after which the screen workflow repeats the intensive setup of Day 3 until the end of the high-throughput screening process.

Trial Screen

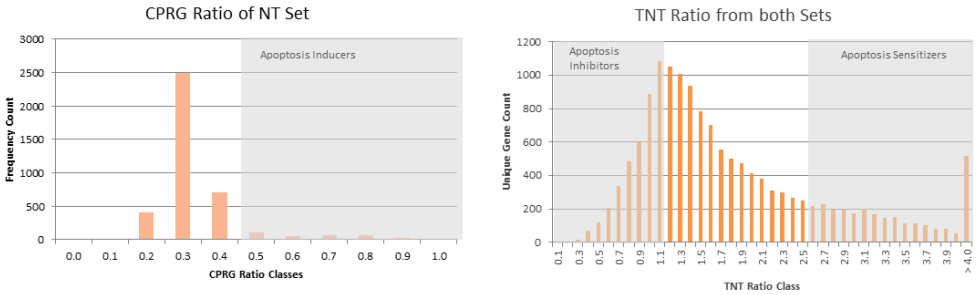
Prior to intensively screening of the NITE library, approximately 14.08% or 4224 cDNA clones from the library were screened. The results were concisely analysed to determine the quality of the screen and candidates.

The non-treated (NT) baseline set consists of 3,919 (92.78%) good quality data points which were above the Triton threshold, with a mean Triton signal of 1.31. The proportion of successful transfection in both NT and treated sets decreased slightly to 89.61%. The mean CPRG ratio of the NT set was 0.285, indicating that the transfection process and subsequent overexpression did not trigger excessively toxicity. 54 cDNA (1.45%) were identified as inducer candidates from the NT set in this trial screen. 757 (20.31%) cDNA were identified as inhibitor candidates based on the TNT ratio, with the mean TNT ratio of the screened library at 1.87 (Figure 28A). The TNT Difference indicator had yet to be discovered at the time point of the trial screen.

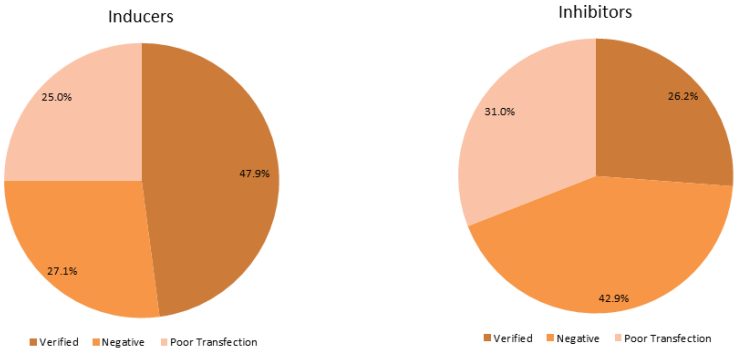
24 of the inducers and 21 of the inhibitors were subjected to repetitive validation based on the CPRG assay (Figure 28B). Unsuccessful transfection accounted for 25% and 31% of the inducer and inhibitor validation, the increase likely to be due to the small sample size. Excluding the non-transfected samples, the inducers were accurately validated with 63.9% of the candidates remaining positive. Therefore, the selection of candidates using a single CPRG ratio value remains fairly accurate. This accuracy was reduced to 37.9% when inhibitors were selected using the TNT ratio, which suggested that the indicator is identifying a high count of false positives.

Both candidate lists were also analysed using DAVID [329] to determine if any major changes were observable in the preliminary dataset (Figure 28C). The most striking difference was found in the localisation GO Terms, where membrane-associated proteins comprised a significant portion of the inducer and sensitizer candidates (34.03% and 33.24%). Inhibitor candidates however displayed higher localisation to the nucleus (16.49%), with membrane association reducing to 15.16%. This finding suggested that sensitizer candidates may be weak inducers which are insufficient to activate apoptosis under the screen conditions. It also highlighted potential differences between the pro- and anti-apoptotic candidates, suggesting that inhibitory response may be more associated with transcriptional activity while apoptosis activation is regulated at the protein level.

A) CPRG Ratio and TNT Ratio of trial screen



B) Candidates were independently validated in a duplicate secondary screen



C) Distribution of candidates based on cellular localisation

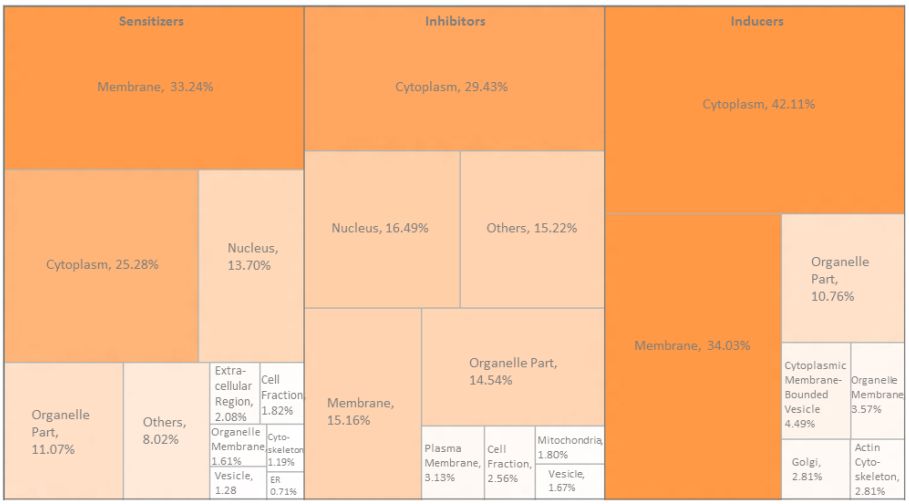


Figure 28: Analysis of the Trial Screen. A) Approximately 14% of the NITE library was subjected to the screen using HEK293T cells and the results analysed. Inducers were identified using CPRG ratio while Inhibitors and Sensitizers were identified using TNT Ratio according to previously determined threshold. B) The identified candidates were subjected to a secondary screen performed in duplicates. 47.9% of the inducer candidates and 35.7% of the inhibitor candidates were positive in the secondary screen. C) The candidates were concisely analysed based on their Gene Ontology terms, with their respective cellular localisation displayed in this treemap. Inducer and sensitizer candidates consist of a significant proportion of membrane proteins, which is reduced in Inhibitor candidates, hinting at potentially successful selection criteria.

Complete Screen of the NITE Library

The primary screen was completed in 38 experiment sets in approximately eight weeks with 25,132 good quality data points generated giving library coverage of 83.77% in a single pass.

The mean CPRG ratio of the non-treated set was 0.279 ± 0.081, which was similar to those calculated in the trial screen. Using the non-treated set, the inducers were first selected with the previously determined CPRG ratio threshold of 0.55. This yielded 377 (1.50%) clones which were subjected to further validation. This selection proportion was similar to those of the trial screen. It was also interesting to note that this selection threshold was just above the 3 sigma range (3 units of standard deviation) calculated at 0.522, reaffirming that these positively selected inducers are likely to be positive (Figure 30). cDNA clones identified as apoptosis inducer candidates were excluded from the inhibitor-sensitizer selection (Figure 29).

The sensitivity of the CPRG assay to silica nanoparticles induced cell death was apparent. Although the treatment set was subjected to the same automation process as the non-treated set, subtle variation between experiments was sufficient to result in a major shift resulting from the secondary treatment factor. This caused some experiments to have a higher toxicity baseline, shifting the data points into two peaks rather than one as observed in the non-treated set. Normalisation of the dataset with Equation 12 did generate a normal distribution (Figure 30B).

TNT Difference (TNTD) was used to select for inhibitors and sensitizers, along with the calculation of selection proportion previously discussed. Excluding transfections which were successful in only one set, the primary screen with the treatment had library coverage of 83.40%. The selection resulted in 3,936 (15.73%) inhibitors and 2,759 (11.03%) sensitizers, which were subjected to further validation (see below).

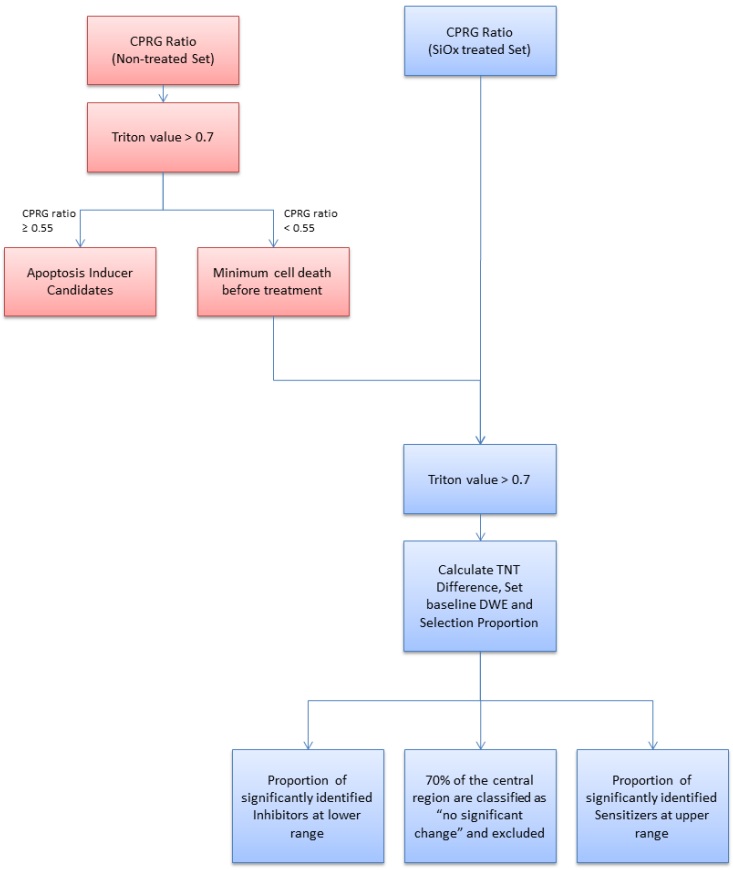
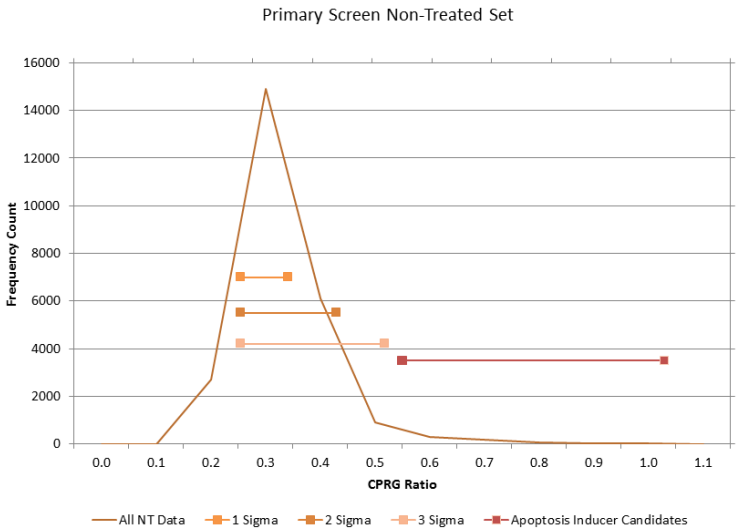


Figure 29: Candidate selection algorithm. The primary screen resulted in two datasets, the non-treated and nanoparticles treated sets. Apoptosis inducer candidates were first selected from the non-treated set using the previously determined CPRG ratio and Triton thresholds. These candidates were excluded and the remaining data points used for selection of inhibitor and sensitizer candidates. The TNT difference was next calculated with the extreme positive outliers in the upper range selected as sensitizers and extreme negative outliers in the lower range selected as inhibitors. The combined inhibitor and sensitizer candidates represent 30% of the data points, with the remaining 70% classified as "no significant change" and excluded from the secondary validation screen.

A) Distribution of CPRG Ratio of the non-treated set



B) Distribution of TNT Difference and Normalised scores for Nanoparticles treated set

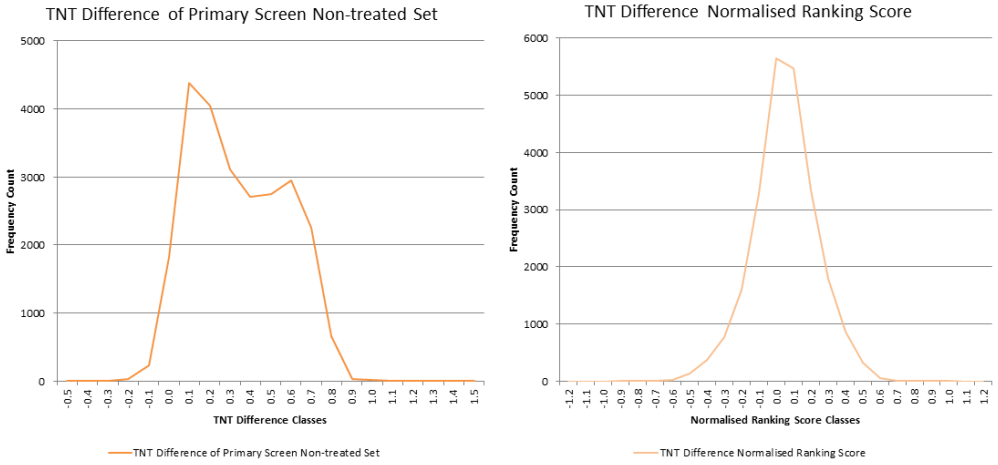


Figure 30: Primary screen data generated using HEK293T cells. A) Distribution of the CPRG ratio in the non-treated set. Previously determined selection criterion of 0.55 for apoptosis inducer was more than three standard deviation from the mean, indicative of a good quality selection. B) Pre-normalised distribution of the TNT Difference indicator (left) and the normalised ranking score (right). Inhibitor and sensitizer candidates were identified with a selection proportion calculated from the deviation of the experiment data from the GFP control threshold.

Validation Screen

The selected candidates were validated through replication. The 377 inducer candidates were independently transfected and assayed up to six times, with the mean CPRG ratio of all inducer candidates at 0.562, compared with the GFP controls of 0.341. This indicated enrichment for pro-cell death candidates. Using t-test, the p-value for each candidate was calculated and only those with p-value less than 0.05 were selected. 300 (79.58%) inducer candidates were finally selected to have CPRG ratios, which were significantly increased, compared to the GFP negative controls.

Since the selection of candidates as inhibitors or sensitizers is likely to result in enrichment, the assumption that clones are randomly distributed is no longer valid and prevents the use of DWE of each experiment set to calculate the toxicity baseline from each experiment set. Instead, a plate of GFP controls was transfected with each experiment and the DWE of this control plate serves as the baseline for toxicity.

Triplicate transfection was performed for the inhibitor-sensitizer validation screen, with two plates subjected to silica nanoparticles treatment and one plate as the non-treated baseline. This resulted in two TNTD values which were used to determine the significance from the GFP. Clones that were significantly decreased in the TNTD compared with the GFP DWE baseline were selected as inhibitor candidates, while those showing increased TNTD were classified as sensitizer candidates. Statistical significance was established using the t-test. 1,142 (29.01%) inhibitors candidates were selected with p-value <0.05, with the mean TNTD of 0.271 compared with GFP DWE baseline of 0.511 after silica nanoparticles treatment. Likewise, 626 (22.68%) sensitizers remained significantly increased with mean TNTD of 0.552 against a GFP baseline of 0.412.

Interestingly, additional inducer candidates were detected in the inhibitors and sensitizers shortlisted for validation, 268 (6.81%) and 64 (2.32%) additional inducers respectively. This is likely due to the increased transfection efficiency from the use of the linear PEI transfection protocol in the validation screen, which resulted in higher expression levels of the plasmids. The proportion of positive inducers resulting from the validation screen on the inhibitors and sensitizers is also higher than the 1.50% identified in the primary screen. This suggests that the proportion of the inducers may be higher and hence the clones which are currently within the “grey zone” between the current selection criteria and the negative control baseline may be potential inducers.

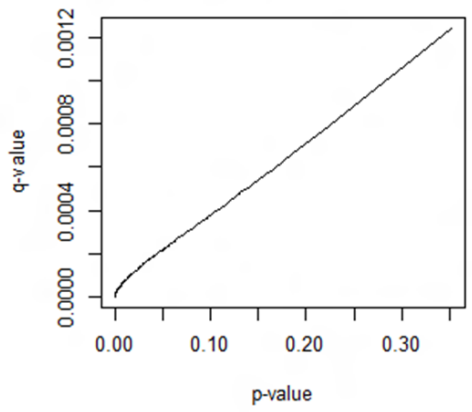
Since the validation screen involved testing of a fairly large number of clones, some of these clones may be statistically significant for the p-value threshold by chance. The false discovery rate (FDR) was calculated for the set of p-values generated for each candidate type in order to correct for multiple testing [330].

The FDR (or q-value) was calculated using the R statistical package and the QVALUE extension [331, 332]. The FDR was used to determine the significance of the p-value selection criteria, and the FDR values for each candidate type are presented in Figure 31. The FDR values appeared to be extremely low (under 0.01) for all candidate types and across a wide range of p-value of up to 0.5. This phenomenon of low FDR values was likely due to the pre-selection step whereby only the significant candidates from the primary screen were selected for validation in the secondary screen. Only repetitions were performed in this secondary validation screen to generate the p-values. It is likely that this selection step had enriched for candidates that were already significant from the baseline, and hence the range of the resulting p-values which in turns translate into low q-values.

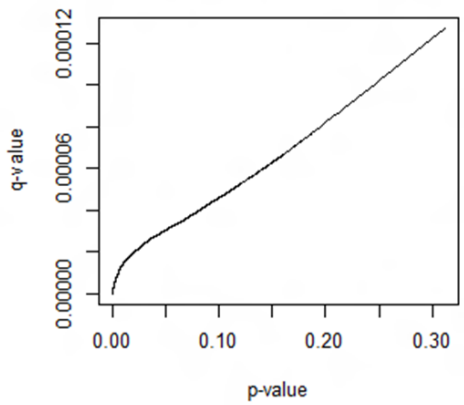
This indicated that the selection thresholds that were established experimentally for each candidate types may indeed be sufficiently to differentiate the phenotype-of-interest from the background. It also indicated that the p-value threshold of 0.05 is sufficient to select for positive candidates that are statistically significant with minimal amount of false positives. The p-value may also be increased to include the other potential candidates although the more stringent p-value criteria of 0.05 is preferred in this scenario since only one experimental test was performed in the primary screen.

The validated candidates were subjected to downstream bioinformatics analysis. The apoptosis inducing candidates isolated in the validation screen were also included in a separate expanded list of inducer candidates.

A) FDR Threshold of Inducers



B) FDR Threshold of Inhibitors



C) FDR Threshold of Sensitizers

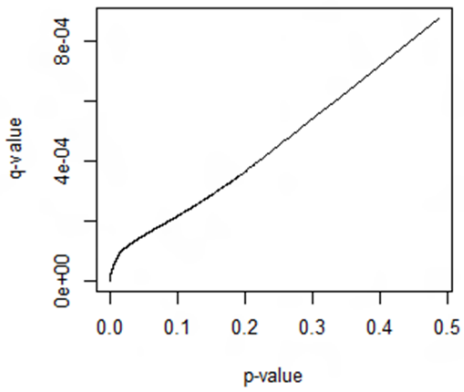


Figure 31: The FDR thresholds were verified for the significance of the p-values. The FDR values for the set of p-values derived from the validation were calculated using R (QVALUE). This is to determine the significance threshold as a result of multiple testing. It was revealed that the FDR values were very low for the p-values, and hence the selection of p-value significance 0.05 remains a validate criteria. The low FDR values are likely to be a result of the validation of selected candidates that were already significantly different from the baselines in the primary screen.

Discussion

The validation of the mouse cDNA apoptosis candidates served as a miniaturised validation process for the workflow downstream of the experimental workflow previously optimised in the setup phase such as DNA isolation, transfection and assay. The aim was to identify alternative methods for the verification of the pro-apoptotic activities of the newly identified clones and experiment with various approaches for bioinformatics analysis of the candidates. Unlike the NITE library which consists of cDNA clones which have been fully sequenced, the mouse cDNA library was established in-house and the identity of the candidates was unknown. Hence an additional sequencing phase was included to verify the identity of the clones.

The sequencing of the mouse cDNA candidates provided the foundation for downstream analysis and handling the sequencing phase which extended into hundreds of sequencing reads was another novel undertaking at our group. A one-pass sequencing strategy was employed to minimise the sequencing costs while the traditional approach of primer design was coded as a macro to facilitate the management of sequencing data. The macro proved to be crucial when handling the increased load of sequencing reads, since each batch of data could be feed into the macro which then quickly generated the lists of primers required for the next sequencing phase. Although this approach was performed for the first time, the approached proved to be efficient and generated high quality sequencing data. Most of the computer-designed primers resulted in high quality sequencing reads which are >500bp in length and >80% of these reads could be assembled without the need for additional data. However, 22 clones required additional processing by using a homologous sequence identified by NCBI BLAST to assemble into a single contig after repeated sequencings and manually designed primers. This is likely due to the presence of secondary structures such as hairpin loops in the sequences, which prevented the proper sequencing under the standard conditions employed by the sequencing centre.

The assembled contigs were made available to Dr Derek Huntley, who performed additional processing to retrieve the associated information such as gene name and ENSEMBL identification numbers. The identifiers for the sequences allowed for the removal of redundancy, and the non-redundant clones which were verified though using alternative apoptosis assays were subjected to further analysis. Bioinformatics analysis was performed in conjunction with Dr Huntley while comparison of the mouse candidate list with disease and cancer associated microarray expression profiles were performed by Sarah Langley and Dr Enrico Petretto. The bioinformatics analysis revealed that seven genes from the candidate list had a known role in apoptosis, reaffirming the

quality of the candidate list. The localisation of many of the candidates to the mitochondria and endoplasmic reticulum, which are traditional organelles associated with apoptosis further support the pro-apoptosis function of the candidates. This implicated the candidates, most of which are novel, in the regulation of apoptosis.

Differential expression analysis of disease and cancer associated microarray profiles also identified candidates which were consistently down-regulated in cancer profiles including the haemoglobin gene (Hba-a2) in breast, colon and prostate cancer profiles analysed. Expression of UDPGT and Tim22 were reduced in colon cancer while ALEX (Gnas) was down-regulated in breast cancer [284]. The overall expression profile of the candidate list indicated considerable heterogeneity, which may suggest the functional importance of the pro-apoptotic candidates in the altered cellular profile in the diseased state. Overexpression of genes with apoptotic activity and sensitizes the cancer cells were known to contribute to the cancer metabolism. For example, the oncogene myc which exhibits both proliferative and pro-apoptotic signals is frequently up-regulated in cancer [333]. Interestingly, Atp1a1 was also found to be down-regulated in the Alzheimer's disease profile analysed. Furthermore, using Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com) revealed that four of the candidates may potentially signal via the NFkappaB complex [284]. The collaboration with the bioinformatics and microarray specialists offered valuable exposure to the analysis approaches which was important for the analysis of the NITE derived candidate.

Efforts were also expended to explore alternative validation assays for the apoptosis activity, with the focus on high throughput formats, sensitivity and cost-efficiency. Assay sensitivity is an important factor since the RISCI screen uses an enzymatic assay which is sensitive to apoptosis or cell death in general. Utilisations of alternative assays which are less sensitive have the unintended consequences of invalidating a potential candidate. Here, the preferred assay of choice was the Roche Cell Death Detection ELISA^{PLUS} assay which is extremely sensitive for late-stage apoptosis detection. The increased sensitivity and commercial availability is compromised by a starkly increased consumable costs. Nonetheless the assay was applied to the mouse candidates (Figure 22A).

Assays detecting the activation of caspase-3, a well-established marker of apoptosis, were also tested. A variety of colorimetric and fluorogenic caspase-3 substrates as well as commercially available kits were investigated, including colorimetric caspase-3 substrate I

(Merck, 235400), fluorogenic caspase-3 substrate II (Merck, 235425), and caspase-3 assay kit (Sigma-Aldrich, CASP-3-C). These colorimetric and fluorogenic substrates lacked the sensitivity to determine caspase-3 activity in a 96-well cell-based format and no signal was detectable even for the RIP control. These generally required the use of substantial amount of cell lysate which could not be generated from a 96-well format, which may explain the lack of substrate conversion. Furthermore, caspase-3 as an effector caspase serve to activate the caspase cascade and its expression level or activity may not be sufficiently high for the required sensitivity. The Caspase-Glo® 3/7 Apoptosis Quantification Kit (Promega) was successful in detecting caspase-3 activation under compound treatment in 96-well format, but the signal was insignificant and unstable for transfections of pro-apoptosis genes. This is likely to be due to a lack of normalisation mechanism, which was built into the CPRG assay to account for the discrepancies generated by the transfection process. The sensitivity, costs and the lack of access to a luminescence compatible instrument limited the application of this assay. A FRET-based reporter for caspase-3 activation was also generated and proved to be able to detect caspase-3 activation under confocal imaging, but again failed to be translated into the 96-well format due to the limited sensitivity of the current instrument. The caspase-3 activation was eventually detected indirectly via the accumulation of its substrate, PARP. PARP is cleaved by caspase-3 [334] and the normal and cleaved form at 113 kDa and 89 kDa may be detected conveniently using western blotting and quantified using image analysis software such as ImageJ. The western blot approach was successful in detecting caspase-3 activation for the mouse candidates (Figure 22A). The western blot approach was more laborious but remained scalable to simultaneous processing of hundreds of clones within a manageable budget, the major consumable being the antibodies. The apoptosis activity of the candidates as detected by the commercial ELISA assay and western blot approach resulted in similar signal profile, highlighting the quality of both approaches. Furthermore, the majority of the candidates validated by the two alternative assays remained positive (approximately 90%), further highlighting the quality of the CPRG assay in detecting apoptosis signal as compared to other forms of cell death such as necrosis.

The validation screen for the mouse candidates served as an important learning point where the experience with the high throughput workflows were scaled and extended to include the management of the post-experiment processes such as validation assays, data handling and analysis.

The selection criteria and thresholds for candidate selection were also finalised prior to the execution of the screen. The traditional method of selecting for pro-apoptosis inducer candidates remained consistent and had become well established by this point. This selection method involved the use of a fixed threshold, which was experimentally determined to be 0.55 being approximately 10% lower than the mean CPRG ratio of the weak control, caspase-2. The inclusion of the Triton signal threshold at 0.7 also served to improve the quality of the selection, since any signal below this threshold increased probability of false positives as the Triton signal approaches the limit of the CPRG substrate reference value. An approximation of this effect may be observed in Figure 24A where the frequencies of the negative and positive controls remained equal, hence an approximately equal chance of any candidate identified without a Triton signal filter of being either a true or false positive. This pair of selection thresholds formed the basis of selection for inducer candidates in the primary screen, and the effectiveness of these thresholds was proven in the validation screen where up the successful inducer candidates were independently repeated for six times with about 80% positive rate.

The selection for the inhibitor and sensitizer candidates was approached in a similar manner, where a single CPRG ratio threshold was determined for the apoptosis inhibitor and negative controls after silica nanoparticles treatment. This approached had a major disadvantage since the CPRG ratio after treatment did not account for any potential toxicity before the treatment (the non-treated CPRG ratios). This introduced biasness within the selection, resulting in a greater probability that clones with a higher non-treated baseline being excluded from inhibitor selection (and being positive as a sensitizer). It also made it difficult, if not impossible, to establish a theoretical threshold for the sensitizers since we do not have any known sensitizer controls at the time of the screen.

This fixed and linear approach was eventually discarded in favour of the TNT ratio, which is the quotient of the treated CPRG ratio against the non-treated CPRG ratio baseline. The TNT ratio effectively corrected the treatment signal for any signal resulting from overexpression of the cDNA. The rate of change between each pair of CPRG ratios hence provided an elegant indicator to measure the effect of each cDNA on apoptosis; an anti-apoptotic effect occurs when the cDNA sufficiently slow the rate of change after treatment, resulting in little change or a reduced (less than 1.0) TNT ratio (TNTR), while the sensitizers could conveniently be observed on the opposite effect where the rate of change is increased. The TNT ratio was also statistically significant in differentiating between the inhibitor (mIκBa) and negative (GFP) controls. However, the compression of each pair of CPRG ratios as a result of the division minimised the effect of inhibition while expanding the range of the sensitizers, making the indicator bias and placed greater emphasis

on the lower decimal places of the experimental values. Thus the TNT ratio is good for candidates that exhibit inhibitory or sensitizing effect of a greater magnitude. This indicator eventually evolved into the TNT difference (TNTD), which was proposed following a consultation with Dr Enrico Petretto. The TNTD quantified the effect on apoptosis using the magnitude of change (the difference in absolute values of the CPRG ratio pairs) rather than a rate of change. The TNTD proved to be considerably more stable compared with the TNTR, and improved the statistical significance (p-value) by more than 2,000,000 fold. This resolved the choice of indicators for the inhibitor and sensitizer candidate selection, and the TNTD thresholds were experimentally determined to be 1.157 and 1.512 for the inhibitor and negative controls after nanoparticle treatment. The assumption of equal effect magnitude was made for the sensitizer candidates, setting the TNTD threshold at approximately 1.867.

However, implementation of the screen at high throughput format introduced further surprises not anticipated in the extensive setup and optimisation phase. As the actual screen progresses, it became increasingly clear that minor differences between experimental conditions were having an effect on the level of apoptosis, with experiment sets being pro- or anti-apoptotic as a whole. Any factor may contribute to this fluctuation in general including the cell culture conditions and the minute changes in the final amount of nanoparticles as a result of the normal pipetting errors. This problem was compounded by the loss of control values due to equipment technicality. The control plasmids for apoptosis were introduced in a separate pipetting step by the robot, which using a blast of air from increased pressure to eject the liquid into the wells. Since only six pipette tips of the 96-well dispenser were used for the controls, the pressure had an increased tendency to dissipate via unused pipette tips. The control dispenses correctly for the first few plates, but repeated dispensing steps caused minor accumulation of the control plasmid solutions in the six control tips to build over time and increasing the preferential dissipation of the pressure via the unused tips. Eventually, the control tips become blocked; this process introduced unintended fluctuation for the control values, rendering them useless.

By using the fixed threshold determined previously, the selection is now biased when all experimental sets were aligned. The pro-apoptosis set would naturally yield only sensitizers while the anti-apoptosis set would isolate only inhibitors when the fixed thresholds are applied. Furthermore, the results were accumulated as the screen progresses, creating yet another conundrum. The current set of results was derived from experiments performed over the past days and may for example indicate a slightly pro-apoptotic phenomenon. Applying these new assumptions experimentally would require upstream changes to be made to a completely different

set of experiment. This approach is not dissimilar to using the current share price to predict the movement of the future prices. The end result would be a mess of different experiment conditions made in an attempt to “guess” the desired baseline, and is likely to widen the fluctuations. The decision was made to keep the condition stable and address the deviations mathematically.

As a result, this caused the selection to be made in a dynamic approach to exploit the observation of the fluctuation between experiment sets. By taking the assumption that the NITE library clones are randomly distributed, the central tendency TNTD of all clones from each experiment set would provide an indication of the “natural” level of apoptosis within each set. Any appropriate measure could be used to quantify this central tendency, including the most common arithmetic mean or simply average. The arithmetic mean weigh each component equally, hence is easily scaled by “outlier values” away from the true central tendency. Alternative approaches involve the trimming of these outlier values prior to calculating the mean which was not appropriate either, since these outlier values of the experiment sets are the inhibitor or sensitizer candidates. The solution was found in the distance weighted estimator (DWE) where the magnitude of each value from every other value was used to weigh the contribution of the value towards the central tendency. The DWE allowed for the quantification of the central tendency, which in the screen setting approximate the level of apoptosis experienced within each experiment set. The selection strength of the assay for selecting sensitizers degrades as the sets become increasingly pro-apoptotic (higher DWE), but the significance of selection is shifted to the inhibitors. The reverse is also true for selection of sensitizers in low DWE sets. Hence the final selection isolates a greater proportion of the candidate at the opposite end to the DWE. The predetermined control values were used as the reference or baseline, and the DWE of each set of TNTDs used to scale each set of values for comparison, and this dynamic approach was eventually converted into Equation 12. The scaled set of values was then used for the GSEA analysis.

The implementation of the screen was impeded with various problems both experimentally and for the interpretation of the results, despite the effort spent on optimisation. However, these were eventually resolved.

The actual screen was executed in phases, scaling up from the trial phase which provided further indication in conjunction with the results from the mouse validation screen that the setup was sufficiently robust and accurate in identifying the phenotype-of-interest. The candidates selected from the primary screen were validated by repetition and subjected to further analysis.

Chapter 8: Analysis

Background Information

The apoptosis inducer candidates were analysed as two different gene lists: 1) Inducer Main List which comprised of the experimentally validated candidates and 2) Inducer All List comprised of the validated candidates from the Main List and pre-validated candidates identified from the secondary validation screens of the inhibitors and sensitizers. The inhibitor and sensitizer lists were analysis as a single combination each. The candidate gene lists were subdivided into two sub-lists each comprising of either protein coding or non-coding genes. Protein coding and non-coding genes were subjected to different analysis.

Protein coding genes referred to cDNA sequences which could be mapped to an experimentally verified or hypothetical protein sequences. A list matching the cDNA sequences to known genes with its corresponding Ensembl identifier was generated by Dr Derek Huntley from the Bioinformatics Services, Imperial College London. This was combined with an alternative list mapped by online databases using the db2db function of bioDBnet [335], part of which were manually verified to be similar to the BLASTx output generated by NCBI BLAST. The protein coding sub-lists were subjected to functional annotation by DAVID, Gene Set Enrichment Analysis (GSEA), Ingenuity® Pathway Analysis (IPA, Ingenuity® Systems, www.ingenuity.com) and GeneMANIA network analysis.

Any outstanding cDNA sequences not mapped were classified as non-coding sequences; this non-coding list was analysed for alternative regulatory mechanisms including non-coding regulatory RNA like microRNAs and potential of the open reading frames to translate into mini-proteins.

The analysis approach is summarised as Figure 32.

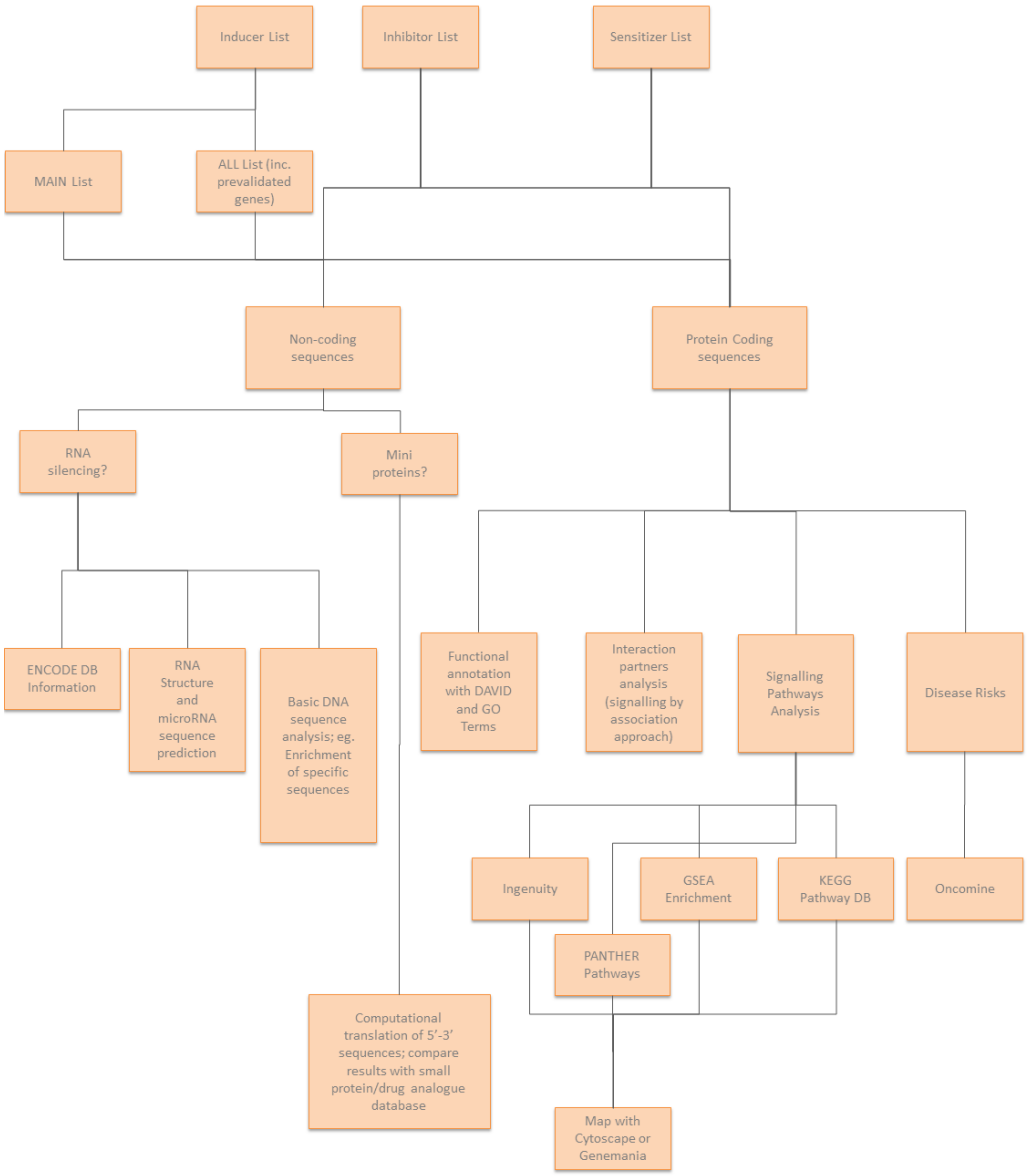


Figure 32: Analysis strategy. The gene lists were classified into protein coding and non-coding cDNA sequences. Protein coding sequences were analysed using functional annotation and enrichment approaches such as DAVID and GSEA while non-coding genes were analysed for the possibility of micro-RNAs or mini-proteins as the mode of action.

Results

DAVID Function Annotation

Analysis of the gene ontology (GO) terms [310] by the DAVID functional annotation tool [329, 336] revealed functional clusters enriched based on biological processes, cellular localisation and molecular functions.

The enriched biological processes GO terms for each candidate type were presented as a tag cloud (Figure 33), where the font size is scaled proportionally with the frequency counts. Both the "Inducer Main" and "Inducer All" lists were shown to be enriched for terms associated with apoptosis and cell death, the former comprising of the validated apoptosis inducer candidates and the latter inclusive of pre-validated candidates. Terms such as apoptosis (GO:0006915), programmed cell death (GO:0012501) and cell death (GO:0008219) indicated increase in frequency counts with the addition of the pre-validated candidates, although the specific term for induction of apoptosis (GO:0006917) was no longer enriched. The pre-validated candidates could possibly represent high pro-cell death regulatory diversity compared with the direct execution of programmed cell death. Transport-associated processes for various organic and inorganic ions including intracellular transport (GO: 0046907), ion transport (GO:0006811), L-glutamate transport (GO:0015813), organic anion transport (GO:0007010) and amine transport (GO001587) were also highly represented. The false discovery rate (FDR) of L-glutamate transport was improved from 19.86% to 8.11% with the inclusion of the pre-validated candidates, while other transport-associated terms like acidic amino acid transport, organic anion transport also have FDR of under 25%. Regulation of defence against virus (GO: 0002230 and GO 0050691) together with cytoskeleton organisation (GOL0007010) also had low FDR at 15.29%, 23.65% and 17.90% respectively. The use of apoptosis as a response to viral infection of the host is well established [337] and supports the identity of the inducer candidates, while the various transport functions may hint at the integral nature of metabolism and associated sensors to cross-talk with apoptosis signalling and activation in the event of any homeostatic imbalance.

Inhibitor candidates associated with the reduction of cell death signal induced by the LUDOX® silica nanoparticles were enriched for RNA processing terms with the general term RNA process (GO:0006396) highly represented with 37 counts and 7.84% FDR, followed by specific terms such as RNA splicing via transesterification reactions (GO:0000377 and GO:0000375) and nuclear mRNA splicing via spliceosome (GO:0000398). Uniquely, the inhibitors were enriched for response-associated processes with the response to metal ion (GO:0010038) represented with 11 gene counts and FDR of 10.18%. The response to metal ion was the second most significant terms ranked by FDR,

while response to inorganic substance (GO:0048872) was only 2.45% above the guideline FDR threshold used in this analysis. This suggests that the silica nanoparticles were inducing a similar effect to metal ions where toxicity of ions such as iron, copper cadmium and nickel was shown to involve oxidative mechanisms involve reactive oxygen species and depletion of antioxidants such as glutathione [338]. A general review of all biological processes enriched regardless of the FDR indicated enrichment of terms such as ubiquinone metabolic processes (GO:0006743 and GO:0006744) and quinone cofactor biosynthetic process (GO:0006730) which are associated with oxidation. Furthermore, inhibitor candidates associated with protein methylation and alkylation (GO:0006479 and GO:0008213), biopolymer methylation (GO:0043414) and histone methylation (GO:0016570 and GO:0016571) could perhaps help mediate any potential oxidative stress and manage the stress response. Methylation of non-histone proteins is important in regulation their stability and function [339] while histone methylation has also been established to contribute to transcriptional regulation [340]. Methylation could also potentially tag the proteins to stress granules which are large ribonucleoprotein particles to mediate the stress response [341]. Furthermore, the term negative regulation of protein complex disassembly (GO:0043242) may suggest the possibility of the inhibitor candidates maintaining important cellular functions by preventing the breakdown of protein complexes. Involvement of the inhibitors in estrogen receptor signalling pathway (GO:0030520), homeostasis of number of cells (GO: 0048872), erythrocyte differentiation and homeostasis (GO:0030218 and GO:0034101) further supports the pro-survival characteristics of this candidate type. For example, estrogen receptors are over-expressed in 70% of breast cancer incidents and are involved in disrupted apoptosis [342-344].

Intriguingly, the sensitizers were significantly enriched in biological processes associated with cell morphogenesis (GO:0000902), cell projection organisation (GO:0048858, GO:0032990 and GO:0030030), cellular component morphogenesis (GO:0032989), cell differentiation and development. These terms were below the guideline FDR threshold of 25%, with terms associated with cell morphogenesis, cellular component morphogenesis, cell projection morphogenesis, cell part morphogenesis and cell part organisation all below a FDR of 5%. Many of these terms were associated with neurons, from differentiation (GO:0000904, GO:0048667 and GO:0030182), to development and morphogenesis of neuron projections (GO:0031175, GO:0048666 and GO:0048812), to axonogenesis and axon guidance (GO:0007409 and GO:0007411). However the sensitizer candidates may also be involved in heart development (GO:0007507) and leukocyte differentiation (GO:0002521) albeit with high FDR. Management of the cytoskeleton also indicated high gene counts, with terms such as cytoskeleton organisation (GO:0007010), actin cytoskeleton organisation (GO:0030036) and actin filament-based process (GO:0030029) perhaps in association

with its other development and differentiation functions like axonogenesis. This association with the cytoskeleton was significant in the inducer candidates, and suggest the similarity between the inducer and sensitizer candidates. This connection with the inducers and the large number of enriched development-associated terms suggests that the sensitizers might possibly serve as sensors which activate programmed cell death in the event of erroneous development outcome. Other processes with insignificant FDR includes regulation of ARF protein signal transduction (GO:0032012 and GO:0032312) , regulation of hydrolase activity (GO:0051336) and positive regulation of translation (GO:0045727 and GO:0045948). It remains interesting to note that the sensitizers were associated with translational activity while the inhibitors appear more involved in transcriptional processes.

As illustrated in Figure 34, the validated inducers in the "Inducer Main" list localise exclusively to the membrane being intrinsic to membrane (GO:0016021) and integral to the membrane (GO:0016021), although the inclusion of the pre-validated inducers expanded the localisation terms to include the cell soma (GO:0043025), neuron projection (GO:0043005), microtubule (GO:0005874), integral to mitochondrial outer membrane (GO:0031307) and extrinsic to membrane (GO:0031307). Both intrinsic and integral to membrane localisation remains the two most significant localisation terms, and the gene counts were increased proportionally by approximately two-fold. The FDR for the two membrane terms was also improved from >45% to 15.44% and 23.80% for the integral and intrinsic membrane localisation respectively. The extrinsic to membrane term enriched with the inclusion of the pre-validated inducer candidates reinforced this unique enrichment of membrane associated terms, along with the integral to mitochondrial outer membrane term. It should be noted that the FDRs for the new terms enriched with the "Inducer All" list were above 25% and hence should be interpreted with caution.

The inhibitors enriched for 6 / 13 localisation terms with FDR <25%, with the candidates localised to contractile fibre part (GO:0044449 and GO:0043292)), actin cytoskeleton (GO:0015629), non-membrane-bounded organelle (GO:0043228 and GO:0043232) and sarcomere (GO:0030017). The main difference compared with the inducer candidates is the localisation of the inhibitors to non-membrane-bounded locations, making the inducers and inhibitors appear mutually exclusive within the cellular environment.

The sensitizer localisation appeared to overlap with both the inducers and inhibitors. The only term with FDR under the 25% threshold was the cytoskeleton (GO:0005856) which was also enriched in the inhibitors. All three candidate types seem to be associated with localisation to cell projections

although some sensitizers were also associated with membrane terms such as apical plasma membrane (GO:0045177) and basolateral plasma membrane (GO:GO:0016323).

Functionally, the inducers were associated with a wide range of protein and ion binding functions as well as transporter and enzymatic activities. This includes L-glutamate, acidic amino acid and amine transmembrane transporter activity (GO:0005313, GO:0015172 and GO:0005275), symporter activity (GO:0015293), sodium ion binding (GO:0031402), alkali metal ion binding (GO:0031420) and protein binding such as to the cytoskeleton proteins (GO:0008092) and PDZ domain (GO:0030165). The enzymatic activities include cytidine deaminase (GO:0004126), Ras GTPase activator (GO:0005099) and a range of ligase activities (GO:0016879, GO:0019787 and GO:0016881). However, only the PDZ domain binding and ligase activities among the "Inducer Main" list were significant for the FDR threshold. This implies that the pre-validated inducers could be more functionally diverse resulting in lower enrichment significance.

The inhibitors were significant for actin binding (GO:0003779) and calmodulin binding (GO:0005516) functions. The less significant terms include histone acetyltransferase activity (GO:0004402), transcription factor binding (GO:0008134), transcription regulator activity (GO:0030528) which were in line with the inhibitor biological processes profile, and ligase activities similar to the inducers. Calmodulin is an important calcium binding messenger which binds the ions and alter its interaction with its target proteins [345, 346] and is important in calcium regulated apoptosis signalling [347].

The sensitizers were enriched for receptor regulator activity with a FDR of 13.86%, with a range of protein binding functions with higher FDR including bridging (GO:0030674), cytoskeleton binding (GO:0008092), heat shock protein binding (GO:0031072) and SHD3/SH2 adaptor activity (GO:0005070). This variety of protein binding functions suggests that the sensitizers could serve as the mediator between various proteins to coordinate their function or assist in inter-network cross-talking. The sensitizers were also associated with ARF GTPase activator (GO:0008060) and epidermal growth factor receptor (GO:0005006) activities.

GO Terms for enriched biological processes

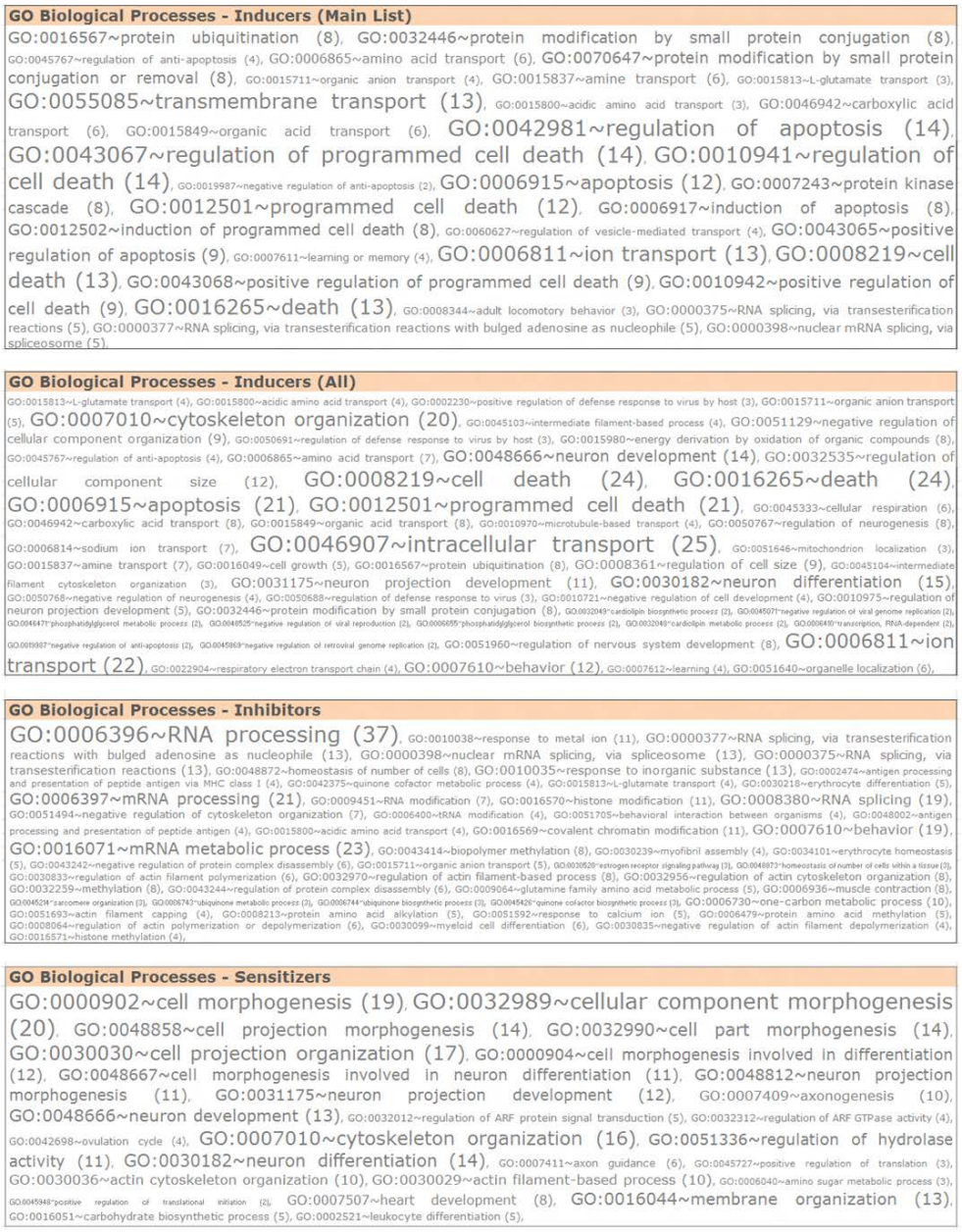


Figure 33: The Gene Ontology biological processes terms displayed as Tag Clouds. The apoptosis inducer candidates in the Inducer Main List were enriched for apoptosis and cell death terms and the counts of death-related terms were increased when the pre-validated candidates were included in the analysis. The inhibitor candidates were enriched for transcriptional-related processes such RNA processing, although enrichment of the sensitizer biological processes associated with apoptosis signalling was less direct consisting of cell morphogenesis, neuron differentiation, translational initiation and membrane organisation. The font size of the terms was presented proportionally to the associated candidate frequency indicated by the brackets.

Figure 34: The Gene Ontology cellular components terms. Inducer candidates from both Main and All lists were associated with membrane type components. Interestingly, the majority of inhibitors were associated with non-membrane bounded organelles. The sensitizers were associated with plasma membrane similarly to the inducers, but were also significant in cytoskeleton and cell projection which are also associated with the inhibitors.

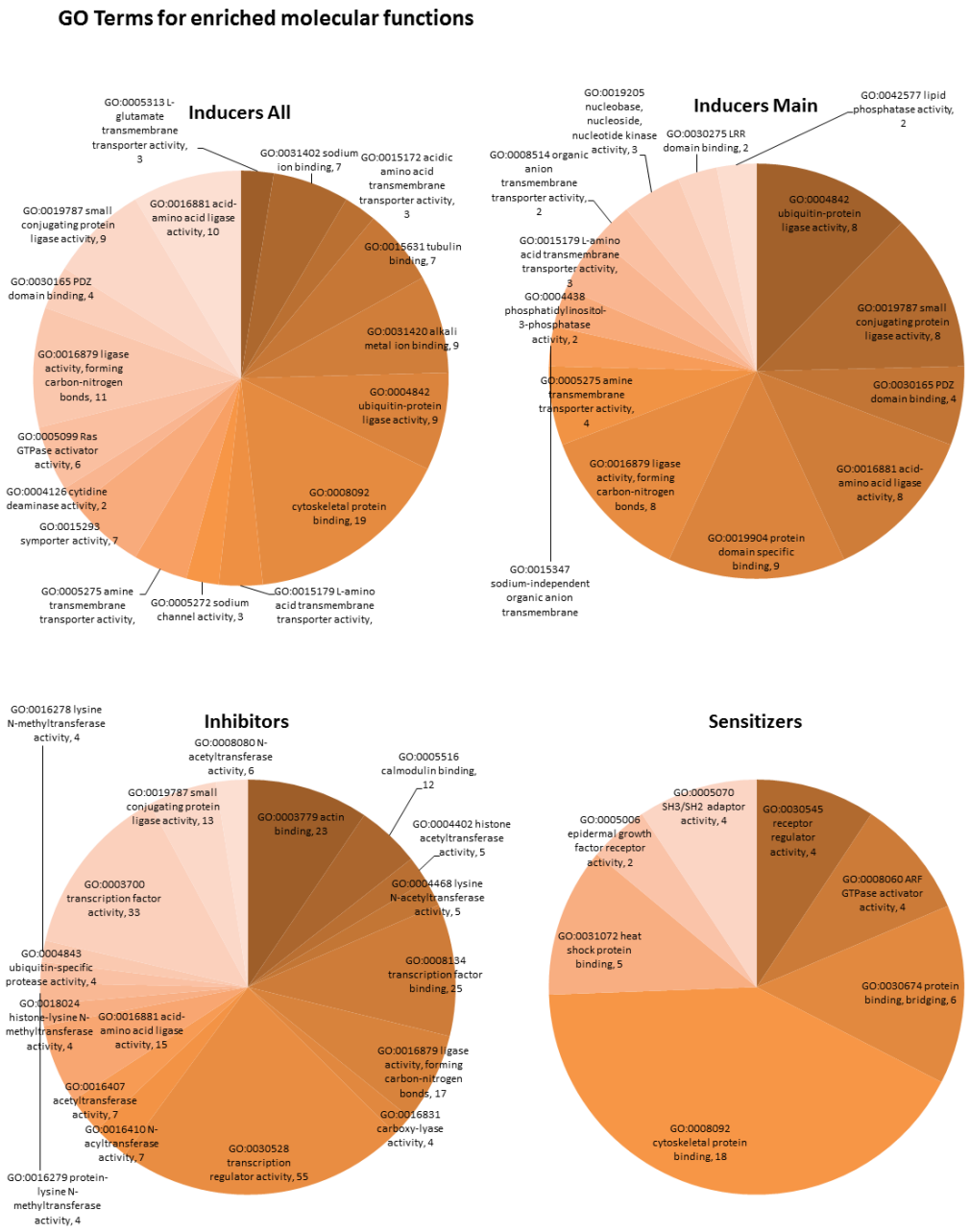


Figure 35: The Gene Ontology molecular functions terms displayed according to the various gene classes from the screen. The inducers and sensitizers appear to be enriched in ion and protein binding functions, while the inhibitors were associated with transcription factor binding and transferase activities. Furthermore, the sensitizers were also enriched in heat shock binding functions and SH2/SH3 adaptor proteins.

Clustering of interaction proteins with the candidates was performed using DAVID on information from the Biomolecular Interaction Network Database (BIND) [348], Molecular INTeraction database (MINT) [349, 350] and Reactome [351].

The inducers were enriched for interactions with members of the ras homolog gene family, various Rho GTPase activating proteins, v-rel reticuloendotheliosis viral oncogene homolog (avian) and deleted in liver cancer 1 (DLC1). These proteins are frequently associated with cancer profiles as oncogenes [352-355] and their ability to suppress apoptosis might be partly attributable to altered interactions with many of the inducer candidates. It should be noted that the FDR did not meet the threshold for any interaction terms of the inducers.

The inhibitors enriched for 103 / 128 protein interaction terms which are within the FDR threshold, almost all of which are associated with transcriptional activities including various small nuclear ribonucleoproteins, splicing factors, RNA binding motif proteins, cleavage and polyadenylation specific factors, small nuclear RNA auxiliary factors, X-box binding nuclear transcription factor, polymerase (RNA) II (DNA directed) polypeptides and THO complex. This association with transcriptional activity continues beyond the FDR threshold, although some interacting proteins of other molecular functions were also enriched such as syndecan 4, receptor-interacting serine-threonine kinase 3 and cell division cycle and apoptosis regulator 1. Syndecan 4 functions as a receptor in intracellular signalling while receptor-interacting serine-threonine kinase 3 and cell division cycle and apoptosis regulator 1 are associated with apoptosis signalling. The enrichment of transcriptional activities associated terms in both the inhibitor candidates and their interacting proteins clearly reinforces the extent of transcriptional mechanisms involvement in apoptosis inhibitory functions.

The three interacting proteins for the sensitizers within the FDR threshold are nucleolin, heterogeneous nuclear ribonucleoprotein A1-like 3 and nucleophosmin 1 which are associated with translational functions. Nucleolin and nucleophosmin 1 are involved in the synthesis and maturation of ribosomes [356, 357]. Nucleophosmin 1 is also known to be crucial in cancer such as leukemia and attributed with tumour suppressor functions [358, 359]. Heterogeneous nuclear ribonucleoprotein A1-like 3 may be involved in mRNA processing functions [360]. The sensitizers were also associated with a variety of eukaryotic translation initiation and elongation factors, ubiquitin B and C, and docking proteins.

Protein domain enrichment was performed using DAVID with information from the InterPro [361], SMART [362, 363] and Protein Information Resources (PIR) [364] databases (Figure 36).

The inducer candidates displayed enrichment for Major Facilitator Superfamily MFS-1 (IPR011701) and RING-type Zinc finger (IPR001841) domains with an FDR under 25%. The MFS transporters are involved in the transport of small solutes in response to chemiosmotic ion gradients [365] while the RING-finger is a specialised class of zinc finger domains involved in mediating protein-protein interactions [366], many of which bind ubiquitination enzymes and their substrate thereby targeting their substrate for degradation [367]. Other notable protein domains include the Hly-III related (IPR004254) domains of which the family comprised of integral membrane proteins including hemolysin-III homologues, the Leucine-rich repeats (LRR) (SM00013 and SM00082) which fold into an arc shape and provide a framework for protein-protein interactions, AAA ATPases (SM00382) which are involved in a range of roles from cell-cycle regulation to protein proteolysis and intracellular transport, EFh-hands (SM00054) associated with calcium binding, and a range of immunoglobulin domains (SM00408, IPR013106, IPR03598 and IPR013783).

Immunoglobulin folds (IPR003006, M00407 and IPR003597) were highly enriched in the inhibitors. Ig-like domains are associated with functions such as cell-cell recognition, cell surface receptors and the immune system [368]. C6HC-type Zinc finger (SM00647), G8 domain (IPR019316) found in disease proteins in polycystic kidney disease and non-syndromic hearing loss [369], WD40 repeats (SM00320, IPR001680, IPR015943 and IPR019775) implicated in signal transduction and transcription regulation of cell cycle and apoptosis [370], PWWP domains (SM00293 and IPR000313) with a role in cell growth and differentiation and protein-protein interactions [371], PLAC (protease and lacunin) domain (IPR010909) associated with peptidase activity, TUDOR domains (SM00333, IPR002999 and IPR018351) associated with zinc binding and found in many proteins colocalised with ribonucleoprotein or single-strand DNA –associated complexes [372], peptidase C19 ubiquitin carboxyl-terminal hydrolase 2 (IPR001394) and pyridoxal phosphate-dependent decarboxylase (IPR002129) are other protein domains associated with the inhibitors.

The enrichment of the immunoglobulin folds between the pro-and anti-apoptotic candidates is an interesting find, and may suggest a link between apoptosis and the intracellular antibody-mediated degradation. TRIM21 for example is an intracellular antibody effector of the pathway and binds both immunoglobulin G and M resulting in ubiquitination [373]. The candidates with these immunoglobulin folds may represent novel effectors of the intracellular antibody-mediated proteolysis, or regulate apoptosis by tagging their target for degradation.

The sensitizers were associated with the TUDOR domains (SM00333 and IPR00299) and C-terminal chaperone DnaJ domains (IPR002939), with both domains represented under the FDR thresholds. Other domains enriched include ArfGAP (SM00105 and IPR001164), Leucine-rich repeat (SM00369 and IPR003591), actin/actin-like (SM00268 and IPR004000), EXTL2 alpha-1,4-N-acetylhexosaminyltransferase (IPR015338), spindle associated domains (IPR012943) and A-kinase anchor (IPR008382). ArfGAP family proteins are strongly implicated in tumour invasion and malignancy with a critical role in suppression of superoxide production [374].

The sensitizers appear to have an overlap in the domain enrichment with the inducers and inhibitors, for example the leucine-rich repeats with the inducers and the TUDOR domains with the inhibitors. However, the sensitizers also included unique enrichment for domains associated with DnaJ which is a chaperone associated with binding of unfolded proteins to prevent aggregation and possesses protein folding activity [375]. The inducers and inhibitors also shared similarities, most notably the enrichment for immunoglobulin domains. The three candidate types appear to enrich for domains associated with protein-protein interactions, which might be speculated to aid the candidates in achieving their pro- or anti-apoptotic behaviour.

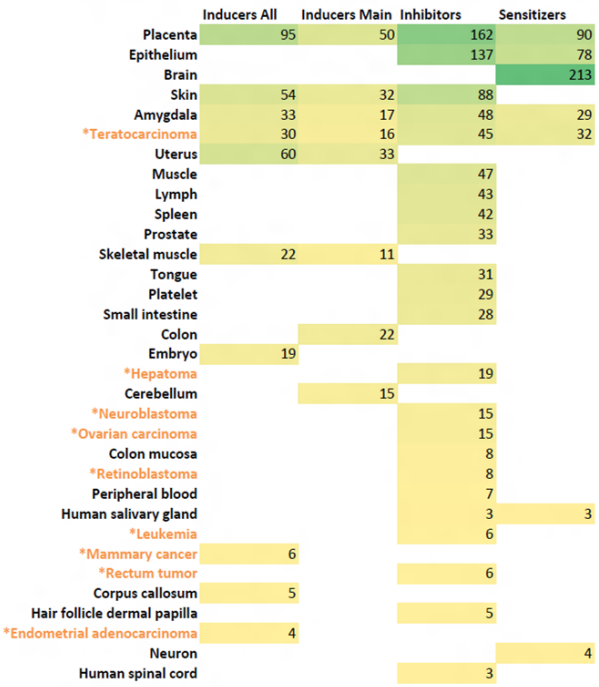
DAVID was also used to identify the enrichment for the tissue specificity of each candidate class using the UniProt tissue specificity annotation (Figure 37A). All three candidate types were found to be enriched in the placenta, amygdala and teratocarcinoma, beyond which they appear to be mutually exclusive. The inducers were found to be expressed in the uterus, skeletal muscle, colon and embryo of normal tissue types. The tissue expression of the inhibitors appeared with greater diversity including muscle, lymph, spleen, prostate, tongue, platelet, small intestine, colon mucosa, peripheral blood, salivary gland and the spinal cord. The inhibitors were also strongly enriched in the epithelium with 137 gene counts, a phenotype also observed for the sensitizers with 78 genes. The inhibitors and inducers also overlapped in expression in the skin, with 88 and 54 genes respectively. The sensitizers were unique in enriching for 213 genes expressed in the brain and 4 counts in the neuron, which might be a result of their role in neuron morphogenesis and development of neuron projects. The sensitizers were also uniquely expressed in other tissue types.

Interestingly, the three candidate types were found to be expressed in teratocarcinoma, with similar frequency between 30 and 45. The inducers were expressed in mammary cancer and endometrial adenocarcinoma while the sensitizers were not found to be expressed in other cancer tissues apart from teratocarcinoma. The inhibitors were however extremely well represented in the cancer tissue types with 114 genes associated with cancer tissues compared with 40 and 32 for the inducers and sensitizers respectively (Figure 37B). The cancer tissue-associated terms enriched amongst the

inhibitors included hepatoma, neuroblastoma, ovarian carcinoma, retinoblastoma, leukaemia and rectum tumour. The high association with cancer tissues was supportive of their anti-apoptotic effect identified from the screen. Moreover, it could be hypothesized that the pro-survival signalling exploited by cancer cells could similarly be employed by the normal cells in response to LUDOX® silica nanoparticles induced cell death.

Figure 36: Enrichment of Protein Domains. The inducers were associated with MFS-1 domains associated with solute transporters and the RING-type zinc finger domains involved in DNA and protein binding. The inhibitors were enriched for immunoglobulin domains and interestingly, the C6HC-type zinc finger which is also involved in protein-protein interactions, while the TUDOR and DnaJ c-terminal domains were enriched amongst the sensitizers.

A) Tissue Expression Profile



	Inducers All	Inducers Main	Inhibitors	Sensitizers
Placenta	95	50	162	90
Epithelium			137	78
Brain				213
Skin	54	32	88	
Amygdala	33	17	48	29
*Teratocarcinoma	30	16	45	32
Uterus	60	33		
Muscle			47	
Lymph			43	
Spleen			42	
Prostate			33	
Skeletal muscle	22	11		
Tongue			31	
Platelet			29	
Small intestine			28	
Colon		22		
Embryo	19			
*Hepatoma			19	
Cerebellum		15		
*Neuroblastoma			15	
*Ovarian carcinoma			15	
Colon mucosa			8	
*Retinoblastoma			8	
Peripheral blood			7	
Human salivary gland			3	3
*Leukemia			6	
*Mammary cancer	6			
*Rectum tumor			6	
Corpus callosum	5			
Hair follicle dermal papilla			5	
*Endometrial adenocarcinoma	4			
Neuron				4
Human spinal cord			3	

B) Frequency of cancer tissue-associated terms

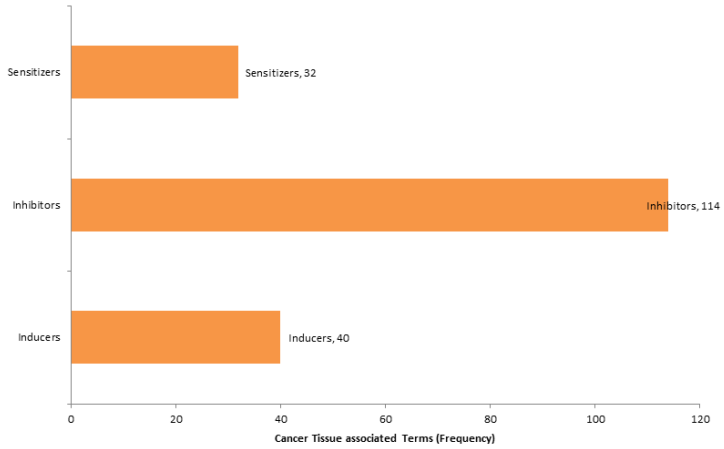


Figure 37: Tissue expression profile associated with each candidate type. A) All four candidate gene lists were highly enriched for expression in the placenta, amygdala and teratocarcinomas, although terms unique to each candidate type were also enriched. Cancer associated tissue types were labelled in orange. B) Cancer tissue-related term frequency. The inhibitor candidates were highly enriched for cancer tissue-associated terms, with up approximately three fold increase compared to the inducer and sensitizer candidates.

Gene Set Enrichment Analysis

The primary data were analysed using Gene Set Enrichment Analysis (GSEA) V2.0 [376, 377]. A ranked list for the non-treated dataset was generated using the CPRG ratio while the nanoparticles treated dataset was ranked using the normalised TNTD score calculated based on Equation 12. The non-treated dataset corresponds to the inducer candidates while the treated dataset was used for enrichment associated with inhibitors and sensitizers.

Both datasets were analysed against three curated Molecular Signatures Database v3.0 collections: 1) c2 curated gene sets comprising of pathway databases, 2) c3 motif gene sets consisting of conserved cis-regulatory motifs and 3) computational gene sets defined with expression neighbourhoods focused on 380 known cancer genes. The default settings of 1,000 permutations, weighted enrichment statistic with minimum and maximum set size of 15 and 500 genes was used. The GSEA results were complimentary to the formerly identified list since the approach was based on the statistical significant of a priori defined set of genes, hence would be able to identify enrichment information such as signalling pathways without every component of the pathways passing the thresholds.

With the c2 collection, the NFAT pathway and Hypertrophy of the heart gene set (BIOCARTA_NFAT_PATHWAY) were enriched with the highest normalised enrichment score (NES) of 2.030, normalised p-value of 0.001198 and 15 / 19 gene counts in the core enrichment (Figure 38A). Nuclear factor of activated T-cells (NFAT) is a family of transcription factors comprising of four members localised to the cytoplasm, two (NFAT3 and NFATc4) of which are expressed in adult heart and regulates cardiac hypertrophy signalling [378, 379]. The association between a gene set involved in heart failures with genes displaying pro-cell death increase of CPRG ratios is supportive of the quality of the primary data. Furthermore, two gene sets associated with genes up-regulated in ATG16L1 deficiency (CADWELL_ATG16L1_TARGETS_UP, NES 1.929, NOM p-value 0.004662, 24 counts) and up-regulated in response to IL6 (DASU_IL6_SIGNALING_UP, NES 1.766, NOM p-value 0.008403, 19 counts) also supported the association of increased CPRG ratios with increased cell death. ATG16L1 is a gene involved in autophagy [380], which when excessive, can lead to cell death. Interleukin-6 (IL6) is a cytokine involved in inflammatory response [381]. While IL6 signalling is commonly associated with apoptosis suppression [382], it has also been shown to activates apoptosis under conditions such as STAT3 inhibition [383]. Three transmembrane transport associated gene sets involved in SLC-mediated transmembrane transport (REACTOME_SLC_MEDIATED_TRANSMEMBRANE_TRANSPORT), inorganic cation/anion SLC transporters (REACTOME_INORGANIC_CATION_ANION_SLC_TRANSPORTERS) and transmembrane

transport of small molecules (REACTOME_TRANSMEMBRANE_TRANSPORT_OF_SMALL_MOLECULES) was also identified and this was in line with membrane localisation enriched in the inducers by DAVID. Amongst the five cancer associated gene sets enriched in the top twenty terms, three sets indicated genes with increased CPRG ratios were down-regulated in colon carcinoma tumours (GRADE_COLON_CANCER_DN), 3D cultures of preinvasive breast cancer cells (RIZKI_TUMOR_INVASIVENESS_3D_DN) and diffuse large B-cell lymphoma (SHIPP_DLBCL_VS_FOLLICULAR_LYMPHOMA_DN) consistent with the predicted behaviour of pro-cell death genes. Interestingly, these pro-cell death genes were also up-regulated in prostate cancer (TOMLINS_PROSTATE_CANCER_UP) and stage 2 colon cancer (BARRIER_COLON_CANCER_RECURRENCE_UP), which could support the hypothesis that these genes have significant roles in cancer metabolism and signalling and could serve as drug targets since they may predispose the cancer cells to apoptosis.

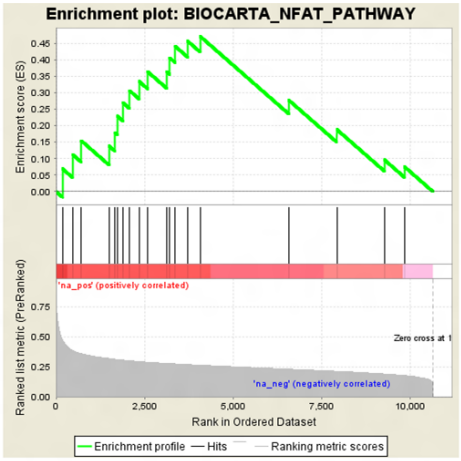
The inhibitors selection tail (negatively enriched by GSEA) was highly enriched in cancer associated signalling which formed nine of the top twenty enriched terms. Six of these cancer-associated terms were for up-regulation of genes in breast cancer (NADERI_BREAST_CANCER_PROGNOSIS_UP), embryonic carcinoma tumours (KORKOLA_EMBRYONAL_CARCINOMA_UP), seminoma tumours (KORKOLA_SEMINOMA_UP), hepatocellular carcinoma (LEE_LIVER_CANCER_MYC_TGFA_UP and LEE_LIVER_CANCER_E2F1_UP) and esophageal adenocarcinoma (WANG_ESOPHAGUS_CANCER_VS_NORMAL_UP), which supported the apoptosis suppression capabilities of the inhibitor candidates. The top signalling profile enriched by GSEA was the gene set which distinguish cells expressing activated beta-catenin (CTNNB1) oncogene from the normal cells (BILD_CTNNB1_ONCOGENIC_SIGNATURE) [384] which had the highest normalised enrichment score of -1.87 and 14 genes among the core enrichment (Figure 38B). The knockdown of HIF1A and HIF2A (ELVIDGE_HIF1A_TARGETS_UP and ELVIDGE_HIF1A_AND_HIF2A_TARGETS_UP) also resulted in the up-regulation of genes on the inhibitor tail, perhaps as a compensatory effect involving the action of 2-OC-dependent dioxygenases on non-HIF target genes. This gene set was also similar to a gene expression profile induced by hypoxia. Finally, the knockdown of cofactor of BRCA1 (COBRA1) also led to the down-regulation of a gene set associated with the inhibitor tail. COBRA1 is important in regulating the expression of genes at various chromosomal locations in breast cancer, thereby implicating the genes in another cancer signalling mechanism.

The sensitizer tail had enriched for much less profiles associated with cancer unlike the inhibitors, with only three occurrences in among the top two enriched gene sets. Two of the three cancer-associated gene sets indicated down-regulation of the candidates in the sensitizer tail

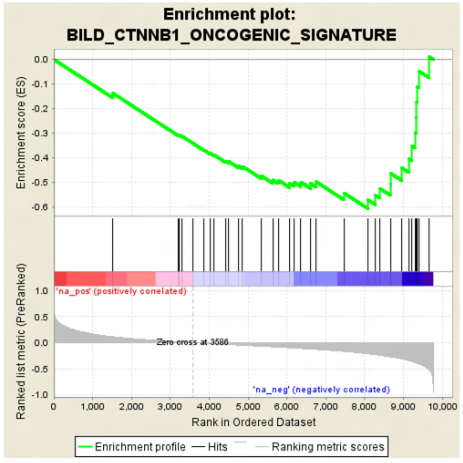
(BOYAULT_LIVER_CANCER_SUBCLASS_G123_DN and LANDIS_BREAST_CANCER_PROGRESSION_DN) in liver and breast cancer respectively. The third cancer-associated gene set was associated with the gene signature of S2 subtype of hepatocellular carcinoma (HCC), with implications in proliferation, MYC and AKT1 activation.

The sensitizers were instead enriched for signalling pathways associated with cellular functions, with NCAM1 signalling (REACTOME_NCAM1_INTERACTIONS and REACTOME_NCAM_SIGNALING_FOR_NEURITE_OUT_GROWTH) as the top gene set with NES of 1.95 (Figure 38C). Candidates located at the sensitizer tail were also found to be up-regulated in inflammatory response to lipopolysaccharide (LPS) (SEKI_INFLAMMATORY_RESPONSE_LPS_UP), hepatocyte growth factor signalling not associated with AKT1 signalling (XU_HGF_SIGNALING_NOT_VIA_AKT1_48HR_UP), up-regulated by CD40 receptor (BASSO_CD40_SIGNALING_UP) and also in response to interleukin-6 (IL-6) stimulation (DASU_IL6_SIGNALING_UP). Translational activity was also prominent in the sensitizer tail, with four gene sets (REACTOME_TRANSLATION_INITIATION_COMPLEX_FORMATION, KEGG_RIBOSOME, REACTOME_VIRAL_MRNA_TRANSLATION and REACTOME_FORMATION_OF_THE_TERNARY_COMPLEX_AND_SUBSEQUENTLY_THE_43S_COMPLEX) . Signalling by Rho GTPases (SIG_REGULATION_OF_THE_ACTIN_CYTOSKELETON_BY_RHO_GTPASES) and PDGF (REACTOME_SIGNALING_BY_PDGF) was also identified, along with regulation of peroxisome (KEGG_PEROXISOME), regulation of beta-cell development (REACTOME_REGULATION_OF_BETA_CELL_DEVELOPMENT) and DNA replication in oocyte (KEGG_OOCYTE_MEIOSIS). Another interesting pathway profile was for the genes involved in the RNA transcription and replication of Influenza virus (REACTOME_INFLUENZA_VIRAL_RNA_TRANSCRIPTION_AND_REPLICATION). Unlike the inhibitors where a clear association with a range of cancer signalling profiles was found, there does not appear to be any clear indication of association with apoptosis among the sensitizers. However, signalling pathways associated with translation and development of neurons was similar to the functional annotation results by DAVID (see above).

A) The NFAT Pathway was the top gene set for the inducers



B) The inhibitors exhibited oncogenic signature associated with CTNNB1 activation



C) NCAM1 interactions associated with neurone development was enriched in the sensitizers

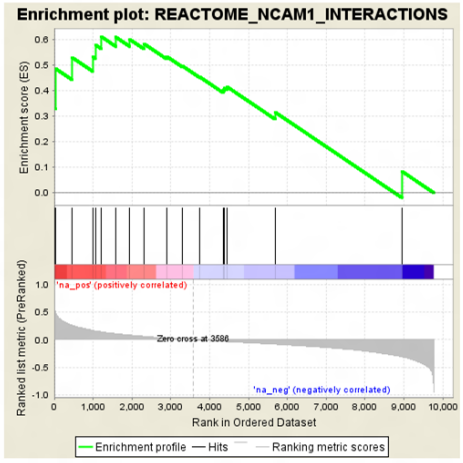


Figure 38: GSEA enrichment plot for the top gene set associated with each candidate type. The enrichment score is the maximum deviation from zero with the genes prior to enrichment score is known as the leading edge. A) The NFAT pathway has the highest normalised enrichment score (NES) of 2.03 among the inducer tail, with 15 / 19 genes in the core enrichment. B) The inhibitors exhibited oncogenic signature associated with cells expressing activated CTNNB1, a known oncogene. The CTNNB1 signature was enriched with NES of -1.86 and the core enrichment including 14 / 35 genes. C) The NCAM1 interactions gene set was associated with the sensitizers with NES of 1.95 and 6 / 17 genes in the core enrichment.

The c3 collection offered insights into transcriptional regulatory mechanisms of the candidate types. The genes in the inducers tail of the non-treated dataset were found to be target genes of Sp1 transcription factor (V$SP1_01 and V$SP1_Q2_01) as their upstream promoter regions contained the GGGGCGGGGT motif recognised by Sp1. Other promoter regions were also identified including those regulated by hepatic nuclear factor (HNF1) (V$HNF1_01), HMX1 (V$HMX1_01), CEBPA (V$CEBP_C), SREBF1 (V$SREBP1_02), HOXA4 (V$HOXA4_Q2) and FOXN1 (V$WHN_B). Some of the candidates at the inducer tail were also found to be targets of microRNA including MIR-122A (ACACTCC,MIR-122A), MIR-409-3P (AACATTC,MIR-409-3P), MIR-299-SP (GTAAACC,MIR-299-5P), MIR-346 (GGCAGAC,MIR-346) and MIR-518A-2 (TTTGCAG,MIR-518A-2). For example, the genes SMARCD1, CD320, NEGR1, TBC1D108, BAI2, MARK1, SLC39A8, CDC42BPB, GNPDA2, SLCO5A1, CALM3, HAND2, ANKRD13C could potentially be targeted by the MIR-122A microRNA; this represented a core enrichment of 13 / 41 genes. Seven of the terms indicated similarity in the promoter regions but the associated transcription factor was currently unknown.

Using the silica nanoparticles treated dataset, candidates at the inhibitor tail were enriched for ten promoter regions associated with ESRRA (V$ERR1_Q2), JUN (V$AP1_Q6, V$AP1_Q6_01, V$AP1_01), POU3F1 (V$TST1_01), TCF4 (V$TCF4_Q5), SRF (V$SRF_C), VDR (V$VDR_Q6), MIF (V$MIF1_01) and SPI1 (RGAGGAARY_V$PU1_Q6). Both JUN and SPI1 are known oncogene genes [385, 386] and the controls of the anti-apoptosis candidates suggests that these oncogenic transcription factors could contribute to the response towards the silica nanoparticle treatment. Five microRNA sequences were identified as potential regulators of the inhibitor candidates: MIR-198 (TCTGGAC,MIR-198), MIR-205 (ATGAAGG,MIR-205), MIR-190 (ACATATC,MIR-190), MIR-216 (TGAGATT,MIR-216) and MIR-90 (CCAGGTT,MIR-490). Three promoter regions were not known to be associated with known transcription factors.

On the sensitizer tail, 12 / 20 of the top enrichment terms were associated with known transcription factors including: TCF1 (RGTTAMWNATT_V$HNF1_01), EN1 (V$EN1_01), NF1 (V$NF1_Q6), FOXA1 (V$HNF3ALPHA_Q6), EGR3 (V$EGR3_01), STAT1 (V$STAT1_01), PAX5 (V$PAX5_01), CART1 (V$CART1_01), CEBPB (V$CEBPB_02), TGIF (V$TGIF_01), GATA1 (V$GATA_Q6), NR2F2 (V$DR1_Q3), HNF4A (V$HNF4_DR1_Q3) and POU2F1 (V$OCT1_07). 2 / 20 of the enrichment terms were associated with microRNA comprising of MIR-31 (ATCTTGC,MIR-31) and MIR-320 (CAGCTTT,MIR-320), while four promoter regions with unknown transcription factors were also enriched.

It is interesting to note that between the three classes of candidates, the known transcription regulatory mechanisms were exclusive to each class and do not overlap with other transcription factors.

Finally, using the c4 collection enabled the clustering of candidates to known cancer gene neighbourhoods on the genome to be analysed.

The inducers were found to be localised to the neighbourhood of GPX4 (MORF_GPX4), PSMC2 (MORF_PSMC2), ETV3 (MORF_ETV3), S100A4 (GNF2_S100A4), BNIP1 (GCM_BNIP1), PPP2R4 (MORF_PPP2R4), PRKAR1A (MORF_PRKAR1A), ATOX1 (MORF_ATOX1), LMO1 (MORF_LMO1), THRA (MORF_THRA) and STK17A (MORF_STK17A).

The inhibitors were found near cancer-related genes such as ITGA2 (MORF_ITGA2), BAG5 (GCM_BAG5), CSNK1D (GCM_CSNK1D), PTPN6 (GNF2_PTPN6), G22P1 (GNF2_G22P1), SIRT2 (GCM_SIRT2), REV3L (MORF_REV3L), CD53 (GNF2_CD53), SELL (GNF2_SELL), FSHR (MORF_FSHR), ELAC2 (GNF2_ELAC2), CCNA2 (GNF2_CCNA2), HAT1 (GNF2_HAT1), TYK2 (GNF2_TYK2), PRKCA (MORF_PRKCA) and MLLT10 (MORF_MLLT10).

The sensitizers were found in the region of TPT1 (GCM_TPT1), EIF4A2 (MORF_EIF4A2), BECN1 (MORF_BECN1), CCNF (MORF_CCNF), DNM1 (GNF2_DNM1), TST (GNF2_TST and MORF_TPT1), CCNA1 (GNF2_CCNA1), MAPT (GNF2_MAPT), NPM1 (GCM_NPM1), PTX3 (GNF2_PTX3), MLF1 (GNF2_MLF1),

As with the enrichment from the c3 collection, the three candidate types appear to be mutually exclusive although the implications of being in the proximity of cancer-associated neighbourhoods may likely be dependent on the type of cancers.

Ingenuity® and Alternative Signalling Pathway Analysis

Ingenuity® Pathway Analysis is a curated proprietary knowledgebase widely adopted as an integrative analysis platform to examine complex biological and chemical relationships. IPA incorporates data from a range of experimental platforms, providing insight into molecular and chemical interactions, signalling pathways, cellular and disease phenotypes. The IPA knowledgebase comprises curated information on genes, proteins, chemicals, drugs, biomarkers and other molecular relationships to quickly build biological models or help drive novel hypothesis.

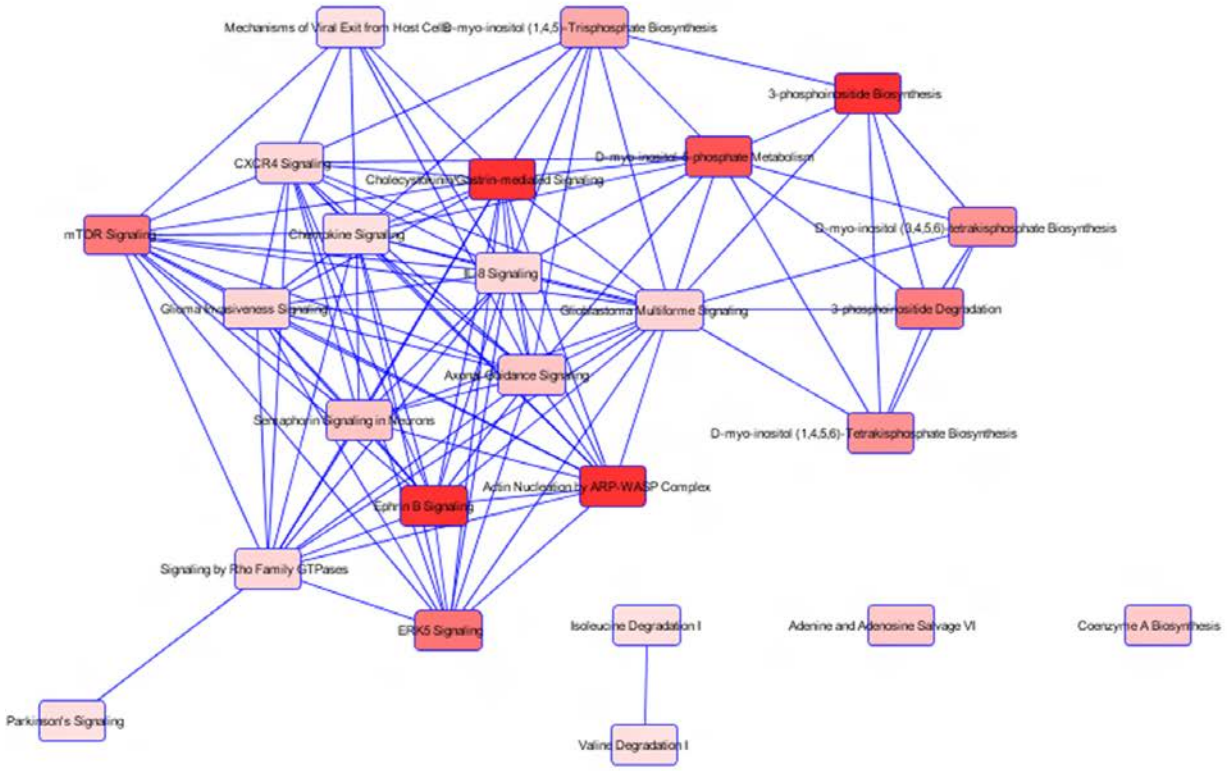
Here, the functional information, where available, was used to reinforce, or supplement, the annotation results generated with DAVID. In particular, the signalling pathways generated by IPA were combined with those enriched by DAVID, GSEA and GeneMANIA plugin [387] on cytoscape 2.8 [388] to build potential signalling networks and identify important signalling mechanisms.

Inducers

The inducer candidates were well enriched for phosphorylated inositol signalling. 3-phosphoinositide biosynthesis and D-myo-inositol-5-phosphate metabolism pathways were among the top pathways identified by IPA (Figure 39), while DAVID (Figure 40) identified phosphatidylinositol signalling system (hsa04070) and inositol phosphate metabolism (hsa00562). Phosphoinositide is defined as any phosphorylated inositol-containing compounds involved in cell activation and calcium mobilisation in response to hormones. The four phosphoinositide associated signalling pathways yielded eight unique genes (PLCB2, PI4KB, PRKCB, PTEN, DGKA, CDIPT, PPM1F and PTPN1) from the screen out of 483 protein coding inducers.

PLCB2 is a calcium-binding protein involved in signal transduction through the production of the second messenger molecules diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP3). PLCB2 is also involved in the activation of phospholipase C activity. The IP3 molecule is soluble and freely diffusible across the cytoplasm to the endoplasmic reticulum (ER) where it activates the release of calcium, thereby increasing intracellular calcium concentration which is a trigger of apoptosis [389]. DAG remains in the plasma membrane where it facilitates the translocalisation and activation of protein kinase C (PKC). Phosphatidylinositol 4-kinase beta (PI4KB) is another protein involved in signal transduction, and phosphatidylinositol biosynthetic and signalling processes. PI4KB catalyses the phosphorylation of phosphatidylinositol (PI) which is the first committed step in the synthesis of PI3 [390, 391]. Protein kinase C beta type (PRKCB) is activated by calcium or DAG and is associated with a range of cellular functions including oxidative stress-induced apoptosis, regulation of androgen receptor-dependent transcription, regulation of the B-cell receptor (BCR) signalosome, insulin signalling and proliferation of endothelial cells [392].

A) Integrated map of the enriched signalling pathways



B) Pathway chart of the top signalling pathways

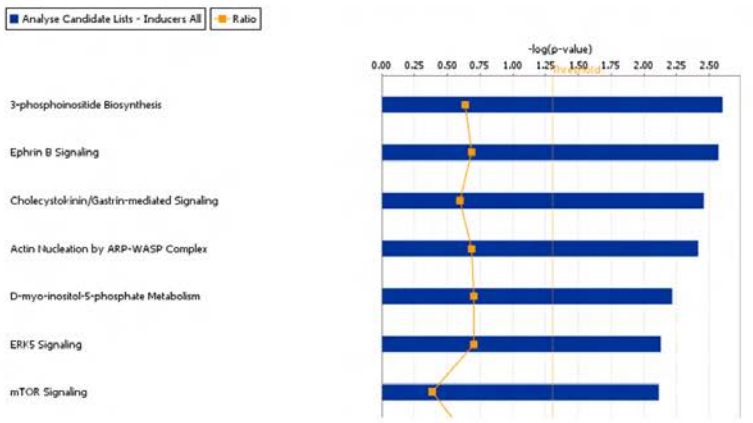


Figure 39: Signalling pathways identified for the Inducers. A) An overview showing all the signalling pathways and their relationships. The intensity of the shade of red increases with the significance of the signalling pathway. B) Pathway chart showing the most significant signalling pathways. The yellow ratio line represents the proportion of the candidate genes identified among each curated gene set.

Figure 40: Enrichment of the signalling pathways identified by DAVID. The signalling pathways enriched by DAVID were ranked according to FDR and then plotted as a pie chart using the gene counts associated with each signalling pathway.

PTEN is a known tumour suppressor involved in apoptosis, lipid metabolism and neurogenesis. Its dual-specific phosphatase function includes protein phosphatase where it dephosphorylate proteins phosphorylated at the tyrosine, serine or threonine residues, while its lipid phosphatase activity dephosphorylates the D3 positions of the inositol ring of various phosphoinositides [393, 394]. This lipid phosphatase activity is important for its tumour suppressor function [395]. It is a crucial modulator of the AKT-mTOR signaling pathway, hence controls the process of neurogenesis. The mono-ubiquitinated PTEN localised to the nucleus possesses enhanced apoptotic potential compared to the non-ubiquitinated form located in the cytosol [396-398].

Diacylglycerol kinase alpha (DGKA) converts DAG into phosphatidate to initiate the resynthesis of phosphotidylinositols, in the process terminating the PKC activity [399]. CDP-diacylglycerol--inositol 3-phosphatidyltransferase (CDIPT) is another enzyme involved in the biosynthesis of phosphatidylinositol and could function to limit excessive cellular phosphatidylinositol content through its phosphatidylinositol:inositol exchange reaction [400]. Protein phosphatase 1F (PPM1F) is a phosphatase that dephosphorylates CaM-kinases and therefore inactivates them; PPM1F is also known to be pro-apoptotic [401]. Tyrosine-protein phosphatase non-receptor type 1 (PTPN1) is a tyrosine-protein phosphatase involved in the regulation of endoplasmic reticulum unfolded protein response [402], which is known to lead to apoptosis [403].

These phosphoinositide associated pathways, in combination with the mTOR signalling identified by IPA, would constitute the PI3K/AKT/mTOR signalling pathway. Within the PI3K/AKT/mTOR pathway, activation of PI3 kinases (PI3K) in turns activates protein kinase B (AKT) which then activates mTOR; this pathway is heavily up-regulated and exploited in cancers as a proliferative mechanism to by-pass apoptosis [404, 405]. Here, it was discovered that many members associated with this signalling pathway also possess a pro-apoptotic functions, which could be exploited as treatment focal points in cancer therapies.

Nine genes were identified for the mTOR pathway: DIRAS3, MRAS, PRKAG2, PRKCB, RHOA, RHOT1, RPS3A, RPS6KA1 and RPS6KA3. GTP-binding protein Di-Ras3 (DIRA3) possesses GTP binding and GTPase activity, and is involved in the regulation of cyclin-dependent protein kinase activity, gene expression by genetic imprinting and signal transduction mediated by small GTPase [406]. DIRA3 was also found to be expressed in normal ovarian and breast epithelial cells but not the cancer cells [406]. Ras-related protein M-Ras (MRAS) is involved as a signal transducer during the control of cell proliferation and is a weak activator of the MAP kinase signalling pathway [407]. 5'-AMP-activated protein kinase subunit gamma-2 (PRKAG2) is an energy sensor protein kinase involved in the regulation of cellular energy metabolism [408].

RhoA regulates the signal transduction between the cell surface receptors and to the assembly of actin stress fibres and focal adhesions [409]. The involvement of RhoA in apoptosis is conflicting, where it is associated with the negative regulation of apoptosis in neurons and the positive regulation of NF-kappaB and MAPK which are generally pro-survival proliferative pathways [410, 411]. However, RhoA along with p38 MAPK were shown to be pro-apoptotic under depleted cellular cholesterol environment [412] while Rho1, the Drosophila melanogaster homologue of RhoA, was demonstrated to induce apoptosis via the JNK signalling pathway [413]. Here, we demonstrated that the ectopic expression of RhoA constitute a pro-apoptotic signal. Mitochondrial Rho GTPase 1 (RHOT1) is localised to the mitochondria outer membrane where it is involved in mitochondrial trafficking processes which transport and organise mitochondria. Overexpression of RHOT1 has been observed to increase apoptosis [414, 415]. 40S ribosomal protein S3a (RPS3A) is involved in translational initiation and associated with the induction of apoptosis [416]. Ribosomal protein S6 kinase alpha-1 (RPS6KA1) and Ribosomal protein S6 kinase alpha-3 (RPS6KA3) are serine/threonine-protein kinases which act downstream of ERK (MAPK1/ERK2 and MAPK3/ERK1) signalling and mediate stress-induced activation of transcription factors such as CREB1, NR4A1/NUR77 and ETV1/ER81 [417]. They promote cell proliferation and survival via mTOR signalling and suppression of pro-apoptotic activity of DAPK and BAD [418, 419].

MAPK signalling pathway (hsa04010) was identified by DAVID while ERK5 signalling was identified by IPA, with both pathways comprised of members from the mitogen-activated protein kinase (MAPK) family. Both pathways are activated by growth factors and play a role in the regulation of cell proliferation [420]. RPS6KA1, STMN1, NF1, GNA12, MRAS, FGFR1, PRKCB, RPS6KA3, PLA2G6 and ELK4 are associated with the classical MAPK signalling while five genes (ELK4, GNA12, MRAS, RPS6KA1 and RPS6KA3) are associated with ERK5 signalling. All five genes enriched for ERK5 signalling is a subset of the ten genes associated with the extended MAPK signalling pathway, while RPS6KA1, RPS6KA3, PRKCB and MRAS were also identified with the PI3K/AKT/mTOR signalling pathway.

Among the genes involved in MAPK signalling, six were unique to this pathway. Stathmin (STMN1) is involved in the regulation of the microtubule filament system by preventing the assembly and promoting the disassembly of microtubules. STMN1 binds to two alpha/beta-tubulin heterodimers and is known to interact with KIST [421, 422]. Guanine nucleotide-binding protein subunit alpha-12 (GNA12) is part of the G proteins family associated as regulators and signal transducers for various transmembrane signalling systems such as the MAPK and Rho signalling pathways. ETS domain-containing protein Elk-4 (ELK4) is a DNA binding protein with transcription cofactor activity [423,

424]. Fibroblast growth factor receptor 1 (FGFR1) is involved in cell proliferation, differentiation and migration. It acts as the surface receptor for the fibroblast growth factors and possesses tyrosine-protein kinase activity capable of phosphorylating targets such as PLCG1, FRS2, GAB1 and SHB upon ligand binding to activate their signalling cascade. PLCG1 activation triggers DAG and IP3 production which activates the IP3K signalling pathway while SHB and FRS2 mediates the MAPK signalling [425, 426]. Neurofibromin (NF1) is a likely regulator of Ras capable of stimulating the GTPase activity of Ras [427] while 85 kDa calcium-independent phospholipase A2 (PLA2G6) catalyses the release of fatty acids from phospholipids and is also associated with cardiolipin biosynthesis [428, 429]. Both NF1 [430] and PLA2G6 [431] were previously demonstrated to be pro-apoptotic.

Gene sets featuring intracellular transport pathways were also enriched using GSEA, which included SLC mediated transmembrane transport, amino acid and oligopeptide SLC transporters, inorganic cation anion SLC transporters and transmembrane transport of small molecules. In total, 46 unique genes were associated with transport pathways as enriched by GSEA. These solute-carrier (SLC) transporters facilitate the movement of a specific substrate either against or with their concentration gradient through conformation change of the transporter protein [432]. While the direct association of SLC transporters with apoptosis is not clear, SLC transporters are known to play an important role in resistance of cancer cells towards chemotherapy by reducing the cellular accumulation of drugs. Some SLC transporters such as folate, nucleoside, and amino acid transporters, in contrast increase the sensitivity to chemotherapy by mediating the uptake of hydrophilic drugs [433]. It could be speculated that the ectopic expression of these transporters resulted in an imbalance of the intracellular homeostasis of their substrate, allowing for pro-apoptotic signals to manifests. Such ectopic expression may be reminiscent of the up-regulation of these transporters in disease pathology.

The NFAT signalling pathway was also enriched by GSEA (BIOCARTA_NFAT_PATHWAY) with 15 genes among the core enrichment. It is closely associated with the IP3 signalling, whereby the initial calcium release triggers a greater calcium influx through calcium release activated channels (CRAC), which then activates the NFAT transcriptional activity [434]. This pathway is heavily implicated in cardiac failure, which is orchestrated by cytokines and growth factors acting through various signalling cascade such as the MMAPK, PKC, low molecular weight GTPases (Ras, RhoA and Rac) and G-proteins [435]. Nuclear Factor of Activated T-cells (NFAT) occupies a central role in these pathways [436]. Activation of NFAT is known to mediate apoptosis, with pro-apoptotic signalling including up-regulation of COX2 and the FAS death pathway [437, 438], and the association of inducer candidates with this pathway vindicates their pro-apoptotic activity.

Signalling by Rho GTPases (REACT_11044) was identified by DAVID with eight gene counts (CHN1, SYDE2, MYO9A, RHOT1, RHOA, KALRN, FGD2 and ARHGEF16). The Rho family of small GTP-binding proteins are signalling molecules activated by stimuli ranging from cytokines, growth factors, hormones, G-proteins and other biologically active substances. They are involved in processes such as reorganisation of actin cytoskeleton, regulation of transcription activity, vesicle trafficking, morphogenesis, apoptosis and tumourigenesis [439].

KALRN, FGD2 and ARHGEF16 are known to exhibit pro-apoptotic activity. Kalirin (KALRN) promotes the exchange of GDP by GTP and activates specific Rho GTPase family members to trigger regulatory signalling involved in the control of neuronal shape, growth and plasticity [440]. FYVE, RhoGEF and PH domain-containing protein 2 (FGD2) also promotes the exchange of GDP for GTP and activates CDC42, a member of the Ras-like family of Rho- and Rac proteins. CDC42 then activates JNK1 signalling. FGD2 is also able to bind a range of phosphoinositides. Rho guanine nucleotide exchange factor 16 (ARHGEF16) acts as a Guanyl-nucleotide exchange factor of the RHOG GTPase to promote the exchange of GDP for GTP [441] and could also activate CDC42 [442].

Other interesting pathways identified by IPA are Ephrin B Signalling, Cholecystokinin/Gastrin-mediated Signalling and Actin Nucleation by ARP-WASP Complex, while those from GSEA includes Phase II conjugation pathways, IL6 signalling, neuroactive ligand receptor interaction and response to gonadotrophins.

Ephrin B is a subfamily of the Eph tyrosine kinase (Tk) receptor family with six members involved in the regulation of cellular function such as cell-cell adhesion and are important in tissue development [443]. Cholecystokinin/Gastrin-mediated Signalling involves gastrointestinal peptides such as gastrin and cholecystokinin (CCK) which act as messenger molecules for signal exchange within the organism. For example, Gastrin is a circulating hormone able to stimulate acid secretion from the parietal cells; it is also a growth factor. The binding of these gastrointestinal peptides to their surface receptors (GPCRs) results in production of DAG and IP3, and the subsequent activation of IP3 signalling pathways, while Rho and MAPK signalling are also involved.

Phase II conjugations are biotransformation pathways involving transferases which transfer cofactors groups to the target substrate to increase the excretory potential of the compounds. The pathway generally results in inactivation or detoxification, although bioactivation may also occur.

As part of an integrated network of genes identified by IPA, the functions associated with the top network terms including haematological disease, hereditary disorder, infectious disease, drug metabolism, lipid metabolism, small molecule biochemistry, cellular development, cellular assembly

and organisation, cell-to-cell signalling and interaction, cellular function and maintenance, cellular movement and development disorder. Analysis of the integrated networks of known interactions revealed significant interactions among the inducer candidates, notably via UBC and Gpc (Figure 41).

IPA also provides information on the upstream regulators of the candidates. The inducers do not appear to be regulated under similar upstream signals, but clusters of genes able to be triggered by the same signal could be identified. The top five upstream regulators for the inducer candidates are MAGI2, tropomyosin, ZNF148, PTIO and MGEA5 (Figure 42). MAGI2 is a membrane-associated guanylate kinase which acts upstream of PTEN and RHOA, two of the inducer candidates. MAGI2 could for example improve the ability of PTEN to inhibit Akt1 activation [444]. Tropomyosin is upstream of TPM2 and TPM3, all of which are involved in the troponin complex associated with the calcium dependent regulation of vertebrate striated muscle contraction. ZNF148 is a zinc finger protein involved in transcriptional regulation and acts on ITGAM, PTCRA and VIM while PTIO is a chemical reagent which targets ITGAM and NCAM1. PTIO is a stable radical scavenger for nitric oxide (NO) and possesses significant inhibitory activity against NO biological actions. MGEA5 is an enzyme which specifically cleaves N-Acetylglucosamine but not N-Acetylgalactosamine from glycopeptides [445] and targets ABLIM1, ADAM19, AHNAK2, APOBEC3G, CD99L2, CDCA7, DAB2, FGFR1, PHKA2 and RPS6KA3 from the inducer candidates; it could potentially act as a regulator of apoptosis by changing the glycosylation status of its targets.

The top diseases and disorders enriched by IPA among the inducer candidates are inflammatory response, gastrointestinal disease, hepatic system disease, neurological disease and cancer.

The top toxicity terms include increases Bradycardia, decreases depolarisation of mitochondria and mitochondrial membrane, pro-apoptosis, cardiac fibrosis and TGF-β signalling while the top toxicological function includes increased levels of AST, red blood cells and haematocrit.

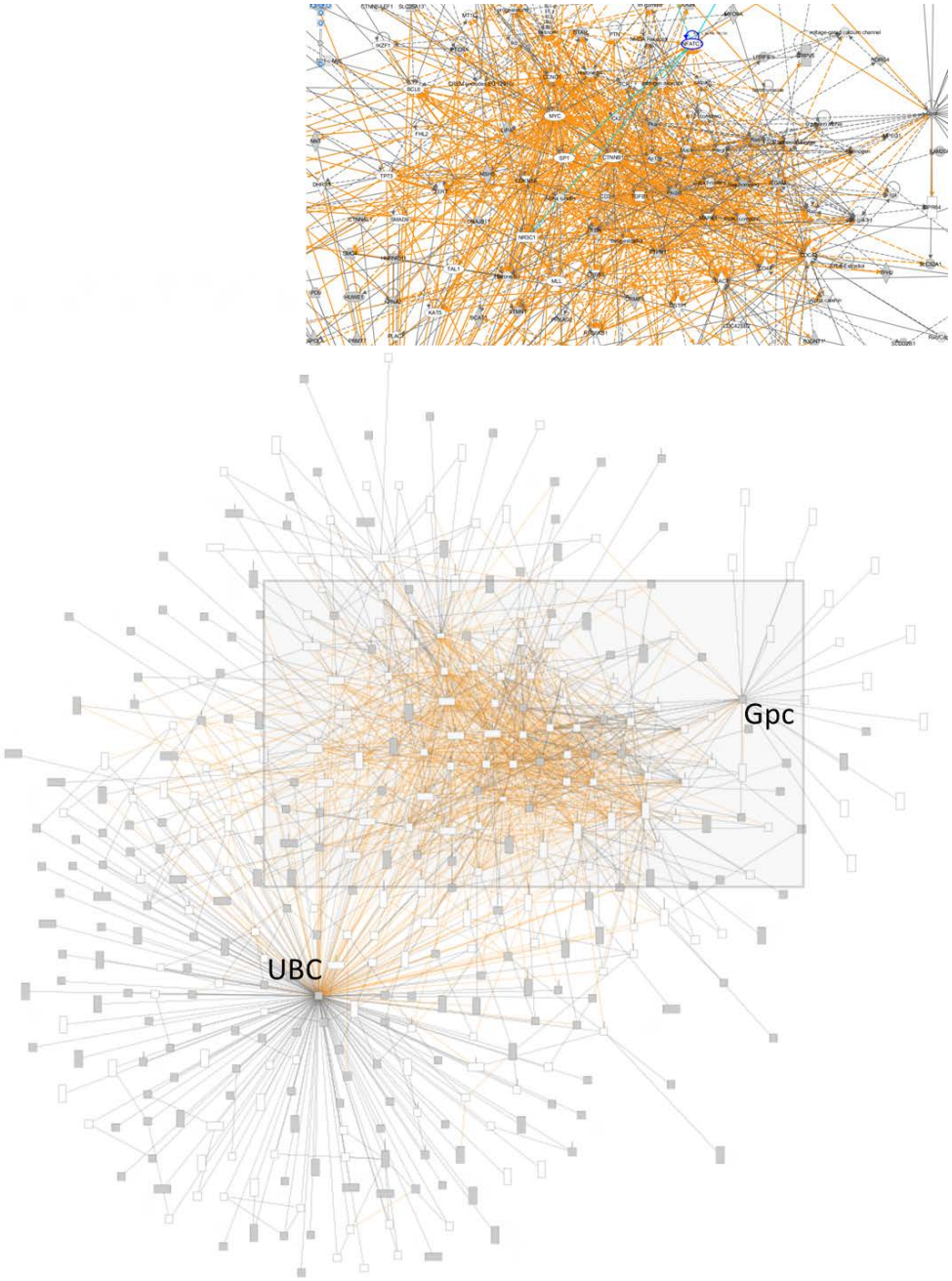


Figure 41: Integrated network for all known inducer interactions. All the networks identified by IPA were merged into a single network, with the orange lines indicating significant interactions. The proteins with the most interactions are potential signalling hubs and changes in the expression or activity of these hubs could potentially affect its interacting partners. An enlarged view of the region of interest with the highest density of significant interactions is inset in the top right corner. Candidate genes are shaded grey.

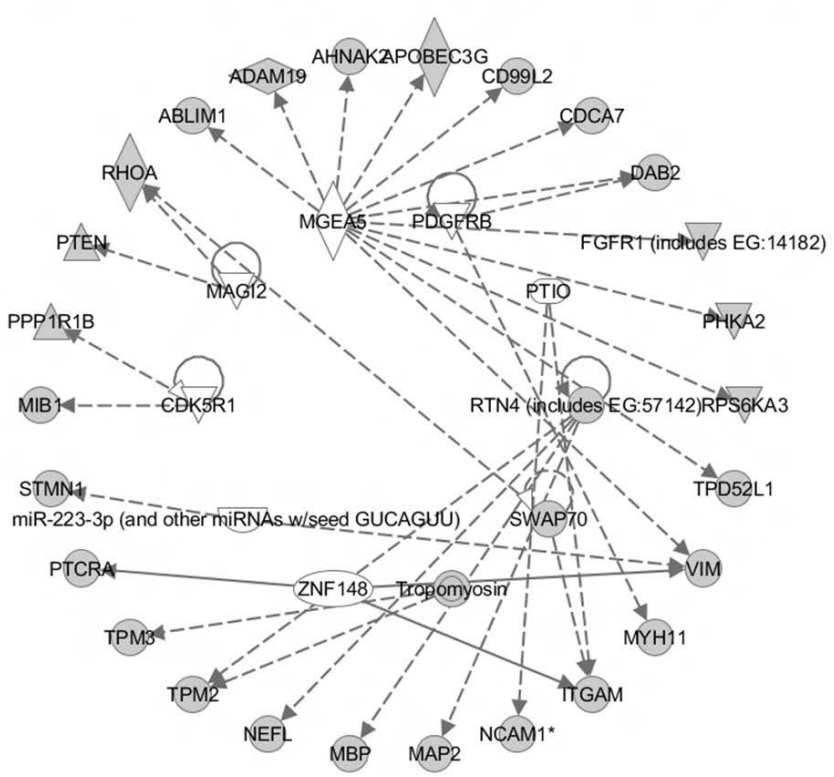


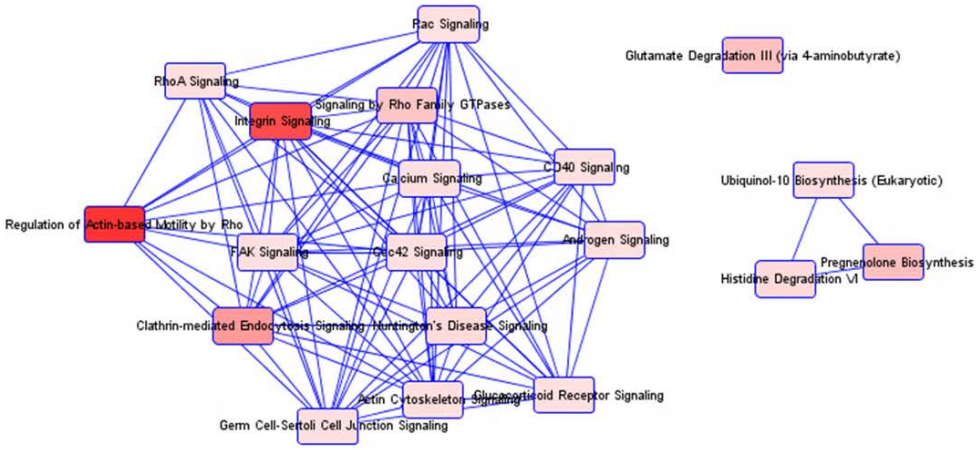
Figure 42: Top ten upstream regulators for inducers. Upstream regulators may be identified using IPA although the inducer candidates did not feature control of a large cluster of target genes by a single regulator. The five most significant upstream regulators are MAGI2, tropomyosin, ZNF148, PTIO and MGEA5.

Inhibitors

Using DAVID (Figure 40), five signalling pathways were identified and these are spliceosome (hsa03040), processing of Capped Intron-Containing pre-mRNA (REACT_125), cytoskeletal regulation by Rho GTPase (P00016), Gene Expression (REACT_71) and Alzheimer disease-amyloid secretase pathway (P00003). The majority of the enrichment by GSEA was for gene sets associated with cancer profiles, and only two (ZHANG_RESPONSE_TO_IKK_INHIBITOR_AND_TNF_DN and ELVIDGE_HIF1A_AND_HIF2A_TARGETS_UP) were linked to known signalling networks. IPA (Figure 43) yielded multiple pathways including regulation of actin-based motility by Rho, integrin signalling, clathrin-mediated endocytosis signalling, pregnenolone biosynthesis, glutamate degradation III (via 4-aminobutyrate), signalling by Rho family GTPases, Huntington's Disease signalling, histidine degradation VI, thiamin salvage III, histamine biosynthesis and oleate biosynthesis II (Animals) pathways.

Spliceosome (hsa03040) and processing of capped intron-containing pre-mRNA pathways (REACT_125) were both associated with transcriptional activity, in particular mRNA splicing. This process removes the noncoding introns from the protein-coding exons resulting in a final matured mRNA transcript. Spliceosomes are comprised of a dynamic family of particles which assemble on the mRNA precursor to help it achieve the conformation for the splicing to proceed. The standard spliceosome is comprised of five small nuclear reibonucleoproteins or snRNPs (U1, U2, U4, U5 and U6) and other spliceosome-associated proteins (SAPs). Both of these mRNA processing pathways in combination with the more general gene expression network (REACT_71) which covers the pathways from regulation of various transcription processes from the transcription of genomic DNA to how the new RNA transcripts are processed yielded 30 unique genes from the inhibitor candidates. The processing of capped intron-containing pre-mRNA is a subset of the gene expression network, with all 11 associated genes also enriched in the latter pathways. DHX15, HSPA8, MAGOHB, PRPF38B, PRPF40B and TRA2B are the six genes unique to the spliceosome pathway while the gene expression pathways consisting of 17 unique genes (CARS2, CCAR1, DHX9, EEF1G, HNRNPH1, IARS2, LARS, MAMLD1, NARS2, PAPOLA, POLR2F, RPL10A, RPS3, SUPT16H and TAF11). Six genes (NHP2L1, PCBP1, PRPF8, SF3A1, SRSF2 and SRSF7) are common to the spliceosome and gene expression pathways; Poly(rC)-binding protein 1 (PCBP1) is a DNA binding protein with a preference for cytosine while the remaining five are various splicing factors associated with spliceosome formation.

A) Integrated map of the enriched signalling pathways



B) Pathway chart of the top signalling pathways

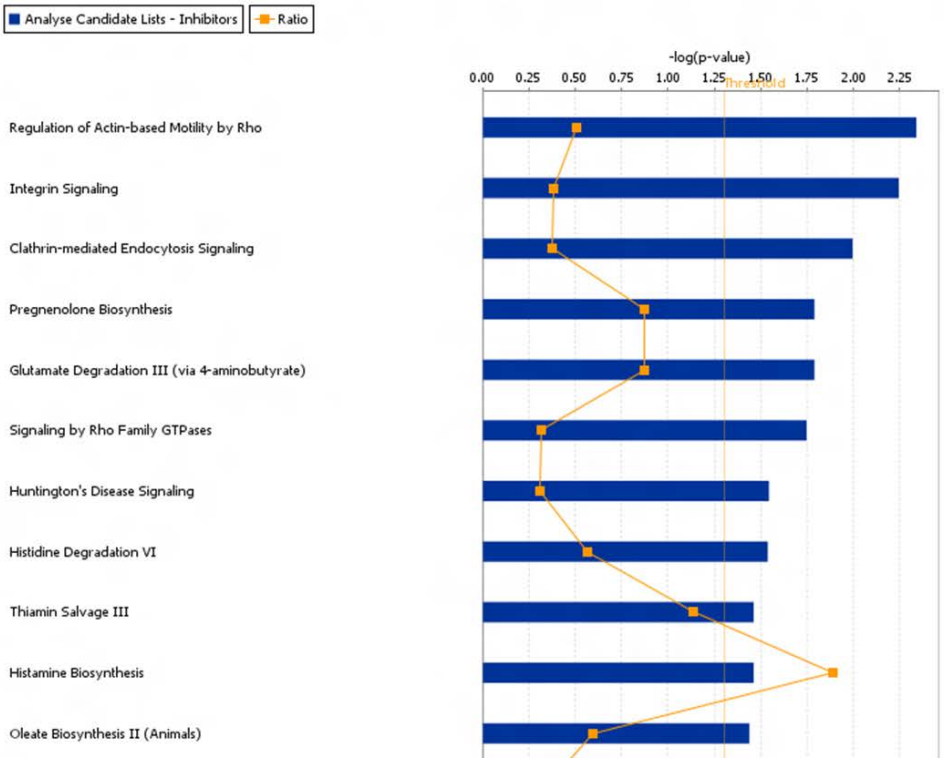


Figure 43: Signalling pathways identified for the Inhibitors. A) An overview showing all the signalling pathways and their relationships. The intensity of the shade of red increases with the significance of the signalling pathway. B) Pathway chart showing the most significant signalling pathways. The yellow ratio line represents the proportion of the candidate genes identified among each curated gene set.

Regulation of the cytoskeleton by Rho was identified by both DAVID and IPA with the following unique genes: DIAPH3, SSH3, TTN, ARHGAP32, SSH1, ASPM, MYH11, TUBB6, ARPC5, RHOF, ACTG2, BAIAP2, GSN, MYL9, PI4KA and RND2. This is likely to be closely associated with the signalling by Rho family GTPases pathway, which was also identified in the inducer candidates. The Rho signalling pathway was associated with 8 inducer candidate genes, while 14 was enriched among the inhibitor candidates (ACTG2, ARPC5, BAIAP2, DIAPH3, GNB1, GNG7, MAPK1, MYL9, PARD3, PI4KA, PIK3C3, RHOF, RND2 and SEPT4). The genes associated with Rho signalling are distinct to the inducer and inhibitor candidates, and this observation reinforces the involvement of Rho signalling in cell proliferation and apoptosis. It could also highlights the pro- and anti-apoptotic points along the Rho signalling pathway as indicated by the respectively candidate types, which could be exploited as treatment or biomarkers in the relevant clinical setting.

Alzheimer disease-amyloid secretase pathway and Huntington's disease signalling pathway form another interesting pair of enriched signalling pathways. Both are neurodegenerative diseases manifesting as a result of accumulation of protein aggregates of beta amyloid or mutant huntingtin. Six genes (PCSK5, CHRM3, CACNB2, APBA2, MAPK1 and PCSK6) are associated with the Alzheimer disease-amyloid secretase pathway and 13 are involved in Huntington's disease signalling (CAPN2, CAPN3, CLTC, CPLX2, GNB1, GNG7, GOSR1, HDAC2, HSPA8, HTT, MAPK1, PIK3C3 and POLR2F), both sets of genes do not appear to overlap apart from MAPK1. Mitogen-activated protein kinase 1 (MAPK1) is a serine/threonine kinase central to the MAP kinase signal transduction pathway, and depending on the signals and cellular context, the MAP signalling cascade could mediate a range of biological functions from cell proliferation, survival and differentiation through to the regulation of transcription, translation and cytoskeletal rearrangements. MAP signalling was previously enriched in the inducers, but may play a pro-survival role in association with the inhibitor candidates.

Alzheimer disease-amyloid secretase pathway involves the cleavage of the amyloid precursor protein by beta-secretase (BASE) then subsequently by gamma-secretase to generate the beta amyloid fragment found in senile plaques in Alzheimer's disease patients' brains. This pathway is different from the normal cleavage pattern of the amyloid precursor, which generates the benign p3 fragment first by alpha-secretase cleavage (TACE, ADAM10 or MDC9) then by gamma-secretase (presenilin complex) [446, 447]. Within the genes identified with this pathway, PCSK5 and PCSK6 possess endoprotease activity which cleaves at the RX(K/R)R consensus motif [448, 449]. CHRM3 is the muscarinic acetylcholine receptor that mediates cellular responses such as the inhibition of adenylate cyclase, phosphoinositides degradation and regulation of potassium channels through the action of G proteins, and hence is capable of influencing the various associated signal transduction

pathways [450]. CACNB2 is the beta subunit of voltage-dependent calcium channels which contributes to the calcium channel function by shifting the voltage dependencies of activation and inactivation and controlling G protein inhibition [451], APBA2 is associated with putative protein transport functions and MAPK1 is one of the key signalling nodes of the MAPK signal transduction.

Huntington's disease (HD) generally involves the mutations of the huntingtin gene (htt) which translate into extended polyglutamine tracts (polyQ) resulting in insoluble misfolded protein aggregates. The HD mutation also causes abnormal protein interactions such as the interference of the PSD-96 substrate interaction with its NMDA receptor by the mutant htt, leading to increased calcium influx. The HD mutation leads to sequestration of proteins including transcription factors. The mutant htt is also able to disrupt gene transcription by recruiting transcription factors like CBP, TBP and Sin3A via its cleaved N-terminal fragments.

CAPN2 is a calcium-regulated non-lysosomal thiol-protease able to catalyse the proteolysis of substrates involved in signal transduction and cytoskeletal remodelling [452]. CAPN3 is also calcium-regulated protease and has been demonstrated to exhibit an anti-apoptotic activity through NF-kappaB-dependent expression of the anti-apoptotic factor c-FLIP [453]. CLTC is a major protein component of coated pits and vesicles, CPLX2 negatively regulates the formation of synaptic vesicle and is down-regulated in Huntington disease [454, 455]. GNG7 is a G-protein involved in signal transduction, GOSR1 is involved in ER-Golgi intra-organelle transport [456], HDAC2 catalyses the deacetylation of lysine residues on the core histones (H2A, H2B, H3 and H4) [457] and can negatively influent apoptosis via its control of p53 [458]. HSPA8 is a chaperone associated with protein folding function and stress response, and is capable of acting as a repressor of transcriptional activation such as the inhibition of the transcriptional coactivator activity of CITED1 on Smad-mediated transcription [459, 460]. HSPA8 is also regulated by the anti-apoptotic BAG-1 [461], which may facilitate its apoptosis inhibitory behaviour. HTT is the primary protein involved in Huntington's disease and may be involved in microtubule-mediated transport function. HTT is also associated with the negative regulation of apoptosis [462]. PIK3C3 is the catalytic subunit of the PI3K complex which regulates the production of phosphatidylinositol 3-phosphate, which in turn regulates the formation of autophagosomes [463] while POLR2F is a DNA-dependent RNA polymerase also associated with the spliceosome pathway.

It remains unclear how the inhibitor candidates associated with the Alzheimer and Huntington's neurodegenerative disease pathways exhibit their anti-apoptosis effect, although it may be speculated that this is achieved through transcriptional mechanisms or through the pro-survival signals as a result of stress response (HSPA8) or autophagy (PIK3c3). CAPN3, HDAC2 and HTT were

demonstrated to be negative regulators of apoptosis. It is interesting that the inhibitor candidates may confer pro-survival mechanisms under disease conditions such as Alzheimer and Huntington's along with various cancer profiles.

IPA also identified six pathways involved in cellular metabolism: pregnenolone biosynthesis, glutamate degradation III (via 4-aminobutyrate), histidine degradation VI, thiamin salvage III, histamine biosynthesis and oleate biosynthesis II. The genes involved are CYP11A1, MICAL2, GAD1, MYO5B, TPK1, HDC, FADS1 and FADS2. The mechanism by which these pathways associated with synthesis or degradation of biomolecules translates into an anti-apoptotic response is unknown, and may represent novel crosstalk between the metabolic and apoptosis pathways.

Integrin signalling and clathrin-mediated endocytosis signalling were also identified by IPA with 14 and 13 genes respectively. Integrins are receptors which mediate the attachment between the cells and its extracellular matrix while clathrin-mediated endocytosis involves the uptake of surrounding molecules through clathrin-coated vesicles. Integrin signalling is able to activate PI3K and MAPK signalling to modulate cell proliferation or apoptosis, and integrins are able to shift the balance between pro- and anti-apoptosis signals. For example in anoikis, the absence of a survival signal provided by integrin anchorage would trigger apoptosis. Alternatively, integrin signalling is also able to increase resistance to apoptosis directly via pathways which target the Bcl-2 and IAP families of apoptosis regulators [464]. The anti-apoptosis effect via the clathrin-mediated endocytosis pathway may partially be attributed to the reduction of cell surface expression of death receptors [465].

The top upstream regulators of the inhibitor candidates are miR-1 (GGAAUGU), tropomyosin, DL-threo-dihydrosphingsosine, HNF-4A and miR-125bp-5p (CCCUGAG). Unlike the upstream regulator profile of the inducer candidates which appeared to have little correlation, the regulators of the inhibitor candidates seemed to be able to target a wide range of genes. Among these, the transcription factor HNF-4A appears to have the widest target cluster (Figure 44), while two microRNAs (miR-1 and miR-125bp-5p) were also identified. Both microRNAs identified by IPA did not appeared in the GSEA enrichment, likely as a result of the different method of enrichment algorithm employed between the two analysis methods. miR-1 is important in cancer where its expression was shown to be reduced while its target genes were up-regulated [466, 467]. It is also shown to possess tumour suppression activity via the regulation of LASP1 [468], and may be hypothesised to act through some of the inhibitor candidates identified in the screen. The activity of miR-125bp-5p is currently not well established. Tropomyosin is also identified as an upstream regulator of inducer candidates, while DL-threo-dihydrosphingsosine is a chemical inhibitor of sphingosine kinase which is involved in calcium mediated signalling [469].

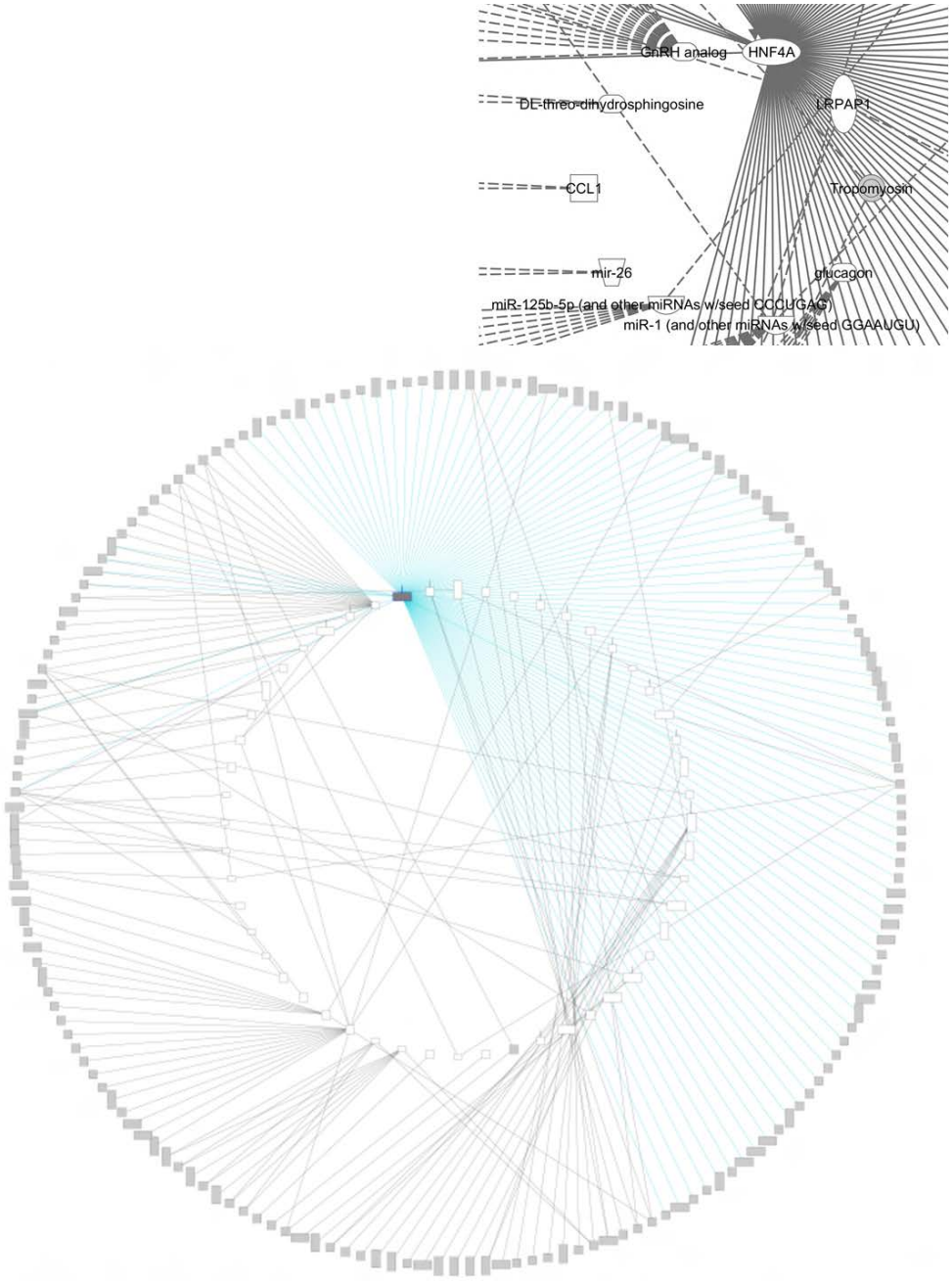


Figure 44: Upstream regulators for Inhibitors. All upstream regulators for the inhibitor candidates with the focus on HNF4A (blue) indicating the wide extend of its target genes, arranged in around the circular peripheral. The top ten upstream regulators are inserted in the top right corner.

The top networks identified by IPA are associated with the listed terms: hereditary disorder, neurological disease, skeletal and muscular disorder; connective tissue development and function, embryonic development, nervous system development and function; cell-to-cell signalling and interaction, tissue development, neurological disease; lipid metabolism, small molecule biochemistry, dermatological diseases and conditions.

The top IPA biological functions associated with the inhibitors are infectious disease, neurological disease, developmental disorder, hereditary disorder and metabolic disease, while toxicological functions includes decreased levels of albumin, increased levels of ALT, increased levels of Hematocrit, increased levels of albumin, increased levels of blood urea nitrogen. Since the inhibitors are acting in response to the silica nanoparticles, these toxicological functions may be proposed as potential biomarkers upon exposure to the same nanoparticles although this hypothesis needs to be verified experimentally.

The top toxicity list includes renal ischemic resistance panel (rat), cardiac necrosis/cell death, increased cardiac proliferation, decreased permeability transition of mitochondria and mitochondrial membrane and cytochrome P450 panel (human).

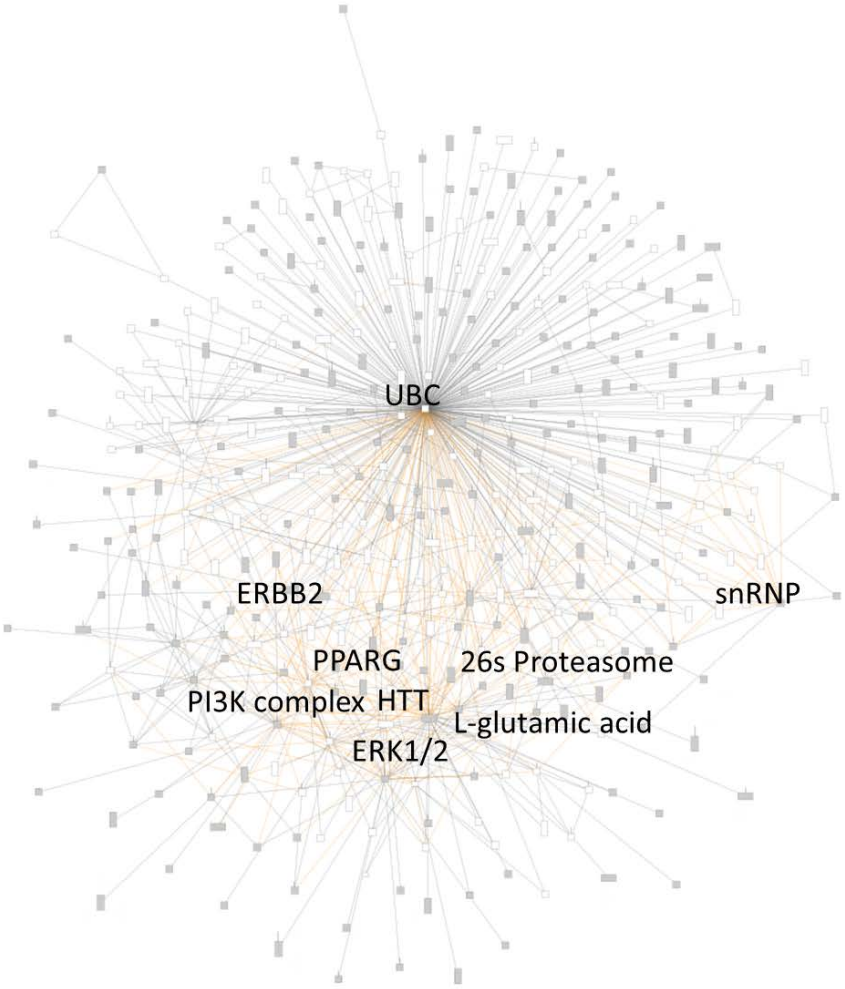


Figure 45: Integrated networks of all known inhibitor interactions. All the networks identified by IPA were merged into a single network, with the orange lines indicating significant interactions. The proteins with the most interactions are potential signalling hubs and changes in the expression or activity of these hubs could potentially affect its interacting partners. Candidate genes are shaded grey.

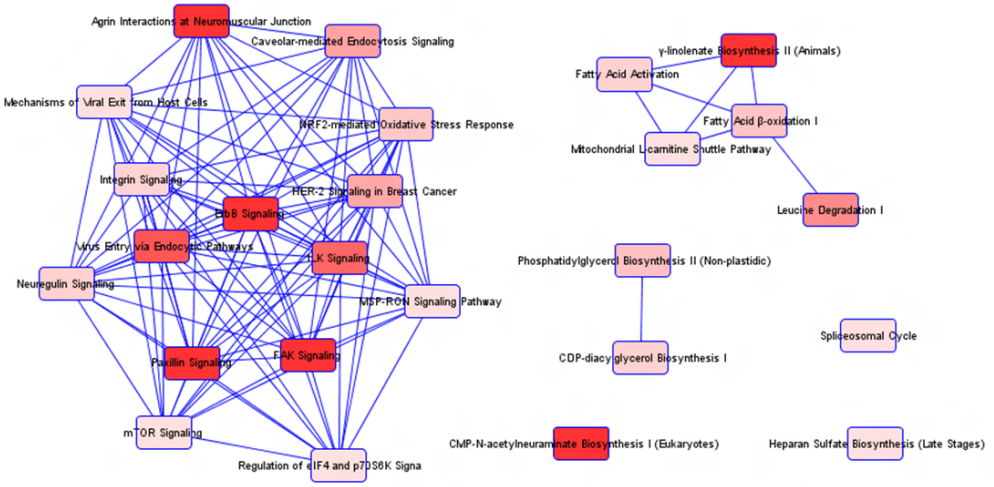
Sensitizers

Pathways involving endocytosis (hsa04144) and regulation of actin cytoskeleton (hsa04810) were identified by DAVID (Figure 40). GSEA enrichment for the sensitizer tail indicated fifteen pathways including NCAM1 interactions, translation initiation complex formation, inflammatory response to LPS, peroxisome, oocyte meiosis, formation of ternary complex and subsequent 43S complex, CD40 signalling, signalling by PDGF, regulation of beta cell development, GTP hydrolysis and joining of the 60S ribosomal subunit, ribosome, viral mRNA translation, regulation of the actin cytoskeleton by Rho GTPases, NCAM signalling for neurite out growth and IL6 signalling.

IPA (Figure 46) identified another twelve signalling pathways including Paxillin signalling, FAK signalling, agrin interactions at neuromuscular junction, CMP-N-acetylneuraminate biosynthesis I (Eukaryotes), y-linolenate biosynthesis II (Animals), ErbB signalling, virus entry via endocytic pathways, ILK signalling, leucine degradation I, caveolar-mediated endocytosis signalling, HER-2 signalling in breast cancer and phosphatidylglycerol biosynthesis II (non-plastidic).

Whereas the transcriptional pathways were associated with the inhibitor candidates, pathways involved in the translation of mRNA into proteins appear to be prevalent in the sensitizer candidates with five pathways enriched by GSEA. 14 unique genes (RPS27A, RPS4X, RPS29, RPS16, RPS6, EIF4G1, EIF5B, FAU, RPS9, RPL9, RPL12, UBA52, RPLP1 and DNAJC3) were identified among the four translation related pathways. RPS27A, RPS4X, RPS29, RPS16, RPS6, FAU, RPS9, RPL9, RPL12, UBA52 and RPLP1 are ribonucleoproteins associated with various aspects of translational activity such as initiation, elongation and termination. EIF4G1 is a translation initiation factor that is part of the complex which recognises the mRNA cap and initiates translation while EIF5B promotes translation initiation by enhancing the binding of formylmethionine-tRNA to ribosomes. UBA52 is interesting in that it is also a precursor for ubiquitin, which is generated via a cleavage at the N-terminal freeing a single ubiquitin molecule from the 60S ribosomal protein L40. UBA52 is involved in apoptosis with both anti-apoptotic and pro-apoptotic (via extracellular signal) capabilities as annotated by Reactome. DNAJC3 is a chaperone involved in the unfolded protein response during ER stress and is activated in response to influenza viral infection [470]. Intriguingly, 11 / 13 genes including DNAJC3 may also be hijacked for the translation of viral mRNA, and the association of these sensitizer candidates may act to prime the cell to or even activate apoptosis as antiviral defence response.

A) Integrated map of the enriched signalling pathways



B) Pathway chart of the top signalling pathways

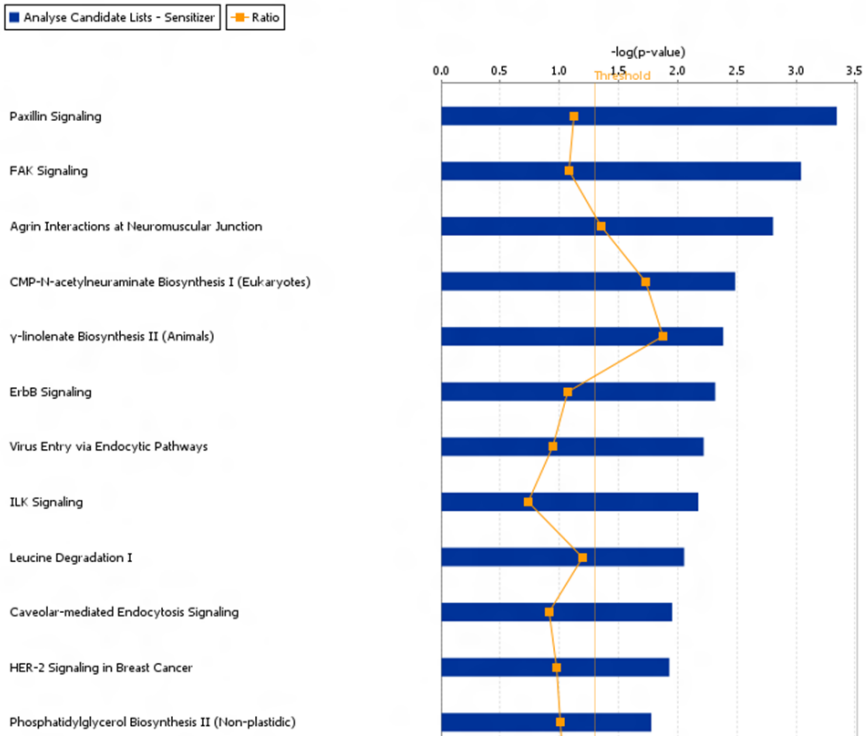


Figure 46: Signalling pathways identified for the sensitizers. A) An overview showing all the signalling pathways and their relationships. The intensity of the shade of red increases with the significance of the signalling pathway. B) Pathway chart showing the most significant signalling pathways. The yellow ratio line represents the proportion of the candidate genes identified among each curated gene set.

Endocytosis is another well represented pathway with three associated terms: endocytosis (hsa04144), caveolar-mediated endocytosis signalling and virus entry via endocytic pathways. 14 genes were associated with these pathways: ASAP2, EGFR, GIT2, EHD2, VPS37B, ARFGAP1, ARAP3, EPN2, ACTB, ACTC1, DNM2, ITGB5, PIK3R1 and PRKCE. 5 / 6 genes associated with viral entry overlapped with the other to endocytosis pathway, although PRKCE is unique to this viral association. Protein kinase C epsilon type (PRKCE) is a serine/threonine-protein kinase with important functions in the regulation of biological processes linked to cytoskeletal proteins including cell adhesion, migration, motility and cell cycle, functions in neuron growth and ion channel regulation. PRKCE is also involved in immune response, cancer cell invasion and regulation of apoptosis. Caveolae-dependent endocytosis is important in the internalisation of the scavenger receptor-substrate complex to initiate apoptosis [471] and endocytosis in general is widely associated with the modulation of virus-induced apoptosis [472]. In addition, an increase in the rate of endocytosis may correlate with increase of apoptosis since more silica nanoparticles are taken up by the cells.

Pathways associated with development are also well represented. NCAM1 interactions and NCAM signalling for neurite out growth as identified by GSEA are associated with the growth of neurons; NCAM1 is a cell adhesion protein involved in neuron-neuron adhesion, fasciculation and outgrowth of neurons. The pathway involving agrin interactions at neuromuscular junction, identified by IPA, is associated with the development of the neuron-muscle interface; agrin (AGRN) is a protein secreted by motor neuron axons during development which binds to the surface of skeletal muscle [473]. The three pathways associated with neuron development enriched for 15 genes (COL9A1, CACNA1C, COL3A1, COL9A3, COL6A3, NCAN, SPTBN4, PTK2, AGRN, ACTB, ACTC1, EGFR, ERBB2, NRG3 and PAK6). COL9A1, COL3A1, COL9A3 and COL6A3 are variants of collagen chains used for cell adhesion functions, SPTBN4 is a structural constituent of the cytoskeleton [474] while ACTB is the cytoplasmic actin, ACTC1 (alpha cardiac muscle 1) are actin molecules also associated with the cytoskeleton. NCAN binds to hyaluronic acid and regulates neuronal adhesion and neurite growth during development. NRG3 is a ligand for the ERBB4 tyrosine kinase receptor [475].

ERBB2 (a.k.a HER2) is a protein tyrosine kinase which is involved in several cell surface receptor complexes and is a key component of a neuregulin-receptor complex. It regulates neuron outgrowth and stabilisation of peripheral microtubules [476]. PTK2 is a non-receptor protein-tyrosine kinase involved in regulating cell migration, adhesion, reorganisation of the actin cytoskeleton, cell cycle progression, proliferation and apoptosis. PAK6 is a serine/threonine protein kinase involved in the

regulation of gene transcription [477] and may potentially exhibit an anti-apoptotic effect through the phosphorylation of BAD. [478]

CACNA1C is a voltage-sensitive calcium channel (VSCC) regulating the entry of calcium ions into excitable cells. It is also associated with various calcium-dependent processes such as muscle contraction, release of hormone or neurotransmitter, gene expression, cell motility, division and death [479]. NCAM signalling is also thought to mediate pro-apoptotic signals induced by IL-1beta in beta cells [480]. Regulation of beta cell development and oocyte meiosis were also development-associated pathways identified by GSEA with 12 genes each. Beta cells are located in the islets of Langerhans in the pancreas and are implicated in diabetes. Oocyte is a female germ cell involved in reproduction.

Signalling by platelet-derived growth factor (PDGF) is also closely associated with the development process, the signalling pathway being enriched by GSEA with 7 genes (COL9A1, STAT1, COL3A1, PDGFB, COL9A3, COL6A3 and CRK). STAT1 is a signal transducer and transcription activator involved in various cellular responses to cytokines, interferons and growth factors, and is capable of induction of apoptosis [481]. PDGFB is a growth factor associated with the regulation of embryonic development, cell proliferation and survival. CRK is an adaptor protein which mediates MAPK8 activation in a Rac-dependent manner [482]; CRK is also a proto-oncogene.

FAK signalling (7 genes: ACTB, ACTC1, ARHGAP26, EGFR, GIT2, PAK6 and PIK3R1) involved the activation of focal adhesion kinase-1 (FAK) by focal adhesion complex associated growth factors and integrins. FAK activation in turns promotes activation of signal transduction pathways which in turns promote cell migration and may have a role in the regulation of p53. FAK signalling is critical in early development and its loss is lethal to the cells. FAK can regulate cell growth and survival through the activation of PI3K/PKD1/Akt/PKB and Grb2/SOS/Ras/Raf-1/MEK/ERK pathways.

Paxillin signalling (8 genes: ACTB, ACTC1, ARF1, GIT2, ITGB5, PAK6, PARVA and PIK3R1) plays an important role in mediating signal transduction from integrin receptors to the actin cytoskeleton. Paxillin is a signal transduction adapter protein that localises with integrin-β1, FAK, vinculin and certain kinases at focal adhesions. It connects the integrin signalling with p38 MAPK and JNK pathways. Its primary function is as a molecular adapter or scaffold protein to create an interface with multiple docking sites at the plasma membrane for interactions between signalling proteins like kinases and structural proteins [483]. These kinases could then phosphorylate the N-terminus of paxillin allowing for recruitment of downstream effectors like Crk which can then modulate gene

expression via various MAPK cascades. Paxillin signalling results in changes in cytoskeleton structures important for events like embryonic development, wound repair and tumour metastasis.

ILK signalling (9 genes: ACTB, ACTC1, FN1, ITGB5, LIMS2, PARVA, PIK3R1, RSU1 and TMSB4X) involves the integrin linked kinase (ILK) which primarily functions to connect integrins to the cytoskeleton. ILK recruits other adaptor molecules into a large complex to regulate actin dynamics and integrin function. ILK is also important in cancer progression, and has been validated as a cancer therapeutic target [484, 485].

The regulation of actin cytoskeleton yielded two pathway terms by DAVID and GSEA with 14 unique genes: SSH3, TTLL3, FN1, ITGB5, ACTB, PAK6, EGFR, TMSB4X, RPS4X, AKT1, PAK4, FSCN1, CFL1 and WASF1. The association of the actin cytoskeleton across the three candidate types is intriguing and suggests that the actin cytoskeleton may be an important area for the control of apoptosis. Indeed, the actin cytoskeleton dynamics may regulate apoptosis signalling through the release of reactive oxygen species and rearrangement of the cytoskeleton from the mitochondria [486].

Three pathways associated with the immune system and inflammatory response were also enriched. CD40 signalling (7 genes: RGS1, TCEB3, ARID5B, PHACTR1, CFLAR, STAT1 and IRF4) centres on CD40, a member of the tumour necrosis factor receptor (TNFR) family and mediates signals between proliferation, differentiation, growth suppression and cell death. The pathway is important for B-cells survival and production of various molecules associated with the immune system such as costimulatory molecules, cytokines, chemokines (interleukins), TNF-alpha and cytotoxic radicals [487]. CD40-mediated signal transduction increases the transcription of target genes involved in host defence against pathogens. Up-regulation of these genes is achieved by activation of multiple pathways such as NF-KappaB, MAPK and STAT3 [488]. IL6 signalling (6 genes: PDPN, SOD2, DNM1, AGRN, CDC25B and STMN2) involves the cytokine interleukin-6 and is achieved through its receptor comprising of the alpha subunits for ligand specificity and GP (glycoprotein) 130. Binding of the ligand initiates the signalling including the activation of JAK kinases, Ras-mediated signalling and IP3K signalling. These signalling pathways translate into a variety of biological responses such as immune response, inflammation, gene activation, cell proliferation, survival and differentiation. Inflammatory response to bacterial lipopolysaccharide or LPS enriched for 7 genes: RAB20, SLC15A3, SLC11A2, SOD2, TNIP1, CXCL5 and CD44. RAB20 is involved in protein transport/recycling. SOD2 destroys superoxide radicals by catalysing the conversion of these superoxide into oxygen and hydrogen peroxide [489]. TNIP1 interacts with zinc finger protein A20/TNFAIP3 and suppresses TNF-induced NF-kappaB-dependent gene expression by impeding RIP- or TRAF2-mediated transactivation signal [490]. CXCL5 is a chemokine involved in neutrophil activation [491] while CD44 is a receptor for

hyaluronic acid involved in cell adhesion. SLC15A3 and SLC11A2 are both transporters involved in the transport of histidine and small peptides for the former, and metal ions such as iron, manganese, cobalt, cadmium, nickel vanadium and lead for the latter.

The ErbB/HER2 signalling pathway implicated in various cancer types including breast, ovaries, brain and prostate was identified by IPA with 7 unique genes: EGFR, ERBB2, NRG3, PAK6, PIK3R1, PRKCE and ITGB5. The ErbB receptor tyrosine kinase family comprises of four members (ErB1 – 4) that are typical surface receptor tyrosine kinases which become activated upon ligand binding. Receptor dimerization and subsequent phosphorylation enables the interaction with adaptor proteins to promote downstream signalling, with the range of effects including cell proliferation, migration, differentiation, and apoptosis. In addition, ErbB family proteins can also act as transcriptional regulators by translocating to the nucleus. For example, ErbB2 is able to interact with importin β1 and Nup358 to translocate to the nucleus via endocytic vesicles, where ErB2 then regulates target genes such as COX-2 [492, 493].

Five metabolic pathways were identified including peroxisome, leucine degradation I, phosphatidylglycerol biosynthesis II (non-plastidic), CMP-N-acetylneuraminate biosynthesis I (Eukaryotes) and y-linolenate biosynthesis II (Animals). Unlike the inhibitor candidates enriched for approximately equal counts of biosynthesis and degradation pathways, the metabolic pathways associated with the sensitizers appear to be largely anabolic. The peroxisome is involved in the breakdown of long fatty acid chains, D-amino acids and polyamines, but may also be involved in biosynthesis of plasmalogens, an ether phospholipid with a large presence in the nervous, immune and cardiovascular systems [494]. Phosphatidylglycerol is a glycerolphospholipid found in pulmonary surfactant. Interestingly, two units of phosphatidylglycerol would form cardiolipin, an important component of the inner mitochondrial membrane and a known trigger for apoptosis [495]. The connection of phosphatidylglycerol synthesis with cardiolipin may be a probable mode of action for the apoptosis sensitizing effect of the associated genes. CMP-N-acetylneuraminate is a common sialic acid in human, and the precursor for the synthesis of other sialic acids. Sialic acids are usually the terminal sugar residue of glycoproteins and glycolipids, but may also mediate cellular recognition and adhesion events during development, immune and inflammatory responses and oncogenesis. y-linolenate is a polyunsaturated fatty acid and a precursor for Δ^6-desaturated fatty acids like arachidonate and eicosapentaenoate which are associated with maintenance of membrane structure and function, and regulation of cholesterol synthesis and transport.

Like the inducer candidates, the sensitizers do not appear to have the regulatory pattern where a large number of genes are targeted by an upstream regulator as in the case of the inhibitor

candidates. The top five upstream regulators for the sensitizers are miR-1 (GGAAUGU), HTF, PLCD4, miR-7-5p (GGAAGAC) and IGFBP2 (Figure 47). HER2 transcription factor (HTF) is an activator protein-2 transcription factor involved in the regulation of ERBB2 gene expression [496]. PLCD4 hydrolyses phosphatidylinositol 4,5-bisphosphate (PIP2) to generated DAG and IP3, in a reaction similar to the PLCD2 identified as an inducer candidate, hence PLCD4 is likely to mediate its downstream target genes via PKC and IP3K signalling. Insulin-like growth factor-binding protein 2 (IGFBP2) binds to insulin-like growth factors (IGFs) to prolong their half-life, and is able to inhibit IGF-mediated growth by altering the interactions between the IGFs and their surface receptors [497]. The association of the microRNA miR-1 with the sensitizers is another intriguing find, since miR-1 was previously found to be an upstream regulator for many of the inhibitor candidates. While the pro-survival activity of miR-1 is better suited as the regulator for apoptosis inhibitor candidates, its association with the sensitizers may potentially acts as a negative feedback signal to prime the cells for apoptosis when its expression is critically reduced for example in the event of cancer development. Another microRNA, miR-7-5p, was also identified although its functions have yet to be established.

The top networks identified by IPA involves the following listed pathway terms: cellular assembly and organisation, cardiovascular disease, cell morphology; cellular assembly cellular function and maintenance, DNA replication, recombination and repair; post-translational modification, embryonic development and tissue morphology; cell morphology, endocrine system development and function and organ morphology; nervous system development and function, organ morphology and cell cycle. Analysis of the integrated network revealed UBC, HNF4A and miR-124-3p to be among the top signalling hubs (Figure 48).

The top biological functions include cancer, haematological disease, infectious disease, reproductive system disease and connective tissue disorders.

The top toxicity list includes irreversible glomerulonephritis biomarker panel (rat), p53 signalling, renal glomerulus panel (human), NRF2-mediated oxidative stress response and hypoxia-inducible factor signalling. The top toxicity functions are increased levels of creatine and red blood cells.

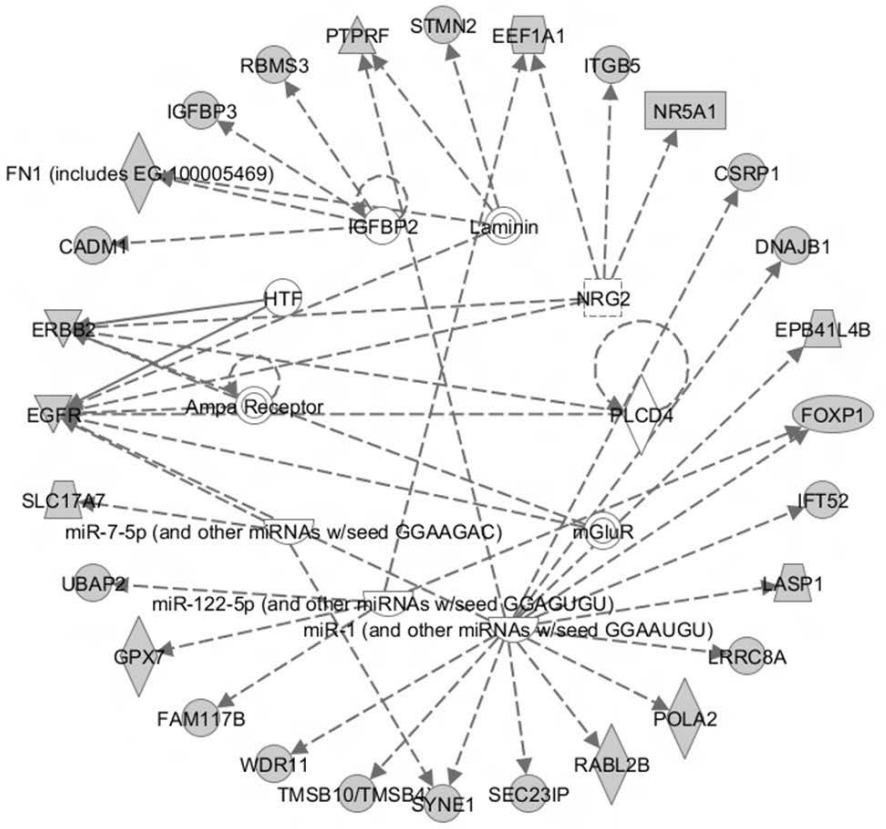


Figure 47: Top ten upstream regulators for sensitizers. Upstream regulators may be identified using IPA although the sensitizers, like the inducer candidates, did not feature control of a large cluster of target genes by a single regulator. The five most significant upstream regulators are miR-1 (GGAAUGU), HTF, PLCD4, miR-7-5p (GGAAGAC) and IGFBP2.

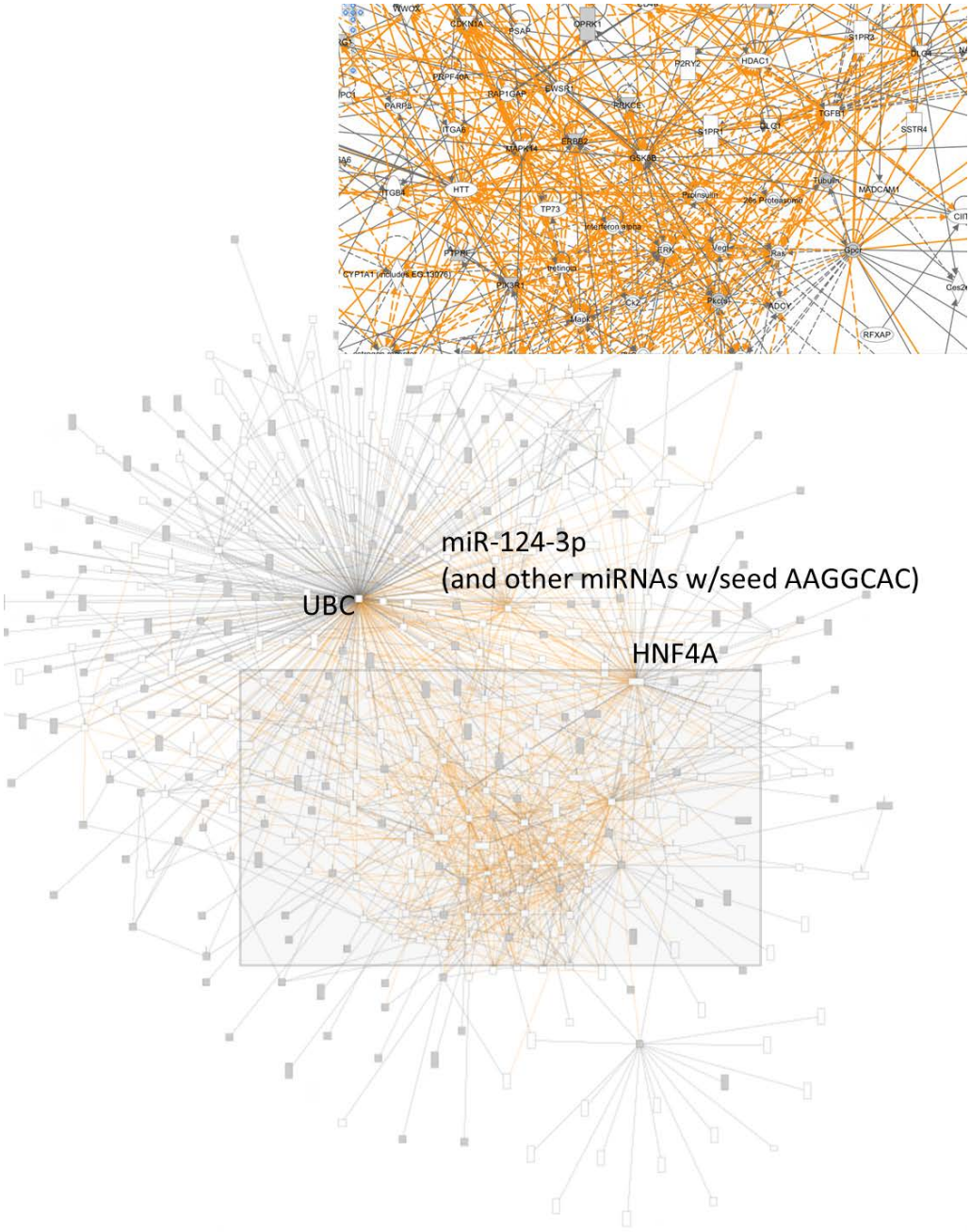


Figure 48: Integrated network for all known sensitizer interactions. All the networks identified by IPA were merged into a single network, with the orange lines indicating significant interactions. The proteins with the most interactions are potential signalling hubs and changes in the expression or activity of these hubs could potentially affect its interacting partners. An enlarged view of the region of interest with the highest density of significant interactions is inset in the top right corner. Candidate genes are shaded grey.

Analysis of Noncoding cDNA Sequences

The cDNA sequences which were not matched to any protein products either by the annotation provided by Dr Derek Huntley, or annotations generated from online databases such as UniProt, Ensembl Biomart, NCBI and bioDBnet were classified as noncoding sequences. The ability of noncoding sequences to regulate cellular functions and processes is an exciting field, with the ENCODE project revealing that the bulk of the previously defined junk DNA may in fact be involved in some form of regulatory mechanisms [498]. The hypothesis of the noncoding sequences acting as various form of RNA regulatory mechanisms such as microRNA was explored together with the potential of the open reading frames coding for mini-proteins which were also known to exert functional activity within cells [499].

The noncoding sequences comprised of 141 inducers, 300 inhibitors and 157 sensitizers. The sequences were submitted to Rfam database, a known collection of structural RNAs including noncoding RNA and cis-regulatory elements maintained by the Wellcome Trust Sanger Institute [500] to identified potential regulatory mechanisms.

Among the inducer candidates, AK022826 was shown to contain the SOX2 overlapping transcript exons 1 to 4 (SOX2OT). SOX2OT is a known long non-coding RNA which is involved in embryogenesis and vertebrate development [501]. Here, we have demonstrated that the overexpression of this transcript is capable of pro-apoptosis activity which may be mediated by the exons 1 – 4 of the SOX2OT transcript. AK127225 and AK124567 contained the GRIK4 3' UTR element and could be speculated to interfere with GRIK4 expression by interacting with its mRNA to change its half-life [502]. GRIK4 encodes a kainate receptor subtype which is a member of the ligand-gated ion channels family and the interference of this ligand-receptor interaction has been demonstrated to trigger neuronal cell death [503]. AK023600 possesses the U6 spliceosomal RNA, a non-coding small nuclear RNA (snRNA) which is a component of the spliceosome, making this cDNA sequence a potential U6 snRNA precursor. Free U6 snRNA is also known to interact with the proteins Prp24 and LSms in *S. cerevisiae* [504]. AK129752 and AK124771 were associated with the snoRNAs from RNA Z195/SNORD33/SNORD32 family and sR21 respectively. Small nucleolar RNA Z195/SNORD33/SNORD32 are snoRNAs able to modify other snRNAs and was found to be overexpressed in non-small-cell lung cancer [505]. AK124781 is associated with rydB RNA, a non-coding RNA found in *E. coli*, *Shigella flexneri* and *Salmonella* species [506] although the function of the rydB RNA is currently unknown. This intriguing link between abundance in common bacterial species and apoptosis induction may suggest that the sequence may act as some kind of internal cellular mechanism to trigger apoptosis in the presence of these foreign bacteria within the cells.

The inhibitor candidates also include associations with snoRNAs; AK125313, AK092456 and AK129752 were associated with Small nucleolar RNA SNORD95, Small nucleolar RNA SNORD75 and Small nucleolar RNA Z195/SNORD33/SNORD32 family respectively. Furthermore, the inhibitors were associated with four spliceosomal RNA: AK023614 (U2 spliceosomal RNA), AK124014 (U6atac minor spliceosomal RNA), AK124213 and AK023600 (U6 spliceosomal RNA). AK022419 was associated with Deleted in lymphocytic leukemia 2 conserved regions 1 – 6, which is a long non-coding RNA (lncRNA) encoded by the DLEU2 gene. The DLEU gene is associated with tumour suppressor activity and is frequently deleted in patients with B-cell chronic lumphocytic leukemia [507].

The inhibitors also resulted in six microRNA sequences, supporting our hypothesis that one of the potential mechanisms for these noncoding sequences was via generation of microRNAs to silence their target genes. Here, these microRNAs are implicated in the regulation of apoptosis, a regulatory mechanism currently novel to apoptosis signalling. The inhibitor candidates AK098506, AK001351, AK129585, AK096776 and AK025763 were associated with microRNAs mir-497, mir-198, mir-650, mir-149 and mir-280. AK098506 also contained two sites associated with the mir-16 microRNA precursor family within the same transcript, an observation which supports the role of AK098506 as the microRNA precursor, and its resulting microRNA mir-16 could potentially act to down-regulate pro-apoptosis genes. potC RNA and Pyrococcus C/D box small nucleolar RNA were two other RNA elements associated with the inhibitors AK024181 and AK058064. potC RNA is a recently discovered conserved RNA structure and a predicted cis-regulatory element (acting within the same DNA molecule) associated with the 5’ untranslated regions of genes encoding membrane transport proteins or peroxiredoxins [508]. Membrane transport proteins are highly enriched among the inducer candidates, and the association of a RNA regulatory element which could inhibit apoptosis through the down-regulation membrane transport proteins is indeed an extremely supportive finding. Furthermore, peroxiredoxins are universal antioxidant enzymes capable of regulating cytokine-induced peroxide levels (and possibly apoptosis via changes in the cellular oxidative state) and mediate signal transduction [509]. The potC RNA motif is currently known to be associated with marine bacteria, and this discovery of its presence in mammalian sequences with two potential target classes both associated with apoptosis is yet another novel finding. Pyrococcus C/D box small nucleolar RNA are non-coding RNA (ncRNA) first identified in the achael genus Pyrococcus with the function of modifying ribosomal RNA (rRNA) and transfer RNA (tRNA) [510]. It has since been found to be localised to the nucleolus in eukaryotic cells which is the primary organelle for the biogenesis of rRNA and tRNA.

The sensitizers comprise an approximately similar proportion between snRNAs and microRNA with two and three associations respectively. Small nucleolar RNA SNORA71 and small nucleolar RNA sR11 were found in AK055434 and AK055108 respectively. AK126603 was associated with the microRNA mir-448 while AK055108 contained two sites associated with mir-335 similar to the observation for AK098506 of the inhibitor candidates. AK125960 was associated with the APC internal ribosome entry site (IRES), an RNA element within the coding sequence of the APC gene. APC is a tumour suppressor gene and the IRES-mediated translation of APC is crucial for the apoptosis cascade [511]. Remarkably, the noncoding sequences of the sensitizers also included three terms associated with amino acid metabolism, and one associated with retroviral stability. These candidates reinforced the functional annotation and signalling pathways enriched for the protein-coding sensitizers, where translation- and viral-associated pathways were similarly enriched. Here, AK128619 contained two sites for the Termite-leu RNA structure associated with genes involved in leucine biosynthesis while AK091631 contained ctRNA (counter-transcribed RNA) associated with translational inhibition [512].

The sensitizer candidate AK093650 contained the FAM13A antisense RNA 1 conserved region 2, which is a long noncoding RNA (lncRNA); FAM13A is associated with the positive regulation of GTPase activity and hence involved in regulating small GTPase mediated signal transduction. Its variant is also known to confer a protective effect on lung cancer cells [513]. AK093650 could serve to directly knockdown and silence its target gene FAM13A and influent the small GTPase signal transduction cascade.

Indeed, this potential to knockdown and silence the target gene via the antisense sequences remained a novel aspect of apoptosis regulation. This observation was also found among the candidate noncoding sequences. In the initial attempts to use the annotation information curated by the ENCODE Project [498], ten inducer noncoding sequences were randomly selected and manually searched using UCSC genome browser. One cDNA clone, AK124552, was identified at three positions on chromosome 5 of the human genome (Positions 69403208 – 69406827, 70278675 – 70282294 and 70403560 – 70407180), all of which covered regions associated with NAIP. NAIP, also known as Baculoviral IAP repeat-containing protein 1 (BIRC1), is a known anti-apoptosis protein which acts as inhibitors of caspase-3, -7 and -9 [514, 515]. When translated with ExPASy a protein-coding ORF could be found in the 3’-5’ (reversed) Frame 2, and a blastx query on the NCBI server revealed that this region comprises of 58% coverage and 99% identity match which included two protein domains NACHT and GVQW. The NACHT domain is found in proteins involved in apoptosis [516]. Since this sequence was cloned under the pME18SFL3 expression plasmid, transcription could only occur in the

5’-3’ direction, generating the antisense sequences which potentially could directly bind the mRNA for NAIP, causing the complete knockdown of the target or resulting is alternative splicing and the loss of the region including the apoptosis-associated NACHT domain. This is the most striking example identified among the noncoding sequences, where a noncoding apoptosis inducer sequence could potentially directly influence the expression of an anti-apoptotic protein, relieving the inhibitory function leading to caspase activation. The incorporation of the ENCODE annotation, and identification of more known noncoding RNA sequences using the GenCODE annotation of noncoding RNA [517] is currently underway, and will be separately discussed in a future publication.

The occurrence of microRNA sequences was briefly approached in two ways. First, the frequency for each of the three unique microRNA upstream regulators miR-1 (GGAAUGU), miR-125bp-5p (CCCUGAG) and miR-7-5p (GGAAGAC) identified by IPA was counted to calculate the occurrence probability; miR-1 was common to both inhibitors and sensitizers while miR-125bp-5p and miR-7-5p were only associated with the inhibitors and sensitizers respectively. The occurrence probability was normalised with the natural probability for the length of the microRNA sequence, in this case 7bp and presented as a percentage change as in Equation 13.

Equation 13: Normalised change in occurrence probability (NCOP)

$$NCOP\ (\%) = \left(\frac{Occurrence\ Count_{7bp\ miR\ Seq}}{total\ count\ of\ bp_{candidate\ type}} \div \frac{1}{4^{length\ of\ miR\ Seq}} - 1 \right) \times 100$$

miR-1 with the seed sequence of GGAAUGU was 48.45% and 31.44% more likely to in the noncoding inducer and sensitizer sequences, compared with 24.10% in the inhibitors; likewise miR-125bp-5p showed increased occurrence among the noncoding inducers and sensitizers by 48.45% and 80.74% compared with 40% in the inhibitors. This suggested that there was an increased presence of the seed sequence of both microRNA upstream regulator associated with the inhibitor, together with the fact that miR-125bp-5p was twice as likely to occur in the sensitizers indicated that the noncoding inducers and sensitizers may be enriched for pro-apoptotic precursor microRNA sequences by targeting apoptosis inhibiting targets. The reverse was also true for miR-7-5bp which was an upstream regulator unique to the sensitizers. The occurrence of miR-7-5bp was increased from 19.12% in the sensitizers to 33.65% in the inhibitors, although the occurrence was also increased in the inducers by a slightly lower magnitude to 29.30%.

The second approach involves the use of all known human microRNA sequences downloaded from MiRBase database (2042 matured sequence entries) [518]. The occurrence of the matured

microRNA sequences (full length microRNA sequences usually between 18 – 24bp) was identified among the noncoding sequences of the three candidate types (Figure 49A). The inducers were associated with four unique microRNA sequences hsa-miR-5095 (AK127225), hsa-miR-5096 (AK127110 and AK055262), hsa-miR-1273g-3p (AK092937 and AK024862) and hsa-miR-3654 (AK097279). The inhibitors were associated with hsa-miR-5096 (AK097083, AK055262, AK056407 and AK130488), hsa-miR-1273f (AK022113 and AK026906), hsa-miR-3149 (AK025763) and hsa-miR-1285-5p (AK092621). Intriguingly, a single inhibitor noncoding sequence (AK098506) was contained four unique matured microRNA sequences: hsa-miR-497-3p, hsa-miR-497-5p, hsa-miR-195-3p and hsa-miR-195-5p; this is an incredibly rare occurrence with a probability of 1.67048E-52. AK098506 is hence likely to be a precursor molecule for the generation of the associated microRNAs. The same may be argued for the other noncoding sequences with the matured microRNA sequence, since the occurrence of a nucleotide sequence with an average length of about 20bp is already an extremely rare event occurring once in every 1,099,511,627,776 base pairs. Nonetheless, the potential of these sequences as microRNA precursors may be dependent on other features such as the formation of stem-loop secondary structures to aid in the process of microRNA generation [519]. No microRNA sequences from the miRBase database were associated with the noncoding sensitizers.

The noncoding sequences were also manually submitted to the Regulatory RNA Motifs and Elements Finder (RegRNA) integrated web server for the identification of homologs of regulatory RNA motifs and elements within the noncoding sequences [520]. An interesting observation from the predicted annotations revealed the potential of some noncoding sequences, such as AK021669, which are capable of binding a single microRNA repeatedly across its entire length (Figure 49B). This suggested that these sequences may potentially function as a decoy to sequester their target microRNA and prevent their functions. Such sequences appeared to be fairly common although the precise frequency was not quantified. RegRNA annotations included motifs for the 5'-UTR, 3'UTR exonic regulatory motifs, intronic regulatory motifs, transcriptional regulatory motifs and microRNA target sites, and are too extensive to discuss individually. These annotations are made available in the supplementary materials for reference.

Concisely, information for the unique microRNAs were mined from miRMaid [521] while their potential targets were predicted using TargetScanHuman Release 6.2 [522, 523]. In total, inducers were associated with four microRNA sequences (all miRBase), inhibitors with sixteen (2 IPA upstream regulators, 8 miRBase and 6 Rfam) and the sensitizers with four microRNA sequences (2 IPA upstream regulators and 2 Rfam). TargetScan predicted 741 targets for the four inducer associated microRNAs, 709 of which are unique (4.51% redundancy). For the 16 microRNAs of the

non-coding inhibitors, 5393 targets were predicted with 3204 unique (40.63% redundancy) while the four microRNAs of the sensitizers had 1745 targets with 1585 unique (9.17% redundancy). The average targets per microRNA for the inducers, inhibitors and sensitizers were 247, 599.22 and 581.67 (only microRNAs with available prediction were included for the averaging). This revealed that the target range of the inhibitors and sensitizer microRNAs were approximately 2.43 times that of the inducers. Although the increased range of the microRNA of the inhibitors could be attributed to the high redundancy within the targets, the sensitizer microRNAs appeared to be targeting a wide range of target with little redundancy.

Another interesting aspect was that the target range overlapped with the protein-coding candidates. For example, the inhibitor microRNAs were hypothesised to target and down-regulate the inducer and sensitizer protein-coding candidates to achieve its anti-apoptotic effect. Predicted targets of the inhibitor microRNAs were then compared with the protein-coding inducers and sensitizers, revealing that 100 protein-coding inducers and 86 sensitizers may in fact be targeted by the inhibitor microRNAs (Figure 50). The reverse is also true, albeit with a smaller effect, for the inducer and sensitizer microRNAs targeting the protein-coding inhibitors. Here, 43 and 78 protein-coding inhibitors may be targeted by the inducer and sensitizer microRNAs respectively.

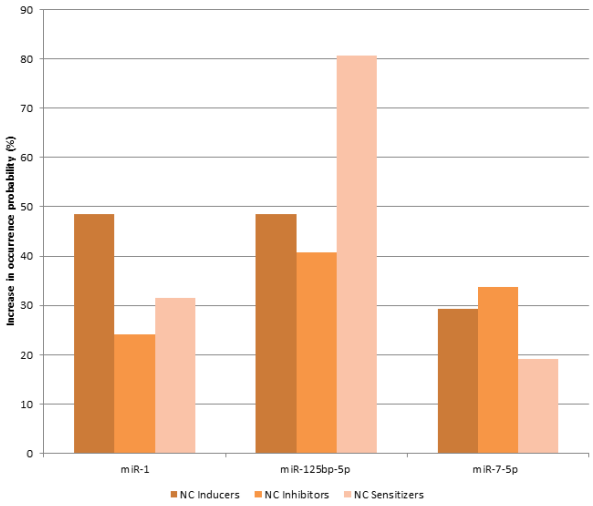
miRMaid also revealed the association of the following microRNAs with cancer: hsa-miR-5095, hsa-miR-5096, hsa-miR-1273g-3p, hsa-miR-1273f, mir-149, miR-7-5p, mir-335 [524], hsa-miR-3654 [525], hsa-miR-3149 [526], mir-16 [527] and mir-650 [528]. However, the functional and regulatory mechanisms of these microRNAs remained to be elucidated experimentally.

The possibility of the noncoding sequences encoding short open reading frames (ORFs) for mini-proteins. These are defined as polypeptides comprising no more than a hundred amino acids and are known to play important roles in a range of cellular processes in both prokaryotic and eukaryotic cells [529]. The 5'-3' ORFs are translated from the cDNA sequences using an Excel macro and the distribution of the mini-proteins analysed.

This distribution appeared to be following its natural pattern where shorter peptides are occurring at greater frequency, which decreases to near zero as the polypeptide length approaches the defined limit of 100 amino acid residues (Figure 51). This pattern remained similar between the three candidate types and is therefore likely to be an indication that the effect may not be significant. Furthermore, attempts were made to align the mini-proteins between 70 – 100 residues in length using ClusterW2. Although there appears to be regions of probably similarity, manual blastp of these overlapping regions between the various mini-proteins did not yield significant products. In addition,

blastp was also performed for approximately 20 of these longer polypeptides (70 – 200 residues); nothing led to known domains or proteins. It is likely that greater depth of analysis is required to look into any potential pattern by analysis of the residue and polypeptide occurrence, and more diverse analysis of the degree of conservation, if any, within these mini-proteins. However, this hypothesis was not developed further within the context of this project.

A) Increase in occurrence probability for the unique IPA enriched microRNA upstream regulators



B) Overview of some microRNA target sites with the AK021669 noncoding sequence

Figure 49: Analysis of the noncoding sequences - microRNAs. A) The occurrence probability of the three unique microRNA upstream regulators identified by IPA was normalised with the natural probability of the length of a nucleotide sequence, in this case 7bp and presented as a percentage change. Here, miR-1 and miR-125bp-5p both inhibitor upstream regulators were shown to be markedly increased among the noncoding inducers and sensitizers; this suggests a regulatory mechanism where these noncoding inducers and sensitizers silence the anti-apoptosis targets to initiate a pro-apoptotic response. The seed sequence for miR-7-5p, a unique upstream regulatory associated with the sensitizers, was also observed to occur with an increased probability among the noncoding inhibitors. B) An example showing some of the microRNA target sites predicted by RegRNA for AK021669. The sequence appeared to be able to bind miR-1 across its entire length, suggesting a decoy mechanism.

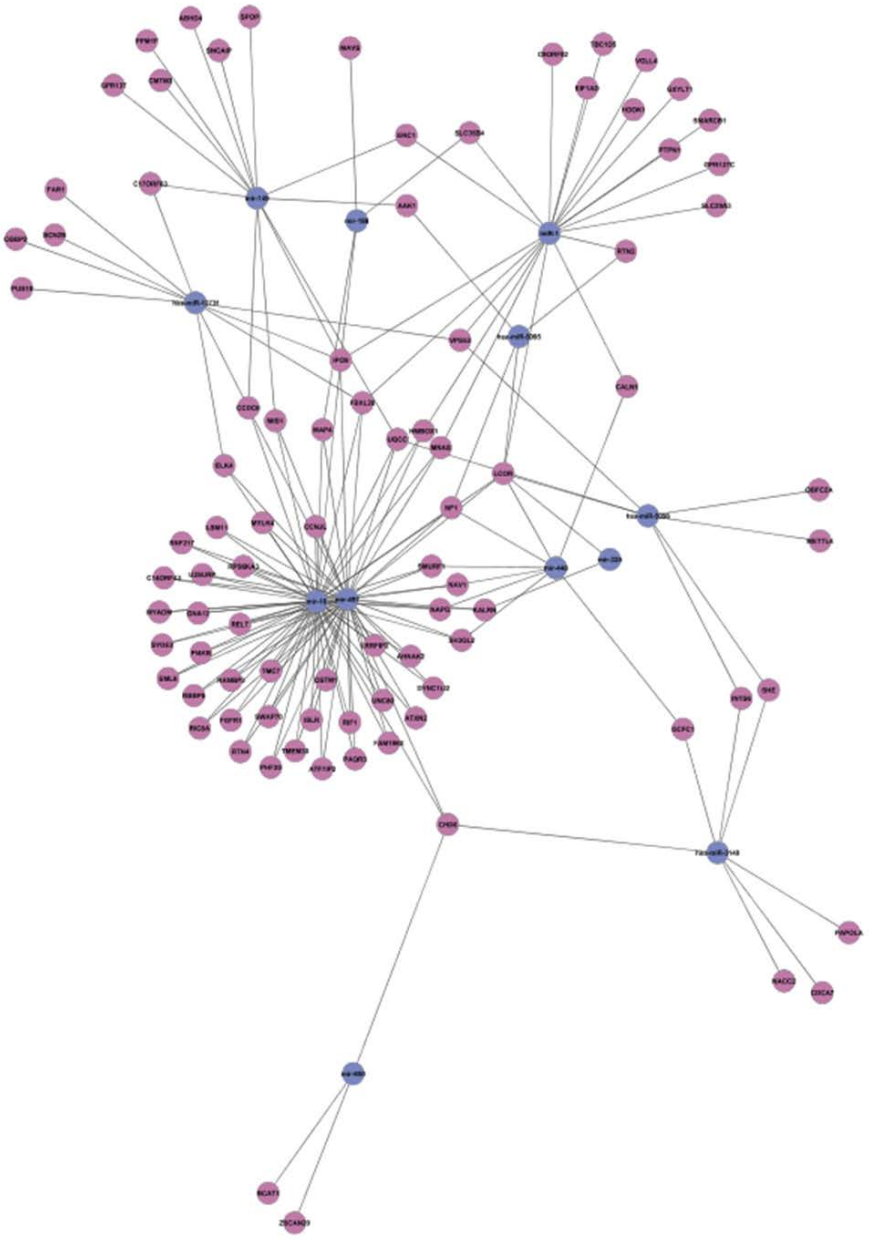


Figure 50: The protein coding inducer candidates are targets of the noncoding inhibitors candidates. Some of the noncoding candidates are potentially the precursor of endogenous microRNAs. TargetScan predicted the potential targets which could be silenced by these microRNAs, revealing here that the protein coding apoptosis inducer candidates (purple) can be knockdown by microRNAs generated from the noncoding inhibitor candidates (blue). The reverse is also true. The large number of interactions suggests an intricate apoptosis signalling network featuring a high degree of cross-talk.

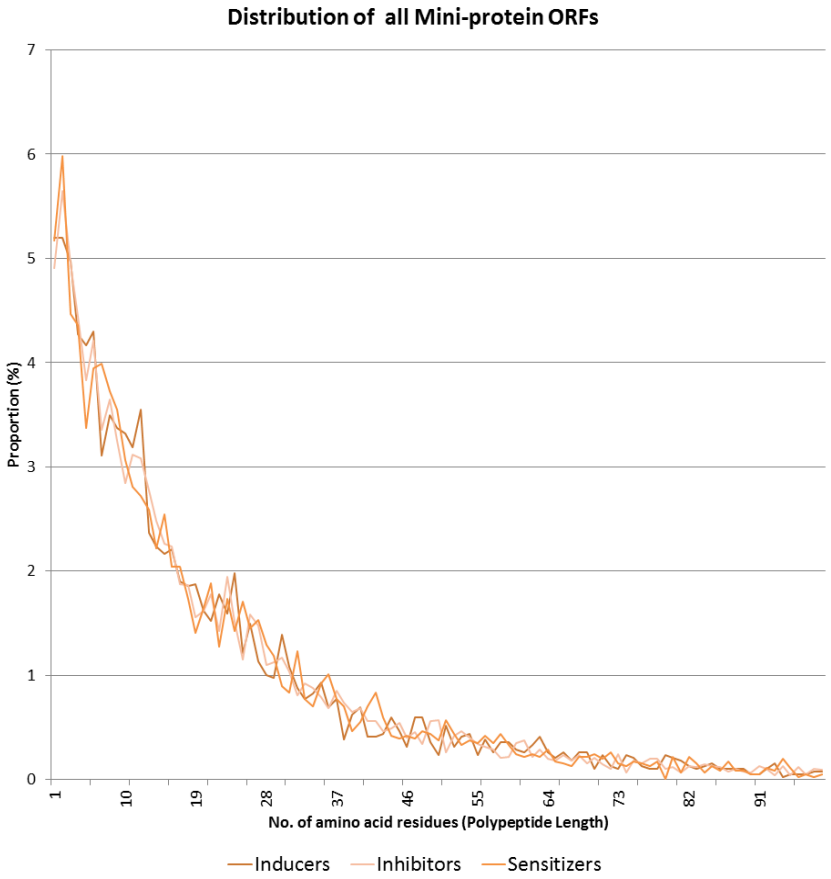


Figure 51: Distribution of all mini-proteins. All ORFs encoding potential polypeptides were translated. The distribution of the potential mini-proteins appear to be following its natural distribution, where the short peptides had increased frequencies which reduces to near zero as it polypeptide length approaches the defined limit of the mini-proteins. This pattern remained similar between the three candidate types, and is likely to indicate that mini-proteins might not be significant in this context.

Discussions

Functional annotation by DAVID generally resulted in distinct characteristics for each of the three candidate types. The enrichment for a high count of genes with known apoptosis activity in both the inducer list was a striking validation of the RISCI screening approach. The inducer and sensitizer candidates with the pro-apoptotic behaviour also appear to involve protein products and the protein translational machinery to a greater degree. This was in contrast to the inhibitors, which appeared to regulate apoptosis primarily via transcriptional mechanisms such as the processing and splicing of RNA.

Furthermore, the inhibitors were closely associated with cancer expression profiles as determined by GSEA; it also has the highest count of genes expressed in cancer-associated tissue profiles. This association with cancer profiles was crucial towards the validation of inhibitors since cancer cells were known to evade apoptosis. It was also interesting to note that the genes which were capable of inhibiting the cell death induced by the LUDOX® silica nanoparticles were in fact similar to those employed that cancer cells. Nonetheless, the enrichment of the set of genes associated with metal ion response offered important insights into the signalling mechanisms potentially employed by normal cells. The response to metal ion activated by the resulting increase in reactive oxygen species (ROS) which is also likely to be implicated in the silica nanoparticles induced cell death.

The pathway analysis also revealed interesting findings. The association of the PI3K, MAPK and Rho signalling cascades all of which could potentially contribute to apoptosis were identified among the inducers. The inhibitions, on the contrary, were involved in Alzheimer disease-amyloid secretase pathway and Huntington's disease signalling pathway both of which are neurodegenerative diseases. The sensitizers were associated with pathways involving the translational machinery and interesting appeared to be associated with various viral terms, which was present even in the noncoding sequences.

While there is a general pattern which fits with the expected profile of the candidates, there remains a healthy level of disagreements. A recurring theme is the observation of that some pro-apoptotic candidates turned out to possess known anti-apoptotic activity and vice versa. These candidates may possess additional roles within apoptosis signalling network which are yet to be discovered. In the end, the association of these known pathways would help to support the general pattern of the candidate lists. But the interesting output of the RISCI screen will be the novel candidates implicated for the first time in regulation of apoptosis. Little is known about these novel candidates, but

together with the integrated analysis they would help to drive new hypothesis and research for these novel candidates.

One of the most surprising outcome of the screen and analysis was the observation that many of the noncoding sequences which did not result in known protein products are in fact functional and equally capable of regulating apoptosis. The ENCODE project, only published recently, served to provide evidence in support of our hypothesis that these noncoding candidates may exert their effect on apoptosis signalling via previously unknown mechanisms.

The analysis of these noncoding sequences revealed that these are likely to involve a range of different regulatory mechanisms. As long noncoding RNA (lncRNA), the noncoding candidates may form complex structures to guide the recruitment of chromatin modifying complex, or facilitate the interactions between two RNA binding protein complexes. Alternatively, they could bind transcription factors to suppress the expression of the target promoter, or directly bind the microRNA sequences to prevent the silencing of the target genes [530]. There are indications from the analysis in support of some of these mechanisms, especially the sequestering of the microRNA since many of the candidates were capable of binding the same microRNA molecules across the entire length (see RegRNA analysis).

The potential to serve as precursors to the microRNA, thus directly involved in the regulation of apoptosis by the knock-down of their counterpart exerting the opposite effect, is also another mode of action. While it was certainly of interest that many of the matured microRNA sequences could be wholly detected among the noncoding candidates in support of the microRNA precursor hypothesis, the single most interesting observation from the noncoding analysis was the reciprocal behaviour of the pro-apoptotic inducer and sensitizer microRNAs have on the protein-coding inhibitor candidates and likewise the inhibitor microRNAs are capable of targeting the protein-coding inducers and sensitizers.

Till date, few of the lncRNA and other noncoding elements were determined experimentally, and none are known to exert a known regulatory function in apoptosis signalling. The identification of 598 noncoding candidates associated with the various aspects of apoptosis signalling could open up a new field in the control of apoptosis regulation through these noncoding elements.

Although candidates from the three classes were subjected to repetitive validation and shown to be statistically significant here, and the extensive analysis based on currently available knowledgebase was in agreement with the predicted characteristics, further experimental validation will be required to firmly elucidate the signalling pathways involved and confirm the apoptosis activity.

Chapter 9: Assay Design

Background Information

In parallel with the optimisation and setup of the RISCI screen for apoptosis inducers, inhibitors and sensitizers, brief attempts were made to establish a suitable assay for the detection of autophagy regulators.

Autophagy is the regulated catabolic breakdown of excess or damaged cellular components and biomolecules, serving as the primary mechanism for the recycling of materials within the cellular system. It is also an important mechanism for cell survival under nutrient starvation, and is also employed for removal of foreign particles such as viruses or bacteria [531]. Excessive autophagy has also become progressively recognised as another form of programmed cell death alongside apoptosis. Intricate crosstalk exists between apoptosis and autophagy signalling. Its signalling network is frequently implicated in various cancer types [532, 533] and in degenerative diseases such as Alzheimer's [534].

Developing an assay for the detection of autophagy regulators using the RISCI screening platform and NITE library could offer new insights into the signalling mechanism of autophagy, and how it interacts with apoptosis regulation.

Microtubule-associated proteins 1A/1B light chain 3B (LC3) is a prominent marker and effector for autophagy. The cytoplasmic protein (LC3-I) is cleaved at the C-terminal during autophagy and translocates to the autophagosome (LC3-II). Fluorescently tagged LC3 is routinely used as an autophagy marker, where the activation of autophagy results in the concentration of fluorescence signal to the autophagosome, which appear as bright green dots. It was hypothesised that by using the C-terminal amino acid sequence in conjunction with a reporter system, the cleavage which is normally observed through microscopy could be adapted into a quantifiable signal suitable for high-throughput format.

Results

Optimisation of Autophagy Induction

The autophagy reporter plasmid pEGFP-LC3 was obtained from Dr Anne Petiot, Université Joseph Fourier, Grenoble. Transfection optimisation was performed for a variety of cell lines including HEK293T, MCF-7, HeLa, HFF-1, A549 and BHK cells. These cell lines were transfected using the optimised calcium phosphate protocol or commercially available transfection kits including QIAGEN Effectene, Attractene, Superfect, Clontech XFect and Polyplus jetPEI according to manufacturer's protocol.

HEK293T cells were easily transfected but the small stature of the cells increased the cytoplasmic fluorescence of the GFP-LC3, which had the effect of masking the localised signal. HFF-1, A549 and BHK were rejected due to inconsistent and low transfection efficiency. HeLa cells were eventually found to be suitable for the detection of autophagy using fluorescence microscopy due to its large cell area. When transfected using jetPEI with up to 0.5ug of active pEGFP-LC3 plasmids and 80,000 HeLa cells, the overexpression of the GFP-LC3 did not result in observable cleavage and localisation of the reporter into the autophagosome.

The various cell lines were also subjected to conditions known to induce autophagy. Work primarily revolve around various type of nutrient starvation, including overnight serum deprivation, culturing of the cells in the amino acid free media Earle's Balanced Salt Solution (EBSS, Invitrogen and Sigma Aldrich), serial dilution of serum-free DMEM with PBS, and even in PBS alone. In addition, ER stressors such as brefeldin-A and thapsigargin, and chemical compound inducers like rapamycin and lithium chloride (10 – 20mM) which inhibits downstream signalling from the target of rapamycin (TOR) proteins were also tested. Incubation in EBSS (Sigma Aldrich) for approximately 90 minutes 16 – 20 hours post-transfection of the GFP-LC3 reporter was shown to induce autophagy above background, quantified by manual counting of nine image frames and more than 300 cells per treatment (Figure 52). However, none of the conditions tested could consistently activate autophagy as assessed by the transient transfection of GFP-LC3.

Lentivirus was used to create a stable cell line of HeLa cells expressing the GFP-LC3 reporter to avoid transient overexpression using the commercial jetPEI kit. The GFP cassette in the pLenti7.3 viral vector was first replaced with a puromycin cassette (termed pLenti7.3puro) to resolve the interfering signal of the GFP and allow for selection of stable clones using the aminonucleoside antibiotic puromycin. eGFP-LC3 DNA sequence was then PCR amplified and cloned into the pLenti7.3puro vector. After infection and puromycin selection, the cells were kept at high dilutions

per well of a 96-well plate to isolate single clones. Subsequently, it was discovered during the expansion of these single clones which occur over a period of two weeks that the stabilised glutamine in the DMEM AQmedia^{TM} (Sigma Aldrich) had an inhibitory effect on the autophagy activation. Indeed, a literature search identified a recent publication which further confirmed that glutamine was a signalling molecule in autophagy [535].

The ammonium (-NH_4) ions appear to play an important role in the activation of autophagy, and a range of ammonium salts were used tested for this effect. 25mM, 50mM, 100mM and 200mM of ammonium acetate, ammonium chloride, ammonium formate and ammonium hydroxide. Ammonium hydroxide was found to be extremely toxic and excluded. Although ammonium acetate and ammonium formate were shown to be able to induce the localisation of GFP-LC3 to the autophagosomes, they displayed toxicity at the minimum concentration of 25mM tested.

It was found that ammonium chloride could readily and consistently activate autophagy, and at 25mM causes the GFP-LC3 to localise to the autophagosome without induction of cell death (Figure 53). The phenotype was assayed using live confocal imaging of the HeLa cells stably expressing GFP-LC3 and western blotting which further confirm the cleavage of endogenous LC3 (Figure 54A). High concentration of 100mM ammonium chloride was toxic and the cells displayed a unique phenotype of being rounded with huge vacuoles visible (Figure 54B). These rounded cells also appeared to remain stable without progression into apoptosis, which some of the surrounding cells did. Apoptotic cells displayed strong and uniform fluorescence signal, with membrane blebbing often observable.

Treatment using 25mM ammonium chloride for two hours was selected as the condition for inducing the autophagy phenotype, although live imaging of the HeLa cells stably expressing GFP-LC3 indicated that localisation of LC3 to the autophagosomes occurred within 60 minutes of treatment (Figure 54C). This condition was also validated in a range of other cell lines such as HEK293T and N2a which is a cell line of neuronal origins to induce cleavage of LC3.

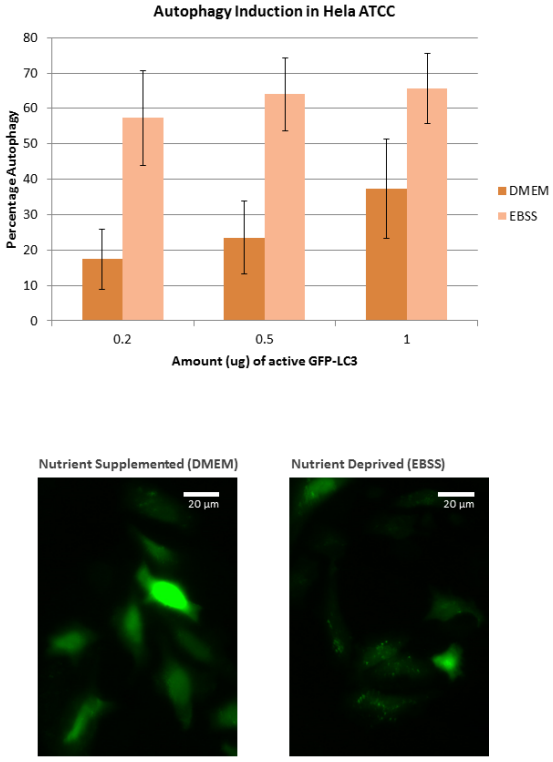


Figure 52: Autophagy induction using nutrient depleted EBSS media. Autophagy in HeLa cells could be observed after 90 minutes incubation in EBSS media 16 – 20 hours post transfection of the GFP-LC3 reporter. The number of cells undergoing autophagy was manually counted against the total population count. In total, nine image frames were captured with more than 300 cells counted per condition. N = 3, Magnification = 400X, scale bar = 20 µm, error bars represent standard deviation of sample size.

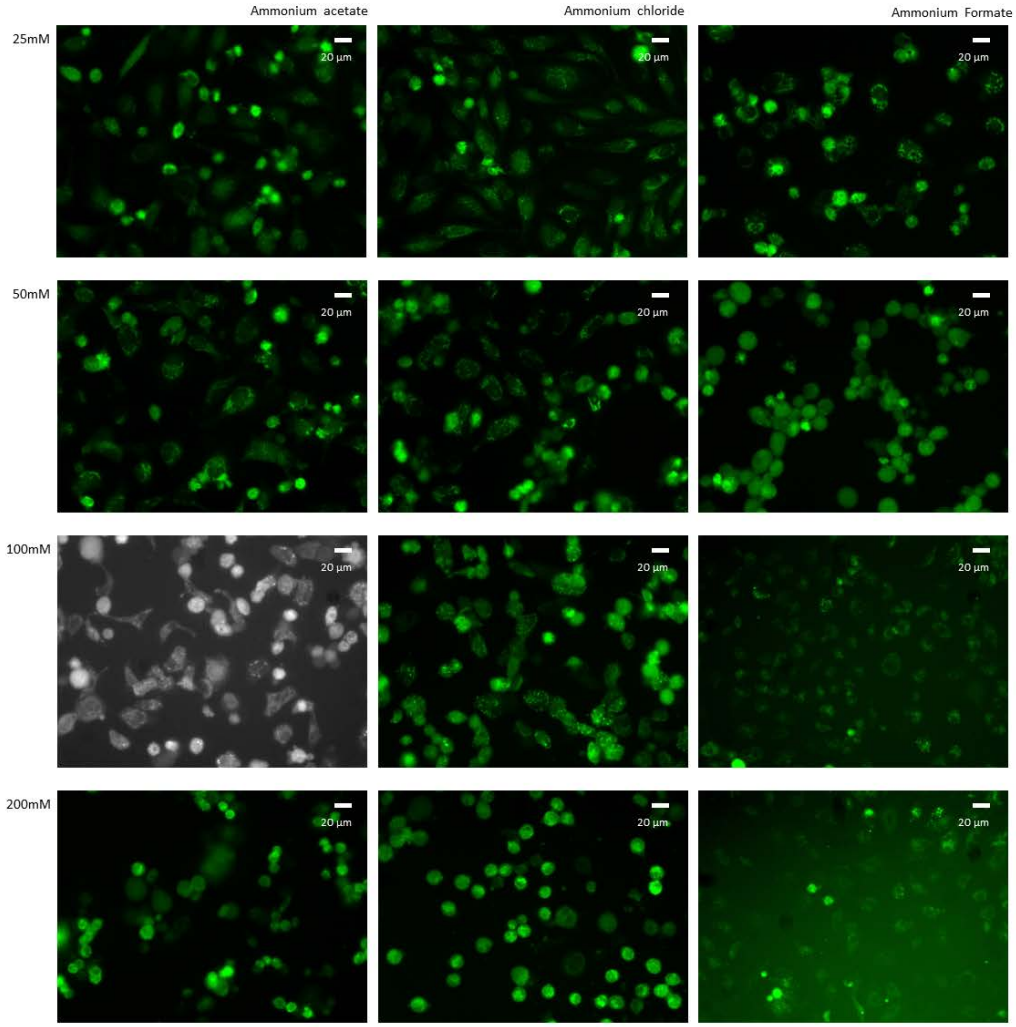


Figure 53: Autophagy induction using ammonium salts. HeLa cells stably expressing the GFP-LC3 reporter were used to screen for the effect of ammonium salts on activation of autophagy. 25mM, 50mM, 100mM and 200mM of each ammonium salts were applied to the cells for 24 hours. Ammonium acetate and ammonium formate were able to induce autophagy but also increased toxicity. 25mM ammonium chloride was found to consistently activate autophagy with little or no cell death observed. Magnification = 200X, scale bar = 20 µm.

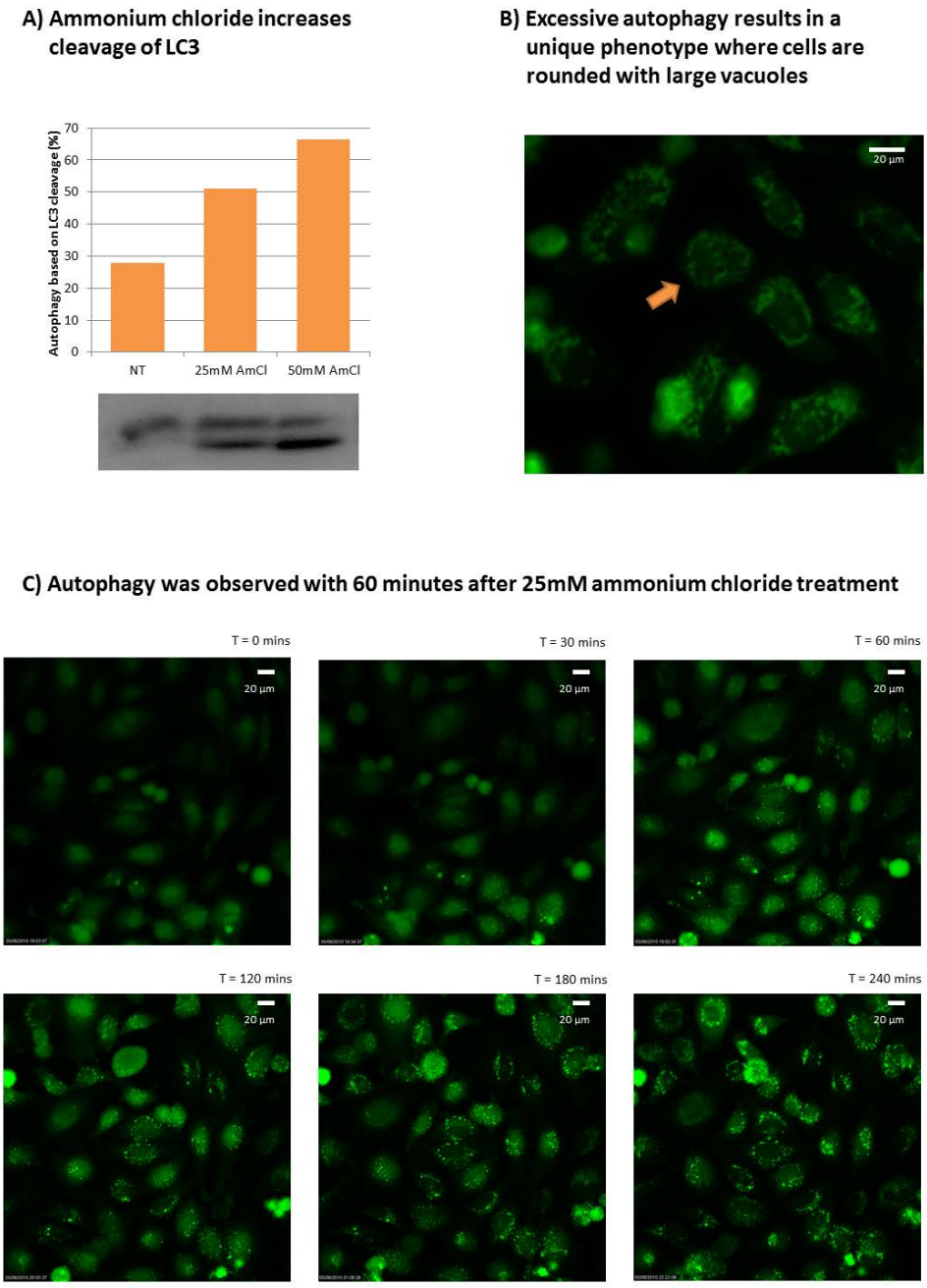


Figure 54: Ammonium chloride was able to induce autophagy. A) Induction of autophagy by ammonium chloride was observed on western blot. Here, endogenous LC3 was observed to be cleaved and phosphorylated in HEK293T cells with increasing concentration of ammonium chloride. B) Image of HEK293T treated with 50mM ammonium chloride for two hours. Vast increase of the autophagosome count and intensity can be observed. Some cells were also observed to assume a unique phenotype of being rounded with large vacuoles. Magnification = 400X, scale bar = 20 µm. C) Time-course confocal microscopy showing the response of HEK293T to 25mM ammonium chloride. Autophagy indicated by the autophagosome formation may be observed after two hours. Magnification = 200X, scale bar = 20 µm.

Cloning of the FRET Reporter System

It was previously published that the cytoplasmic LC3 protein was cleaved at Met121 at the C-terminus by Atg4b [536, 537] to expose Gly120 for lipidation and targeting. Amino acid sequences consisting of residues 111 - 125 of MAP1LC3B (MVYASQETFGMKLSV) was cloned to test for selectivity of the published cleavage site. The human codon optimised linker (SGLRSGGDEVDGGSNS) caspase-3 substrate recognition sequence DEVD was cloned in parallel to act as a cleavage control during the validation of the reporter system.

Two strategies exist to quantify the cleavage of LC3. The first involves the use of fluorescence resonance energy transfer (FRET) between two chromophores linked by the LC3 or DEVD cleavage sequences. The FRET effect can be readily quantified when the two chromophores remain in close proximity imposed by the cleavage sequence. The activation of autophagy or apoptosis would result in the respective cleavage, leading to a loss of FRET signal. This loss of FRET signal may be quantified, although a negative readout may not be an ideal assay. The second strategy exploits the alpha complementary effect of the β-galactosidase enzyme in a similar manner to the blue white genetic screen in *E. coli*. The functionally inactivated mutant with a deletion to amino acid 11 – 41 (the omega fragment) can be reactivated in its catalytic activity by the peptide with amino acid 3 – 90 of the β-galactosidase enzyme (the alpha fragment) [538]. The alpha complementation effect was also previously demonstrated in a mammalian system [539]. It was hypothesised that the fusion of the alpha fragment with the cleavage sequence would disrupt the interaction between the two β-galactosidase fragments and reduces the catalytic activity. As with the FRET system, activation of the target process results in cleavage which releases the alpha fragment to restore the catalytic function of β-galactosidase. As opposed to the FRET system, this assay strategy generates a positive readout which correlates with the phenotype-of-interest.

tCFP and YFP were individually cloned from a tCFP-Calmodulin-YFP FRET construct obtained from Ryota Iwasawa, a PhD student from our group, into the vector pcDNA3.1. An ATG start codon was added to the tCFP at the 5' end for translational initiation while the TAA stop condon was cloned such that it was flanked by the NheI and NotI restriction sites. This construct was labelled as pcDNA-tCFP(NheI)TAA(NotI), which allowed for normal protein translation, and when required the YFP protein can be cloned in frame using the NheI site to create a FRET reporter protein. An NheI restriction site was added before the start codon of the YFP during cloning into pcDNA3.1, and is referred to as pcDNA3-(NheI)YFP. A primer pair was first used to PCR amplify the approximately 700bp generated by either the YFP or CFP sequences from the calmodulin FRET construct. These sequences are then cloned using EcoRI and NotI into pcDNA3.1, with the individual clones

transfected into HEK293T cells and visually screened using fluorescence microscopy. The selected clones were further sequenced to confirm that key mutations, Y66W for CFP and T203Y for YFP, were present before the additional NheI was incorporated using PCR and a repeated cloning step back into pcDNA3.1. To create the LC3 or caspase-3 FRET reporters, primer sequences containing the linker sequences were used to PCR amplify CFP and the PCR amplified sequences cloned into the pcDNA3-(NheI)YFP. The cloning of the CFP and YFP FRET system was performed using standard restriction enzyme digests and T4 ligase methods.

Cloning of the alpha and omega fragments was performed using the In-Fusion™ Advantage PCR Cloning Kit (Clontech) according to the manufacturer's protocol.

Validation of the LC3 and Caspase-3 FRET Reporters

Under the CPF and YFP FRET system, the CFP acts as the donor and the YFP as the acceptor. When the donor and acceptor are far apart, no FRET can occur; hence the excitation of CFP at wavelength 436nm would result in the regular CFP emission at 480nm. However, when both donor and acceptor are in close proximity, the CFP emission at 480nm serves as the excitation wavelength for YFP, resulting in the YFP emission at 535nm. This is known as the FRET effect and by exciting the FRET reporters at 436nm, and measuring the emissions at 480nm and 535nm, the effect can be quantified as the FRET ratio using Equation 14:

Equation 14: FRET Ratio

$$FRET\ Ratio = \frac{YFP\ emssion\ at\ 535nm}{CFP\ emission\ at\ 480nm}$$

Where the FRET effect signal indicated by the YFP emission is normalised by the number of reporter present indicated by the CFP emission signal.

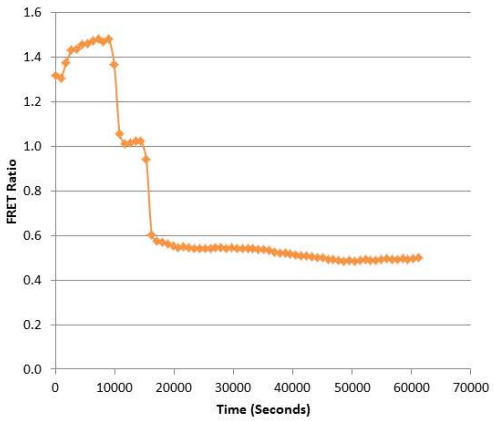
10uM arsenic trioxide was used to induce apoptosis in HeLa cells transfected with the caspase-3 FRET construct and the fluorescence signals were captured every 15 minutes using confocal microscopy (Figure 55). The data was analysed using the Leica LAS AF Lite software. A region of interest (ROI) comprising of two cells was analysed. Prior to apoptosis within the first hours of treatment, the FRET signal strength in the ROI continues to increase as the transiently expressed reporter construct continued to accumulate to a maximum FRET signal of 1.48. 3 hours after treatment, a massive decrease in the FRET signal to 1.05 was observed in one cell, and this was followed by the other cell 1.5 hours later. The loss of FRET signal as a result of caspase-3 activation for each of the cells was 0.425 and 0.420 respectively, which indicated similar expression levels of the FRET construct.

Approximately 2.5 hours after the activation of capase-3, the first signs of membrane blebbing could be observed and massive membrane blebbing continues until the cell finally becomes rounded 10 hours after activation of caspase-3 (See Supplementary Materials).

At the time of testing of the caspase-3 FRET construct, the autophagy induction conditions was yet to be finalised. Therefore, no confocal microscopy data was available. Furthermore, initial western blot analysis suggested that the LC3 FRET constructs was being constitutively cleaved by the cells during transient overexpression (data not shown).

Due to time constraints to initiate the RISCI screen, the LC3 FRET constructs and the alpha and omega fragments complementation were not pursued further.

A) FRET ratio in the region of interest



B) Live confocal imaging of the detection of caspase-3 activation using the FRET reporter

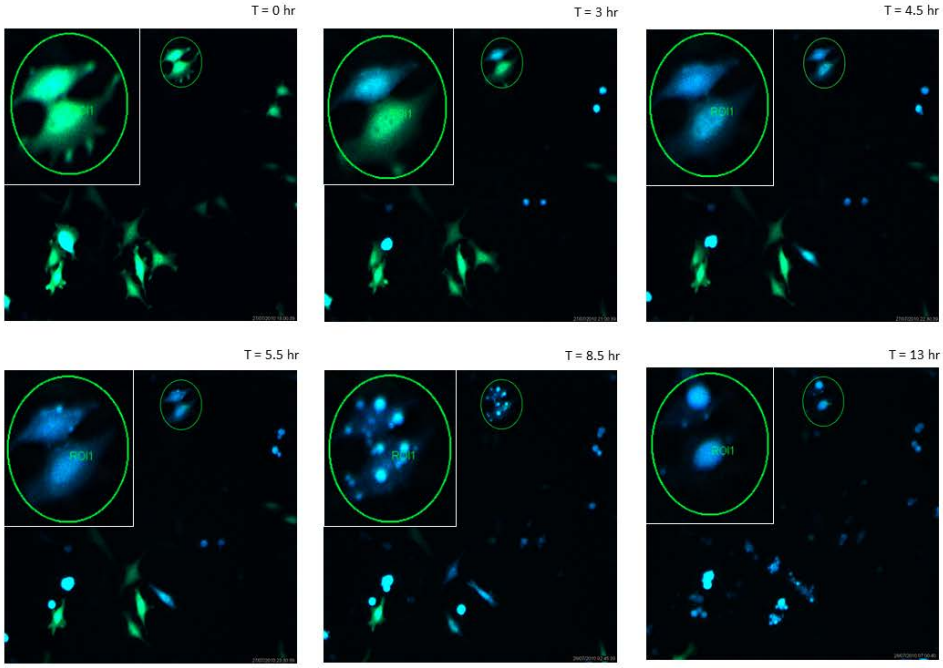


Figure 55: Detection of caspase-3 activation in HeLa cells using the FRET reporter system. A) The calculated FRET ratio within the selected region of interest (ROI). The two cells showed decrease in FRET ratios of 0.42 within the same time period of approximately 50 minutes and follow a similar pattern, suggesting similar level of expression for the reporter protein and extremely high sensitivity of the confocal microscope. B) Time lapse images of the process of apoptosis activation detected by the FRET construct, with an enlarged image of the ROI inset. Co-localisation of the blue and citrine spectrum of the CPF and YFP (FRET) signal resulted in a final green hue. The cleavage of the reporter protein resulted in loss of the FRET signal, returning the fluorescence signal to the blue spectrum. Common apoptosis phenotypes such as membrane blebbing and rounded cell structure continued to be tracked by the labelled cells.

Discussion

Venturing into the realm of autophagy signalling was complimentary to our current speciality research area of apoptosis, since excessive autophagy was acknowledged to lead to an alternative form of programmed cell death.

Transfection was optimised such that overexpression of the GFP-LC3 reporter did not lead to autophagy activation, although it was noted that the presence of GFP-LC3 at high quantities (such as when more than 0.5ug of active reporter plasmids was used) would lead to cellular toxicity within 24 hours. jetPEI (Polyplus) was eventually selected as the optimal transfection reagent.

While this research area was new, the difficulties encountered in trying to elicit an autophagic response was definitely not expected. Well established conditions that were tested include various forms of nutrient and serum deprivation, as well as treatment with chemical compounds were not conclusive due to the presence of labelled autophagosomes in both control and treatment conditions. However, amino acid deprivation using EBSS (Sigma Aldrich) was able to induce a moderate autophagy response.

The acquisition of a Lentivirus system (Invitrogen) also made it possible for the first time to establish stable cell lines with viral particles. The generation of HeLa cells with stable expression of the GFP-LC3 reporter sped up the screening process for identification of autophagy inducing conditions, since transfection steps could be bypassed. The stable cell line was also more sensitive and consistent, with a generally even and standard distribution of the reporter across the cells, which usually differs from between batches of transfection. The generation of the stable cell lines also proved to be a turning as the long incubation during the expansion of the clones led to the depletion of the stabilised glutamine dipeptide alanyl-glutamine (Ala-Gln or AQ) in the AQMedia^{TM} (Sigma Aldrich) used for routine cell culture.

This led to the independent discovery of the importance of glutamine in autophagy signalling, and based on the findings of Christina *et al.* [535] characterisation of an extended range of ammonium salts which could readily induce autophagy by increasing the availability of the ammonium ions. The key to autophagy induction was the increased concentration of ammonium anions, while the cationic counterpart of the salts may play a role in determining the overall toxicity of the treatment. Furthermore, glutamine was also found to induce a similar phenotype when present in twice the normal culture concentration. This implicated the metabolic cycles such as tricarboxylic acid (TCA) cycle in mediating autophagy and vindicates the involvement of glutamine in the process [540, 541].

While cancer cell types were known to actively exploit the TCA cycle metabolism, the involvement of ammonium ions and its sources such as glutamine appears to be a general effect since these molecules were also able to induce the autophagy phenotype in a range of cell lines including the N2a cell line.

Development of the detection assays for autophagy proceeded in parallel with the optimisation of the induction conditions. The aim was to convert the labour-intensive microscopy-based assay of GFP-LC3 into a high-throughput format suitable for use with plate readers. Two approaches were designed to exploit the cleavage of LC3 by ATG4b, a FRET fluorescence based method and another enzymatic assay based on the alpha complementation of β-galactosidase. The use of LC3 cleavage as a quantifiable method of autophagy detection was recently demonstrated at the time, where the LC3 was fused between actin molecules and gaussia luciferase. This fusion prevented the unique luciferase from being secreted into the media environment, and luciferase activity can only be detected during autophagy when LC3 was cleaved [542, 543]. However, it would be difficult to implement this assay on the RISCI platform since a new substrate, luciferin, has to be acquired in substantial quantities on top of the CPRG substrate currently used. The preference was for an assay which either use our current enzyme-substrate selection, or one which does not require the use of additional substrates to minimise the running costs of the screens. Caspase-3 DEVD, a proven sequence for detecting the activation of caspase-3 during apoptosis was used as the positive control for reporter activity.

The FRET based reporter could greatly minimise the costs of a screen by removing the need for substrates, a major cost component. However, as with other fluorescence based methods, the signal generated is proportional to the number of reporter molecules, and hence may compromise on the sensitivity of the assay. The cloning of the FRET reporter system was straightforward, with the CFP and YFP individually cloned with an additional NheI site at the carboxyl and amino terminals respectively. From this pair of plasmids, any linker containing the desired cleavage sequences may be cloned to create its appropriate FRET construct. Here, linkers containing either the LC3 or caspase-3 cleavage sequence were employed, by first using PCR to fuse the linkers to CFP or YFP, then cloning these sequences into the other plasmid. Both LC3 and caspase-3 FRET constructs were verified with restriction digestion, DNA sequencing and fluorescence microscopy.

Decrease in the FRET signal upon the activation of caspase-3 during apoptosis was confirmed for the caspase-3 FRET construct using confocal microscopy. However, the detection of the FRET signal was difficult to achieve using plate readers due to the lower sensitivity. The LC3 construct was not tested for its cleavage upon autophagy induction as the conditions have yet to be finalised, but the

construct was subjected to preliminary western blot analysis by probing with anti-GFP-HRP antibodies (data not presented) which indicated that the fusion protein was constitutively cleaved upon transient overexpression. This, together with on-going literature debates on the viability of LC3 as a quantitative marker of autophagy due to this constitutive cleavage, conflicted with the viability of the LC3 FRET construct. Furthermore, p62 was increasingly being recognised as a quantifiable alternative when monitoring autophagy; p62 (also known as sequestosome 1 (SQSTM1)) was degraded during apoptosis instead of merely localisation to another organelle [544] and hence offer a proportional signal change correlating with autophagy.

The alpha complementation of β-galactosidase was the preferred reporter system, but the cloning of the alpha and omega fragments was hindered by technical issues with PCR and ligation. Eventually, both fragments were cloned using the In-Fusion cloning kit (Clontech) which bypassed the intermediary cloning procedures to proceed directly from PCR product to cloning into vector.

Due to the urgency to implement the apoptosis screen, efforts were redirected towards the direction of implementing and executing this screen. As a result, the LC3 FRET reporter was not validated using confocal microscopy. In additional, both linkers were not added to the alpha fragment, and the alpha complementation effect was not validated in *in vitro* cellular systems.

While the DEVD FRET construct was demonstrated to be functional, the assay for autophagy detection using the LC3 marker was not positive. Furthermore, LC3 may be an excellent phenotypic marker but p62, whose levels changes with autophagy, might offer a better alternative for quantitative methods. This aspect will be independently investigated by a separate project.

Chapter 10: Conclusion

Experience

The RISCI screen was an unique opportunity to explore the interdisciplinary nature of high-throughput technologies, from its engineering and computing aspects which resulted in the robotics and instrumentation of this work that enabled the automation of laboratory procedures, to the intricacy of molecular cell biology experiments and finally, how the eventually generated data came together under the domains of bioinformatics, mathematics and statistics to piece together the associated signalling networks.

A significant amount of time was dedicated to setting up of the screen and troubleshooting, and the initial aim was to get the optimal conditions for each contributing section of the screens such as the bacterial culture, DNA isolation, transfection and assay. This assumption that the best condition for each phase would maximise the final screen output was limited, as the best conditions often conflicted with each other thereby requiring compromises. At the start of the project, the limiting factor for the screen output was the DNA isolation phase, where the daily throughput was eight 96-microplates of plasmid DNA had created spare capacity in downstream processes such as automated transfection. Effort were directed at this phase leading to increased yields and doubling of the daily output to sixteen 96-microplates routinely generated during the screen, with up to twenty-four microplates easily manageable. This increase in output in the DNA isolation phase then made full use of the capacity of the transfection platform, which continuously processed the plates through transfection, treatment, assay and media change protocols during a 10 – 12 hours period. However, the critical workflow bottleneck at this point of the optimisation was shifted further downstream to the assay phase, where the microplates had to be manually processed due to the lack of automation. Under the optimal conditions, the BASY platform could generate twenty-four plates of plasmids but this required an increase in the downstream capacity of the transfection platform and plate reader to be efficient. Hence maximising the potential of one phase simply shifted the limiting step to another point along the entire workflow; the optimal screen setting was determined by the compromises and practicality required to integrate all phases towards equilibrium.

These compromises ultimately drove down to a single major factor, costs. Performing a standard experiment using commercially available kits was already an exorbitant proposition, and attention had to be taken to manage the operating costs of a screen when replicating similar setups on a high-

throughput scale. To put a commercial perspective on the screen, the competent bacterial cells from Invitrogen for the high-throughput transformations would be in the region of £215,000 while QIAGEN 96-well DNA extraction kits would require £136,250 for a single round of screen. The costs of Polyplus jetPEI would be approximately £24,000. These exclude the RNase, CPRG substrate, media and plastic consumables which are fixed costs that cannot be avoided. Under the previously discussed protocols, transformation and DNA isolation was achieved with a budget of £500 each attributable to the chemical reagent costs, while the transfection reagents using either the linear or branched PEI polymers were under £50. The primary screen was estimated to be completed with each of the 30,000 samples costing under £0.25 including consumables. Furthermore, the various key components such as the RNase in the DNA isolation or the CPRG substrate used in the assay were optimised to a cost-efficient level; for example, the RNase was reduced by approximately 20% without affecting the plasmid purity and this translated to a £400 reduction in costs. Equipment maintenance costs were also of a major concern over the course of the project as these are often contingencies not factored into the resources.

The practicality of managing the intensive screen also influenced the output since the process was managed by a single user. The user independent aspect of the transfection platform was exploited and served as the maximal point of reference for workflow optimisation. Experiments were setup and the platform was able to independently process the samples. This frees the user to process and track the DNA isolation protocol on the BASY platform in parallel. This results in sixteen plates of plasmids and thirty two plates of transfections to be performed, moving the limiting step to the assay phase. Any incubation times were assigned for reagent preparations or other preparation for the next experiments. In the event where the workflow became overburdened due to unforeseen circumstances, the steps prior to the overloaded phase were scaled appropriately to allow the bottleneck to be cleared. Scheduling of the consumables and reagents was also managed such that new stocks arrive as the old ones become depleted due to the lack of storage space.

Equipment faults were a major point, which required frequent troubleshooting. Perhaps as a result of intensive use from previous screens, both robotic platforms were prone to malfunctions. The DNA isolation platform was frequently incapacitated in the early trial runs with a unique fault each time. These ranged from minor failures of the electric switches on the robots, to more major issues such as a damaged wire connection between the robot and the computer, PC failures and eventually the breakdown of the robotic centrifuge. Minor issues such as the repairs of broken connection could easily be resolved, but more often the expertise of Dr Volker Kachel and Imperial College engineers to identify and repair the problems was essential. Some equipment such as the cabinet incubator

and plate reader were also made available for communal use which introduced uncertainties into the workflow. In one example, the shaking incubator was put out of use when inappropriately secured or damaged plastic wares employed by users led to solution leakages which damaged the system circuitry. This delayed the screen by almost eight weeks as the circuitry board was no longer commercially available and had to be manually repaired, with the costs totalling over thousands of pounds. In another incident, a user had altered the filter positions within the plate reader without updating the corresponding software, invalidating the works of others since all data generated were of the incorrect wavelengths. The screen was also operating at its peak routine, and this unintended external source of human error meant that the cumulative data generated over four weeks had to be discarded. Both the internal source of error resulting from wear and tear and externally introduced errors constantly impeded on the performance of the screen. Standard aging of the equipment is inevitably but could be minimised if supported by specialised engineers and routine maintenance. Dedicated equipment as employed in commercial screening facilities would greatly reduce the external source of error. Such arrangements were limited in academia and user responsibility towards the equipment remains important towards minimising these sources of errors.

Managing a successful screen required each phase to be well integrated and flow from one step to the next. This was possible with organised schedule and strict time controls, supported by software which assisted with information handling, provided updated status and alleviate the load on data handling. While the popular belief holds true that high-throughput processes are efficient and massively scaled the data generation potential, a tremendous amount of optimisation had to be invested to bring each unique aspect together in a symphonic setup to make possible a screen.

The current RISCI screen was successful in its attempt at identifying the apoptosis inducers traditionally assayed for by the original setup, and extended further this screening capability by integrating an additional apoptosis inducing treatment leading to the identification of potential inhibitors and sensitizers of apoptosis.

Future Prospects

The RISCI platform offers many exciting prospects since both robots were updated to be programmable in executing user generated protocols. The DNA isolation robot could be optimised further to increase its daily throughput to more than forty microplates of plasmids through the use of magnetic DNA extraction beads. Commercially available magnetic beads have been tested to be suitable for integration with the current ultrapure DNA isolation protocol, and these magnetic beads could further reduced the processing time to below two hours by eliminating the centrifugation steps and accelerating the repeated cycles of washings. Integration of commercial magnetic beads handling system such as the KingFisher Flex Magnetic Particle Processors (Thermo Scientific) would further contribute to the increase in throughput.

The transfection with its current capacity of up to forty eight 96-well plates may still manage the increase in DNA isolation output although an additional liquid handling robot would improve the process efficiency. An automated plate stacker could be integrated with the Optima FLUOstar plate reader to handle the individual plate measurements and reduce user involvement. However, the upper limit of the throughput is limited with 96-well format, and a transition towards 384-plates (the current industry preferred format) or higher plate density may be necessary. The use of microfluidics could also augment the potential of the screen [545-547]. Considerations should also be directed towards the "cherry-picking" of potential clones into a secondary plate for screen. Whereas the RISCI platform features automation technologies which is crucial to successful high-throughput screens, automation of processes terminate after the transfection stages. The compilation of candidate clones was manually performed over an extended period of time; this process is both dull and laborious, and should be automated to keep pace with further increase in the screening efficiency upstream of this step.

The RISCI screen platform under its current setup could be exploited to study the genetic and signalling profiles of anti-cancer compounds, many of which activate apoptosis. Understanding the unique signalling triggered by each compound could be a more efficient way of developing combination treatment or polypills containing multiple active ingredients which could substantially reduce the development frequency of tumour resistance [548-550]. Alternatively, the default screening cell line may be varied such that the setup identifies anti-cancer genes. Orctl3 was previously identified as a tumour specific gene under such a similar setup [283]. This approach is currently being pursued and a significantly increased count of anti-cancer genes is likely to be generated due to the diversity of the NITE library.

The analysis of the NITE apoptosis regulator candidates offered countless points for potential research directions. For example, the pathway analysis identified numerous probable “stress points” within a signalling cascade which would be exploited to induce a pro-apoptotic effect for cancer therapy. The up-regulation of the metal ion response genes could perhaps also be verified for the prospect of using them as biomarkers for nanoparticles exposure. The three extensive gene lists for apoptosis inducers, inhibitors and sensitizers could offer numerous research opportunities by exploring their apoptosis signalling pathway.

The most exciting outcome of the NITE screen was the experimental identification of alternative lists comprising of novel noncoding sequences to be involved in apoptosis signalling process. Since there is little known about the involvement of noncoding RNA in the regulation of apoptosis, and multiple hypothetical regulatory mechanisms, the work in this direction would be a novelty with extensive unexplored research opportunities. The experimental validation of the reciprocal relationship between an apoptosis regulating microRNA and its targeting of candidates initiating the opposite effect would also be of great interest. This extends the signalling mechanisms beyond the current protein-based functionality and regulatory controls to include novel non-protein based regulatory mechanisms.

Conclusion

The project has definitely come a full circle, gradually progressing from the initial experiment setup and optimisation phase, to the development of the RISCI platforms, to the shift of focus towards the screen implementation, scaling up to the full screening capacity and finally the validation and analysis of the identified candidates.

The project has offered an exciting mix of novelty with a test of endurance, many aspects of the screen having no precedence. This included the attempt to screen a fully sequenced library, the opportunity to manage its accompanying high-throughput workflows such as a mini-sequencing project and large-scale bacterial transformation, the introduction of an additional pro-apoptosis chemical signal, attempts to isolate apoptosis inhibitors and sensitizers in parallel with inducers, and the dynamic candidate selection criteria.

The end results proved extremely revealing and greatly demonstrated the ability of high-throughput techniques to offer insights into signalling networks. Analysis of the candidates using currently available data revealed that the apoptosis inducing candidates were highly enriched for cell death and apoptosis associated terms, a strong evident in support of our screen setup to identify apoptosis regulators. Furthermore, the inhibitors were found to be distinct from the inducers and sensitizers where the former was highly associated with transcriptional based functions and localised to the nucleus while the latter carried out their functions via protein based mechanisms. This is further demonstrated by the IPA analysis which identified HNF1A, a transcription factor, to be an upstream regulator for a wide range of the inhibitor candidates, an observation that was not present in the inducer or sensitizer candidates. This observation is in line with the current understanding of apoptosis, where the inducers and sensitizers are regulated and activated at protein level to initiate a rapid response while apoptosis inhibition is regulated at transcription level to be metabolically efficient and prevents unintended cell immortalisation.

Extensive cross-talking within the apoptosis network was demonstrated through the IPA integrated network where a vast number of known interactions among the candidates were identified. The inducers displayed a strong enrichment for various transport functions and transmembrane localisation, which was similar to the observation of the previous screen, suggesting an integral nature of metabolism and apoptosis and that the inducers may function as sensors to activate apoptosis in the event of any homeostatic imbalance. The protein binding function of the sensitizers suggests that they could potentially serve as mediators between various proteins to coordinate their

function or assist in inter-network cross-talking. Moreover, the large number of development associated terms enriched among the sensitizers along with their similarity to the inducers suggests that the sensitizers may also hold sensor functions to activate programmed cell death in the event of erroneous development outcome.

In contrast, the inhibitors displayed strong association with various cancer profiles, supporting their anti-apoptosis function. This interesting revelation also indicated that the signalling pathways employed by these cancer cells to avoid apoptosis may indeed have a normal physiological role as a stress response, since the inhibitors were identified for their pro-survival response to silica nanoparticles treatment. Uniquely, the inhibitors were enriched for response to metal ion, a well-established stress response mechanism to reactive oxygen species generated in the presence of the metal ions, indicating that the silica nanoparticles may produce its toxic effects using a similar mode of action.

The pinnacle discovery of the RISCI screen was the ability of noncoding RNA sequences to participate in apoptosis regulation. These once seemingly obscured and unexplained sequences are beginning to take centre stage with increasing interest from the research community and high profile projects such as ENCODE. Using curated databases and available predictive algorithms, insights into their mechanisms could be proposed. Rather than a single mode of action, these RNA sequences are likely to achieve their regulatory effect through a range of mechanisms. Some for example are likely to be microRNA precursors, hence the source of endogenous RNAi generation for target knockdown. Others could potentially form secondary structures which act as scaffolds for transcription complex recruitment or protein-protein complex formation. Many noncoding sequences also displayed ability to bind microRNAs across their entirety, potentially acting as decoys to sequester their target and halt the endogenous gene silencing mechanism. These noncoding sequences could possibly also reciprocally regulate their protein-coding counterparts also identified in the screen based on prediction of their targets; for example a noncoding inducer candidate could generate a microRNA targeting a protein-coding inhibitor candidate, thereby alleviating apoptosis inhibition for the process to advance.

Although the candidates identified from the screen were statistically significant and based on bioinformatics and signalling pathway profiling demonstrated to be highly likely to possess their respective apoptosis regulatory function, further experimental verification would be required to fully elucidate and confirm their role within the vast apoptosis signalling network. Indeed, the primary function of a high-throughput is to massively search available targets to propose an association between the candidates and phenotype-of-interest, providing the starting point for further

hypothesis generation in a discovery-driven approach. So while much could be deduced from existing knowledge about the candidates, the truly interesting part of the project outcome remains those unknown candidates that were implicated in apoptosis regulation for the first time. Dissemination of the consolidated candidate list to the wider research community would help speed up the experimental validation of these candidates and aid other researchers to derive novel hypotheses when the candidates are placed in their research context.

The novelty of the screen at times proved to be persistently difficult to address, and more often than not holds its surprises with additional setbacks despite meticulous planning. This together with the responsibility of managing a unique pair of robot and the need to execute the screen successfully can become exceedingly overwhelming.

The various aspects of the screen had to converge perfectly for a success. The RISCI platform holds true to its capacity to execute the high-throughput workflows, while the assay and selection criteria were sufficiently significant to result in the lists of apoptosis inducers, inhibitors and sensitizers many of which are associated for the first time with apoptosis. Furthermore, the discovery of the potential of noncoding sequences in regulating apoptosis represents another significant outcome of the screen alongside the candidate lists, highlighting the importance of high-throughput approaches such as the RISCI screen in understanding the complex signalling networks.

The process has been trialled and tested in its entirety, and can be conveniently adapted for any workflow. The candidate gene lists together with the prospects of the RNA regulatory mechanisms would make available abundant research directions, allowing this chapter of the RISCI screen to be concluded in favour of the exciting opportunities that lie ahead.

Web Resources

CAP3 Sequence Assembly Program	http://pbil.univ-lyon1.fr/cap3.php
CASVM Caspase Substrate Prediction	http://casbase.org/casvm/index.html
Dependence Receptors Prediction	http://bis.ifc.unam.mx/DependenceReceptors/
DAVID Function Annotation	http://david.abcc.ncifcrf.gov/
UniProt Protein Database	http://www.uniprot.org/
NCBI BLAST	http://blast.ncbi.nlm.nih.gov
Gene List Comparison	http://nemates.org/MA/progs/Compare.html
PANTHER Classification System	http://www.pantherdb.org/
bioDBnet	http://biodbnet.abcc.ncifcrf.gov/
BIND database	http://bind.ca
MINT database	http://mint.bio.uniroma2.it/mint/
Reactome	http://www.reactome.org/
miRBase database	http://www.mirbase.org/
RegRNA Motif Prediction	http://regrna.mbc.nctu.edu.tw/html/prediction.html
Rfam database	http://rfam.sanger.ac.uk/search
miRMaid	http://plugins.mirmaid.org/home

Supplementary Materials

SM 1: Electronic version of the thesis

SM 2: Poster - Investigating apoptosis and toxicity signalling with RISCI

SM 3: Gene Lists and Annotations

SM 4: DAVID Analysis containing all annotations; DAVID Summary.xls summarises the analysis.

SM 5: GSEA; access the index.html within each folder to view analysis

SM 6: Interactions identified by GeneMANIA

SM 7: Ingenuity Pathway Analysis

SM 8: Compiled List of All Enriched Signalling Pathways

SM 9: Rfam Results for noncoding sequence analysis

SM 10: RegRNA Annotations of RNA Regulatory Motifs; access the HTML file of each gene accession number to view annotation

SM 11: TargetScan Prediction of microRNA targets associated with candidates

SM 12: Mini-protein Analysis

SM 13: Confocal Live Imaging - Activation of autophagy by ammonium chloride

SM 14: Confocal Live Imaging - Cleavage of caspase-3 FRET construct

SM 15: Video - Triplicate DNA Transfection Automation

References

1. GlÜCksmann, A., *CELL DEATHS IN NORMAL VERTEBRATE ONTOGENY.* Biological Reviews, 1951. **26**(1): p. 59-86.
2. Lockshin, R.A. and C.M. Williams, *Programmed cell death—II. Endocrine potentiation of the breakdown of the intersegmental muscles of silkmoths.* Journal of Insect Physiology, 1964. **10**(4): p. 643-649.
3. Kerr, J.F., A.H. Wyllie, and A.R. Currie, *Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics.* Br J Cancer, 1972. **26**(4): p. 239-57.
4. Brenner, S., *The genetics of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 71-94.
5. Sulston, J.E. and S. Brenner, *The DNA of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 95-104.
6. Yuan, J. and H.R. Horvitz, *The Caenorhabditis elegans genes ced-3 and ced-4 act cell autonomously to cause programmed cell death.* Developmental Biology, 1990. **138**(1): p. 33-41.
7. Yuan, J., et al., *The C. elegans cell death gene ced-3 encodes a protein similar to mammalian interleukin-1β-converting enzyme.* Cell, 1993. **75**(4): p. 641-652.
8. Li, P., et al., *Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade.* Cell, 1997. **91**(4): p. 479-89.
9. Tsujimoto, Y., *Role of Bcl-2 family proteins in apoptosis: apoptosomes or mitochondria?* Genes to Cells, 1998. **3**(11): p. 697-707.
10. Yuan, J. and H.R. Horvitz, *A first insight into the molecular mechanisms of apoptosis.* Cell, 2004. **116, Supplement 2**(0): p. S53-S56.
11. Mondello, C. and A.I. Scovassi, *Apoptosis: A Way to Maintain Healthy Individuals*

Genome Stability and Human Diseases, H.-P. Nasheuer, Editor. 2010, Springer Netherlands. p. 307-323.

12. Feig, C. and M.E. Peter, *How apoptosis got the immune system in shape.* European Journal of Immunology, 2007. **37**(S1): p. S61-S70.
13. Trump, B.E., et al., *The Pathways of Cell Death: Oncosis, Apoptosis, and Necrosis.* Toxicologic Pathology, 1997. **25**(1): p. 82-88.
14. Majno, G. and I. Joris, *Apoptosis, oncosis, and necrosis. An overview of cell death.* Am J Pathol, 1995. **146**(1): p. 3-15.
15. Wyllie, A.H., J.F.R. Kerr, and A.R. Currie, *Cell Death: The Significance of Apoptosis*, in *International Review of Cytology*, J.F.D. G.H. Bourne and K.W. Jeon, Editors. 1980, Academic Press. p. 251-306.
16. Arends, M.J. and A.H. Wyllie, *Apoptosis: mechanisms and roles in pathology.* Int Rev Exp Pathol, 1991. **32**: p. 223-54.
17. Lemasters, J.J., et al., *Mitochondrial calcium and the permeability transition in cell death.* Biochim Biophys Acta, 2009. **1787**(11): p. 1395-401.
18. Kluck, R.M., et al., *The release of cytochrome c from mitochondria: a primary site for Bcl-2 regulation of apoptosis.* Science, 1997. **275**(5303): p. 1132-6.
19. Susin, S.A., et al., *Molecular characterization of mitochondrial apoptosis-inducing factor.* Nature, 1999. **397**(6718): p. 441-6.
20. Wyllie, A.H., *Glucocorticoid-induced thymocyte apoptosis is associated with endogenous endonuclease activation.* Nature, 1980. **284**(5756): p. 555-6.
21. Fadok, V.A. and P.M. Henson, *Apoptosis: getting rid of the bodies.* Curr Biol, 1998. **8**(19): p. R693-5.
22. Wyllie, A.H., *Apoptosis: an overview.* Br Med Bull, 1997. **53**(3): p. 451-65.
23. Vaux, D.L. and A. Strasser, *The molecular biology of apoptosis.* Proc Natl Acad Sci U S A, 1996. **93**(6): p. 2239-44.

24. Koopman, G., et al., *Annexin V for flow cytometric detection of phosphatidylserine expression on B cells undergoing apoptosis.* Blood, 1994. **84**(5): p. 1415-20.
25. Peng, L., H. Jiang, and C. Bradley, *[Annexin V for flow cytometric detection of phosphatidylserine expression on lymphoma cells undergoing apoptosis].* Hua Xi Yi Ke Da Xue Xue Bao, 2001. **32**(4): p. 602-4, 620.
26. Gavrieli, Y., Y. Sherman, and S.A. Ben-Sasson, *Identification of programmed cell death in situ via specific labeling of nuclear DNA fragmentation.* J Cell Biol, 1992. **119**(3): p. 493-501.
27. Nicoletti, I., et al., *A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry.* J Immunol Methods, 1991. **139**(2): p. 271-9.
28. Levine, B. and D.J. Klionsky, *Development by self-digestion: molecular mechanisms and biological functions of autophagy.* Dev Cell, 2004. **6**(4): p. 463-77.
29. Hannigan, A.M. and S.M. Gorski, *Macroautophagy: The key ingredient to a healthy diet?* Autophagy, 2009. **5**(2): p. 140-151.
30. Thorburn, A., *Apoptosis and autophagy: regulatory connections between two supposedly different processes.* Apoptosis, 2008. **13**(1): p. 1-9.
31. Liu, B., et al., *Polygonatum cyrtonema lectin induces apoptosis and autophagy in human melanoma A375 cells through a mitochondria-mediated ROS-p38-p53 pathway.* Cancer Lett, 2009. **275**(1): p. 54-60.
32. Liu, J., et al., *Beclin1 Controls the Levels of p53 by Regulating the Deubiquitination Activity of USP10 and USP13.* Cell, 2011. **147**(1): p. 223-234.
33. Lemasters, J.J., *V. Necrapoptosis and the mitochondrial permeability transition: shared pathways to necrosis and apoptosis.* American Journal of Physiology - Gastrointestinal and Liver Physiology, 1999. **276**(1): p. G1-G6.
34. Sperandio, S., I. de Belle, and D.E. Bredesen, *An alternative, nonapoptotic form of programmed cell death.* Proceedings of the National Academy of Sciences, 2000. **97**(26): p. 14376-14381.
35. Chinnaiyan, A.M., *The apoptosome: heart and soul of the cell death machine.* Neoplasia, 1999. **1**(1): p. 5-15.
36. Chang, D.W., et al., *Interdimer processing mechanism of procaspase-8 activation.* EMBO J, 2003. **22**(16): p. 4132-42.
37. Nagata, S., *Fas ligand-induced apoptosis.* Annu Rev Genet, 1999. **33**: p. 29-55.
38. Peter, M.E. and P.H. Krammer, *The CD95(APO-1/Fas) DISC and beyond.* Cell Death Differ, 2003. **10**(1): p. 26-35.
39. Luo, X., et al., *Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors.* Cell, 1998. **94**(4): p. 481-90.
40. Li, H., et al., *Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis.* Cell, 1998. **94**(4): p. 491-501.
41. Green, D.R. and G. Kroemer, *The pathophysiology of mitochondrial cell death.* Science, 2004. **305**(5684): p. 626-9.
42. Saelens, X., et al., *Toxic proteins released from mitochondria in cell death.* Oncogene, 2004. **23**(16): p. 2861-74.
43. Jiang, X. and X. Wang, *Cytochrome c promotes caspase-9 activation by inducing nucleotide binding to Apaf-1.* J Biol Chem, 2000. **275**(40): p. 31199-203.
44. Rodriguez, J. and Y. Lazebnik, *Caspase-9 and APAF-1 form an active holoenzyme.* Genes Dev, 1999. **13**(24): p. 3179-84.
45. Hao, Z., et al., *Specific Ablation of the Apoptotic Functions of Cytochrome c Reveals a Differential Requirement for Cytochrome c and Apaf-1 in Apoptosis.* Cell, 2005. **121**(4): p. 579-591.
46. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.
47. Alnemri, E.S., et al., *Human ICE/CED-3 protease nomenclature.* Cell, 1996. **87**(2): p. 171.

48. Thornberry, N.A., et al., *A novel heterodimeric cysteine protease is required for interleukin-1[beta]processing in monocytes.* Nature, 1992. **356**(6372): p. 768-774.
49. Chai, J., et al., *Crystal structure of a procaspase-7 zymogen: mechanisms of activation and substrate binding.* Cell, 2001. **107**(3): p. 399-407.
50. Shi, Y., *Mechanisms of caspase activation and inhibition during apoptosis.* Mol Cell, 2002. **9**(3): p. 459-70.
51. Ni, C.Z., et al., *Conformational restrictions in the active site of unliganded human caspase-3.* J Mol Recognit, 2003. **16**(3): p. 121-4.
52. Thornberry, N.A., et al., *A combinatorial approach defines specificities of members of the caspase family and granzyme B. Functional relationships established for key mediators of apoptosis.* J Biol Chem, 1997. **272**(29): p. 17907-11.
53. Riedl, S.J. and Y. Shi, *Molecular mechanisms of caspase regulation during apoptosis.* Nat Rev Mol Cell Biol, 2004. **5**(11): p. 897-907.
54. Miller, L.K., *An exegesis of IAPs: salvation and surprises from BIR motifs.* Trends Cell Biol, 1999. **9**(8): p. 323-8.
55. Jabbour, A.M., et al., *The p35 relative, p49, inhibits mammalian and Drosophila caspases including DRONC and protects against apoptosis.* Cell Death Differ, 2002. **9**(12): p. 1311-20.
56. Huntington, J.A., R.J. Read, and R.W. Carrell, *Structure of a serpin-protease complex shows inhibition by deformation.* Nature, 2000. **407**(6806): p. 923-6.
57. Liston, P., W.G. Fong, and R.G. Korneluk, *The inhibitors of apoptosis: there is more to life than Bcl2.* Oncogene, 2003. **22**(53): p. 8568-80.
58. Deveraux, Q.L., et al., *X-linked IAP is a direct inhibitor of cell-death proteases.* Nature, 1997. **388**(6639): p. 300-4.
59. Takahashi, R., et al., *A Single BIR Domain of XIAP Sufficient for Inhibiting Caspases.* Journal of Biological Chemistry, 1998. **273**(14): p. 7787-7790.
60. Roy, N., et al., *The c-IAP-1 and c-IAP-2 proteins are direct inhibitors of specific caspases.* EMBO J, 1997. **16**(23): p. 6914-6925.
61. Shin, S., et al., *An Anti-apoptotic Protein Human Survivin Is a Direct Inhibitor of Caspase-3 and -7†.* Biochemistry, 2001. **40**(4): p. 1117-1123.
62. Richter, B.W.M., et al., *Molecular Cloning of ILP-2, a Novel Member of the Inhibitor of Apoptosis Protein Family.* Mol Cell Biol, 2001. **21**(13): p. 4292-4301.
63. Lin, J.-H., et al., *KIAP, a Novel Member of the Inhibitor of Apoptosis Protein Family.* Biochemical and Biophysical Research Communications, 2000. **279**(3): p. 820-831.
64. Chai, J., et al., *Structural Basis of Caspase-7 Inhibition by XIAP.* Cell, 2001. **104**(5): p. 769-780.
65. Huang, Y., et al., *Structural Basis of Caspase Inhibition by XIAP: Differential Roles of the Linker versus the BIR Domain.* Cell, 2001. **104**(5): p. 781-790.
66. Suzuki, Y., et al., *X-linked Inhibitor of Apoptosis Protein (XIAP) Inhibits Caspase-3 and -7 in Distinct Modes.* Journal of Biological Chemistry, 2001. **276**(29): p. 27058-27063.
67. Shiozaki, E.N., et al., *Mechanism of XIAP-Mediated Inhibition of Caspase-9.* Mol Cell, 2003. **11**(2): p. 519-527.
68. Yang, Y., et al., *Ubiquitin Protein Ligase Activity of IAPs and Their Degradation in Proteasomes in Response to Apoptotic Stimuli.* Science, 2000. **288**(5467): p. 874-877.
69. Huang, H.-k., et al., *The Inhibitor of Apoptosis, cIAP2, Functions as a Ubiquitin-Protein Ligase and Promotes in VitroMonoubiquitination of Caspases 3 and 7.* Journal of Biological Chemistry, 2000. **275**(35): p. 26661-26664.
70. Suzuki, Y., Y. Nakabayashi, and R. Takahashi, *Ubiquitin-protein ligase activity of X-linked inhibitor of apoptosis protein promotes proteasomal degradation of caspase-3 and enhances its anti-apoptotic effect in Fas-induced cell death.* Proceedings of the National Academy of Sciences, 2001. **98**(15): p. 8662-8667.
71. Srinivasula, S.M., et al., *A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis.* Nature, 2001. **410**(6824): p. 112-6.

72. MacFarlane, M., et al., *Proteasome-mediated Degradation of Smac during Apoptosis: XIAP Promotes Smac Ubiquitination in Vitro.* Journal of Biological Chemistry, 2002. **277**(39): p. 36611-36616.
73. Hu, S. and X. Yang, *Cellular Inhibitor of Apoptosis 1 and 2 Are Ubiquitin Ligases for the Apoptosis Inducer Smac/DIABLO.* Journal of Biological Chemistry, 2003. **278**(12): p. 10055-10060.
74. Clem, R.J., et al., *c-IAP1 Is Cleaved by Caspases to Produce a Proapoptotic C-terminal Fragment.* Journal of Biological Chemistry, 2001. **276**(10): p. 7602-7608.
75. Fong, W.G., et al., *Expression and Genetic Analysis of XIAP-Associated Factor 1 (XAF1) in Cancer Cell Lines.* Genomics, 2000. **70**(1): p. 113-122.
76. Liston, P., et al., *Identification of XAF1 as an antagonist of XIAP anti-Caspase activity.* Nat Cell Biol, 2001. **3**(2): p. 128-133.
77. Tewari, M., et al., *Yama/CPP32 beta, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly(ADP-ribose) polymerase.* Cell, 1995. **81**(5): p. 801-9.
78. Duan, H., et al., *ICE-LAP6, a novel member of the ICE/Ced-3 gene family, is activated by the cytotoxic T cell protease granzyme B.* J Biol Chem, 1996. **271**(28): p. 16720-4.
79. Zou, H., et al., *Apaf-1, a human protein homologous to C. elegans CED-4, participates in cytochrome c-dependent activation of caspase-3.* Cell, 1997. **90**(3): p. 405-13.
80. Liu, X., et al., *Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c.* Cell, 1996. **86**(1): p. 147-57.
81. Yu, X., et al., *A structure of the human apoptosome at 12.8 A resolution provides insights into this cell death platform.* Structure, 2005. **13**(11): p. 1725-35.
82. Yuan, S., et al., *Structure of an apoptosome-procaspase-9 CARD complex.* Structure, 2010. **18**(5): p. 571-83.
83. Hu, Y., et al., *WD-40 repeat region regulates Apaf-1 self-association and procaspase-9 activation.* J Biol Chem, 1998. **273**(50): p. 33489-94.
84. Diemand, A.V. and A.N. Lupas, *Modeling AAA+ ring complexes from monomeric structures.* J Struct Biol, 2006. **156**(1): p. 230-43.
85. Yuan, S., et al., *Structure of an Apoptosome-Procaspase-9 CARD Complex.* Structure (London, England : 1993), 2010. **18**(5): p. 571-583.
86. Reubold, T.F., S. Wohlgemuth, and S. Eschenburg, *A new model for the transition of APAF-1 from inactive monomer to caspase-activating apoptosome.* J Biol Chem, 2009. **284**(47): p. 32717-24.
87. Aggarwal, B.B., et al., *Human tumor necrosis factor. Production, purification, and characterization.* J Biol Chem, 1985. **260**(4): p. 2345-54.
88. Aggarwal, B.B., B. Moffat, and R.N. Harkins, *Human lymphotoxin. Production by a lymphoblastoid cell line, purification, and initial characterization.* J Biol Chem, 1984. **259**(1): p. 686-91.
89. Haridas, V., et al., *Overexpression of the p80 TNF receptor leads to TNF-dependent apoptosis, nuclear factor-kappa B activation, and c-Jun kinase activation.* J Immunol, 1998. **160**(7): p. 3152-62.
90. Aggarwal, B.B., S.C. Gupta, and J.H. Kim, *Historical perspectives on tumor necrosis factor and its superfamily: 25 years later, a golden journey.* Blood, 2012. **119**(3): p. 651-65.
91. Gaur, U. and B.B. Aggarwal, *Regulation of proliferation, survival and apoptosis by members of the TNF superfamily.* Biochem Pharmacol, 2003. **66**(8): p. 1403-8.
92. Wajant, H., K. Pfizenmaier, and P. Scheurich, *Tumor necrosis factor signaling.* Cell Death Differ, 2003. **10**(1): p. 45-65.
93. Chen, G. and D.V. Goeddel, *TNF-R1 signaling: a beautiful pathway.* Science, 2002. **296**(5573): p. 1634-5.

94. Mahul-Mellier, A.L., et al., *De-ubiquitinating proteases USP2a and USP2c cause apoptosis by stabilising RIP1.* Biochim Biophys Acta, 2012. **1823**(8): p. 1353-65.
95. Boyce, E.G., J. Halilovic, and O. Stan-Ugbene, *Golimumab: Review of the efficacy and tolerability of a recently approved tumor necrosis factor-alpha inhibitor.* Clin Ther, 2010. **32**(10): p. 1681-703.
96. Hartung, H.P. and B.C. Kieseier, *Atacicept: targeting B cells in multiple sclerosis.* Ther Adv Neurol Disord, 2010. **3**(4): p. 205-16.
97. Tsujimoto, Y., et al., *Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11;14) chromosome translocation.* Science, 1984. **224**(4656): p. 1403-6.
98. Vaux, D.L., S. Cory, and J.M. Adams, *Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells.* Nature, 1988. **335**(6189): p. 440-2.
99. Adams, J.M. and S. Cory, *The Bcl-2 protein family: arbiters of cell survival.* Science, 1998. **281**(5381): p. 1322-6.
100. Willis, S.N. and J.M. Adams, *Life in the balance: how BH3-only proteins induce apoptosis.* Curr Opin Cell Biol, 2005. **17**(6): p. 617-25.
101. Mikhailov, V., et al., *Bcl-2 prevents Bax oligomerization in the mitochondrial outer membrane.* J Biol Chem, 2001. **276**(21): p. 18361-74.
102. Mikhailov, V., et al., *Association of Bax and Bak homo-oligomers in mitochondria. Bax requirement for Bak reorganization and cytochrome c release.* J Biol Chem, 2003. **278**(7): p. 5367-76.
103. Chen, L., et al., *Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function.* Mol Cell, 2005. **17**(3): p. 393-403.
104. Certo, M., et al., *Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members.* Cancer Cell, 2006. **9**(5): p. 351-65.
105. Wei, M.C., et al., *Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death.* Science, 2001. **292**(5517): p. 727-30.
106. Sharpe, J.C., D. Arnoult, and R.J. Youle, *Control of mitochondrial permeability by Bcl-2 family members.* Biochim Biophys Acta, 2004. **1644**(2-3): p. 107-13.
107. Grinberg, M., et al., *tBID Homooligomerizes in the mitochondrial membrane to induce apoptosis.* J Biol Chem, 2002. **277**(14): p. 12237-45.
108. Muchmore, S.W., et al., *X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death.* Nature, 1996. **381**(6580): p. 335-41.
109. Schendel, S.L., M. Montal, and J.C. Reed, *Bcl-2 family proteins as ion-channels.* Cell Death Differ, 1998. **5**(5): p. 372-80.
110. Garcia-Saez, A.J., et al., *Peptides derived from apoptotic Bax and Bid reproduce the poration activity of the parent full-length proteins.* Biophys J, 2005. **88**(6): p. 3976-90.
111. Dejean, L.M., et al., *Oligomeric Bax is a component of the putative cytochrome c release channel MAC, mitochondrial apoptosis-induced channel.* Mol Biol Cell, 2005. **16**(5): p. 2424-32.
112. Pavlov, E.V., et al., *A novel, high conductance channel of mitochondria linked to apoptosis in mammalian cells and Bax expression in yeast.* J Cell Biol, 2001. **155**(5): p. 725-31.
113. Gong, X.M., et al., *Conformation of membrane-associated proapoptotic tBid.* J Biol Chem, 2004. **279**(28): p. 28954-60.
114. Thomenius, M.J. and C.W. Distelhorst, *Bcl-2 on the endoplasmic reticulum: protecting the mitochondria from a distance.* J Cell Sci, 2003. **116**(Pt 22): p. 4493-9.
115. Rudner, J., V. Jendrossek, and C. Belka, *New insights in the role of Bcl-2 Bcl-2 and the endoplasmic reticulum.* Apoptosis, 2002. **7**(5): p. 441-7.
116. Thomenius, M.J., et al., *Bcl-2 on the endoplasmic reticulum regulates Bax activity by binding to BH3-only proteins.* J Biol Chem, 2003. **278**(8): p. 6243-50.

117. Pattingre, S., et al., *Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy.* Cell, 2005. **122**(6): p. 927-39.
118. Puthalakath, H. and A. Strasser, *Keeping killers on a tight leash: transcriptional and post-translational control of the pro-apoptotic activity of BH3-only proteins.* Cell Death Differ, 2002. **9**(5): p. 505-12.
119. Gross, A., J.M. McDonnell, and S.J. Korsmeyer, *BCL-2 family members and the mitochondria in apoptosis.* Genes Dev, 1999. **13**(15): p. 1899-911.
120. Huang, Z., *Bcl-2 family proteins as targets for anticancer drug design.* Oncogene, 2000. **19**(56): p. 6627-31.
121. Kang, M.H. and C.P. Reynolds, *Bcl-2 inhibitors: targeting mitochondrial apoptotic pathways in cancer therapy.* Clin Cancer Res, 2009. **15**(4): p. 1126-32.
122. Pechan, P.M., *Heat shock proteins and cell proliferation.* FEBS Lett, 1991. **280**(1): p. 1-4.
123. Helmbrecht, K., E. Zeise, and L. Rensing, *Chaperones in cell cycle regulation and mitogenic signal transduction: a review.* Cell Prolif, 2000. **33**(6): p. 341-65.
124. Takayama, S., J.C. Reed, and S. Homma, *Heat-shock proteins as regulators of apoptosis.* Oncogene, 2003. **22**(56): p. 9041-7.
125. Csermely, P., *Chaperone overload is a possible contributor to 'civilization diseases'.* Trends Genet, 2001. **17**(12): p. 701-4.
126. Soti, C. and P. Csermely, *Aging and molecular chaperones.* Exp Gerontol, 2003. **38**(10): p. 1037-40.
127. Sreedhar, A.S. and P. Csermely, *Heat shock proteins in the regulation of apoptosis: new strategies in tumor therapy: a comprehensive review.* Pharmacol Ther, 2004. **101**(3): p. 227-57.
128. Rogalla, T., et al., *Regulation of Hsp27 oligomerization, chaperone function, and protective activity against oxidative stress/tumor necrosis factor alpha by phosphorylation.* J Biol Chem, 1999. **274**(27): p. 18947-56.
129. Bruey, J.M., et al., *Hsp27 negatively regulates cell death by interacting with cytochrome c.* Nat Cell Biol, 2000. **2**(9): p. 645-52.
130. Samali, A., et al., *Presence of a pre-apoptotic complex of pro-caspase-3, Hsp60 and Hsp10 in the mitochondrial fraction of jurkat cells.* EMBO J, 1999. **18**(8): p. 2040-8.
131. Gupta, S. and A.A. Knowlton, *Cytosolic heat shock protein 60, hypoxia, and apoptosis.* Circulation, 2002. **106**(21): p. 2727-33.
132. Saleh, A., et al., *Negative regulation of the Apaf-1 apoptosome by Hsp70.* Nat Cell Biol, 2000. **2**(8): p. 476-83.
133. Jaattela, M., et al., *Hsp70 exerts its anti-apoptotic function downstream of caspase-3-like proteases.* EMBO J, 1998. **17**(21): p. 6124-34.
134. Gao, T. and A.C. Newton, *The turn motif is a phosphorylation switch that regulates the binding of Hsp70 to protein kinase C.* J Biol Chem, 2002. **277**(35): p. 31585-92.
135. Konishi, H., et al., *Activation of protein kinase B (Akt/RAC-protein kinase) by cellular stress and its association with heat shock protein Hsp27.* FEBS Lett, 1997. **410**(2-3): p. 493-8.
136. Mosser, D.D., et al., *The chaperone function of hsp70 is required for protection against stress-induced apoptosis.* Mol Cell Biol, 2000. **20**(19): p. 7146-59.
137. Sakahira, H. and S. Nagata, *Co-translational folding of caspase-activated DNase with Hsp70, Hsp40, and inhibitor of caspase-activated DNase.* J Biol Chem, 2002. **277**(5): p. 3364-70.
138. Lewis, J., et al., *Disruption of hsp90 function results in degradation of the death domain kinase, receptor-interacting protein (RIP), and blockage of tumor necrosis factor-induced nuclear factor-kappaB activation.* J Biol Chem, 2000. **275**(14): p. 10519-26.
139. Chen, G., P. Cao, and D.V. Goeddel, *TNF-induced recruitment and activation of the IKK complex require Cdc37 and Hsp90.* Mol Cell, 2002. **9**(2): p. 401-10.
140. Nakagawa, T., et al., *Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta.* Nature, 2000. **403**(6765): p. 98-103.

141. Kaufman, R.J., *Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls.* Genes Dev, 1999. **13**(10): p. 1211-33.
142. Reddy, R.K., et al., *Endoplasmic reticulum chaperone protein GRP78 protects cells from apoptosis induced by topoisomerase inhibitors: role of ATP binding site in suppression of caspase-7 activation.* J Biol Chem, 2003. **278**(23): p. 20915-24.
143. Chen, L. and X. Gao, *Neuronal apoptosis induced by endoplasmic reticulum stress.* Neurochem Res, 2002. **27**(9): p. 891-8.
144. Berridge, M.J., *The endoplasmic reticulum: a multifunctional signaling organelle.* Cell Calcium, 2002. **32**(5-6): p. 235-49.
145. Zong, W.X., et al., *Bax and Bak can localize to the endoplasmic reticulum to initiate apoptosis.* J Cell Biol, 2003. **162**(1): p. 59-69.
146. Iwasawa, R., et al., *Fis1 and Bap31 bridge the mitochondria-ER interface to establish a platform for apoptosis induction.* EMBO J, 2011. **30**(3): p. 556-68.
147. Pinton, P., et al., *Calcium and apoptosis: ER-mitochondria Ca2+ transfer in the control of apoptosis.* Oncogene, 2008. **27**(50): p. 6407-18.
148. Mattson, M.P. and S.L. Chan, *Calcium orchestrates apoptosis.* Nat Cell Biol, 2003. **5**(12): p. 1041-3.
149. Soussi, T. and C. Beroud, *Assessing TP53 status in human tumours to evaluate clinical outcome.* Nat Rev Cancer, 2001. **1**(3): p. 233-40.
150. el-Deiry, W.S., et al., *Definition of a consensus binding site for p53.* Nat Genet, 1992. **1**(1): p. 45-9.
151. Contente, A., et al., *A polymorphic microsatellite that mediates induction of PIG3 by p53.* Nat Genet, 2002. **30**(3): p. 315-20.
152. Buzek, J., et al., *Redox state of tumor suppressor p53 regulates its sequence-specific DNA binding in DNA-damaged cells by cysteine 277.* Nucleic Acids Res, 2002. **30**(11): p. 2340-8.
153. Seoane, J., H.V. Le, and J. Massague, *Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage.* Nature, 2002. **419**(6908): p. 729-34.
154. Samuels-Lev, Y., et al., *ASPP proteins specifically stimulate the apoptotic function of p53.* Mol Cell, 2001. **8**(4): p. 781-94.
155. Muller, M., et al., *p53 activates the CD95 (APO-1/Fas) gene in response to DNA damage by anticancer drugs.* J Exp Med, 1998. **188**(11): p. 2033-45.
156. Wu, G.S., et al., *KILLER/DR5 is a DNA damage-inducible p53-regulated death receptor gene.* Nat Genet, 1997. **17**(2): p. 141-3.
157. Attardi, L.D., et al., *PERP, an apoptosis-associated target of p53, is a novel member of the PMP-22/gas3 family.* Genes Dev, 2000. **14**(6): p. 704-18.
158. Thornborrow, E.C., et al., *A conserved intronic response element mediates direct p53-dependent transcriptional activation of both the human and murine bax genes.* Oncogene, 2002. **21**(7): p. 990-9.
159. Sax, J.K., et al., *BID regulation by p53 contributes to chemosensitivity.* Nat Cell Biol, 2002. **4**(11): p. 842-9.
160. Oda, E., et al., *Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis.* Science, 2000. **288**(5468): p. 1053-8.
161. Yu, J. and L. Zhang, *No PUMA, no death: implications for p53-dependent apoptosis.* Cancer Cell, 2003. **4**(4): p. 248-9.
162. Kaeser, M.D., S. Pebernard, and R.D. Iggo, *Regulation of p53 stability and function in HCT116 colon cancer cells.* J Biol Chem, 2004. **279**(9): p. 7598-605.
163. Ding, H.F., et al., *Oncogene-dependent regulation of caspase activation by p53 protein in a cell-free system.* J Biol Chem, 1998. **273**(43): p. 28378-83.
164. Ding, H.F. and D.E. Fisher, *Mechanisms of p53-mediated apoptosis.* Crit Rev Oncog, 1998. **9**(1): p. 83-98.

165. MacLachlan, T.K. and W.S. El-Deiry, *Apoptotic threshold is lowered by p53 transactivation of caspase-6.* Proc Natl Acad Sci U S A, 2002. **99**(14): p. 9492-7.
166. Marchenko, N.D., A. Zaika, and U.M. Moll, *Death signal-induced localization of p53 protein to mitochondria. A potential role in apoptotic signaling.* J Biol Chem, 2000. **275**(21): p. 16202-12.
167. Mihara, M. and U.M. Moll, *Detection of mitochondrial localization of p53.* Methods Mol Biol, 2003. **234**: p. 203-9.
168. Mihara, M., et al., *p53 has a direct apoptogenic role at the mitochondria.* Mol Cell, 2003. **11**(3): p. 577-90.
169. Mayo, L.D. and D.B. Donner, *The PTEN, Mdm2, p53 tumor suppressor-oncoprotein network.* Trends Biochem Sci, 2002. **27**(9): p. 462-7.
170. Mayo, L.D., et al., *PTEN protects p53 from Mdm2 and sensitizes cancer cells to chemotherapy.* J Biol Chem, 2002. **277**(7): p. 5484-9.
171. Oren, M., et al., *Regulation of p53: intricate loops and delicate balances.* Biochem Pharmacol, 2002. **64**(5-6): p. 865-71.
172. Gottlieb, T.M., et al., *Cross-talk between Akt, p53 and Mdm2: possible implications for the regulation of apoptosis.* Oncogene, 2002. **21**(8): p. 1299-303.
173. Bernet, A. and P. Mehlen, *Dependence receptors: when apoptosis controls tumor progression.* Bull Cancer, 2007. **94**(4): p. E12-7.
174. Rabizadeh, S., et al., *Induction of apoptosis by the low-affinity NGF receptor.* Science, 1993. **261**(5119): p. 345-8.
175. Mehlen, P., et al., *The DCC gene product induces apoptosis by a mechanism requiring receptor proteolysis.* Nature, 1998. **395**(6704): p. 801-4.
176. Llambi, F., et al., *Netrin-1 acts as a survival factor via its receptors UNC5H and DCC.* EMBO J, 2001. **20**(11): p. 2715-22.
177. Ellerby, L.M., et al., *Kennedy's disease: caspase cleavage of the androgen receptor is a crucial event in cytotoxicity.* J Neurochem, 1999. **72**(1): p. 185-95.
178. Thibert, C., et al., *Inhibition of neuroepithelial patched-induced apoptosis by sonic hedgehog.* Science, 2003. **301**(5634): p. 843-6.
179. Matsunaga, E., et al., *RGM and its receptor neogenin regulate neuronal survival.* Nat Cell Biol, 2004. **6**(8): p. 749-55.
180. Stupack, D.G., et al., *Apoptosis of adherent cells by recruitment of caspase-8 to unligated integrins.* J Cell Biol, 2001. **155**(3): p. 459-70.
181. del Rio, G., et al., *A novel motif identified in dependence receptors.* PLoS One, 2007. **2**(5): p. e463.
182. Porter, A.G. and R.U. Janicke, *Emerging roles of caspase-3 in apoptosis.* Cell Death Differ, 1999. **6**(2): p. 99-104.
183. Forcet, C., et al., *The dependence receptor DCC (deleted in colorectal cancer) defines an alternative mechanism for caspase activation.* Proc Natl Acad Sci U S A, 2001. **98**(6): p. 3416-21.
184. Savitz, S.I. and D.M. Rosenbaum, *Apoptosis in neurological disease.* Neurosurgery, 1998. **42**(3): p. 555-72; discussion 573-4.
185. Kam, P.C. and N.I. Ferch, *Apoptosis: mechanisms and clinical implications.* Anaesthesia, 2000. **55**(11): p. 1081-93.
186. Loo, D.T., et al., *Apoptosis is induced by beta-amyloid in cultured central nervous system neurons.* Proc Natl Acad Sci U S A, 1993. **90**(17): p. 7951-5.
187. Osborne, B.A., *Apoptosis and the maintenance of homoeostasis in the immune system.* Curr Opin Immunol, 1996. **8**(2): p. 245-54.
188. Ekert, P.G. and D.L. Vaux, *Apoptosis and the immune system.* Br Med Bull, 1997. **53**(3): p. 591-603.

189. Elkon, K.B., *Mechanisms of autoantibody production and their role in disease.* Mt Sinai J Med, 1994. **61**(4): p. 283-90.
190. Emlen, W., J. Niebur, and R. Kadera, *Accelerated in vitro apoptosis of lymphocytes from patients with systemic lupus erythematosus.* J Immunol, 1994. **152**(7): p. 3685-92.
191. Clem, R.J., J.M. Hardwick, and L.K. Miller, *Anti-apoptotic genes of baculoviruses.* Cell Death Differ, 1996. **3**(1): p. 9-16.
192. Young, L.S., C.W. Dawson, and A.G. Eliopoulos, *Viruses and apoptosis.* Br Med Bull, 1997. **53**(3): p. 509-21.
193. Terai, C., et al., *Apoptosis as a mechanism of cell death in cultured T lymphoblasts acutely infected with HIV-1.* J Clin Invest, 1991. **87**(5): p. 1710-5.
194. Peter, M.E., et al., *AIDS and the death receptors.* Br Med Bull, 1997. **53**(3): p. 604-16.
195. Petitjean, A., et al., *TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes.* Oncogene, 2007. **26**(15): p. 2157-65.
196. Earnshaw, W.C., L.M. Martins, and S.H. Kaufmann, *Mammalian caspases: structure, activation, substrates, and functions during apoptosis.* Annu Rev Biochem, 1999. **68**: p. 383-424.
197. Pitti, R.M., et al., *Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer.* Nature, 1998. **396**(6712): p. 699-703.
198. Kaufmann, S.H. and G.J. Gores, *Apoptosis in cancer: cause and cure.* Bioessays, 2000. **22**(11): p. 1007-17.
199. Haupt, S., et al., *Apoptosis - the p53 network.* J Cell Sci, 2003. **116**(Pt 20): p. 4077-85.
200. Boeckler, F.M., et al., *Targeted rescue of a destabilized mutant of p53 by an in silico screened drug.* Proc Natl Acad Sci U S A, 2008. **105**(30): p. 10360-5.
201. Walczak, H., et al., *Tumoricidal activity of tumor necrosis factor-related apoptosis-inducing ligand in vivo.* Nat Med, 1999. **5**(2): p. 157-63.
202. Lejeune, F.J., C. Ruegg, and D. Lienard, *Clinical applications of TNF-alpha in cancer.* Curr Opin Immunol, 1998. **10**(5): p. 573-80.
203. Oezaslan, M., M. Heggen, and P. Strasser, *Size-Dependent Morphology of Dealloyed Bimetallic Catalysts: Linking the Nano to the Macro Scale.* J Am Chem Soc, 2011. **134**(1): p. 514-524.
204. Stone, A.J., *The Theory of Intermolecular Forces*. 1997: Oxford University Press, USA.
205. Powers, K.W., et al., *Research strategies for safety evaluation of nanomaterials. Part VI. Characterization of nanoscale particles for toxicological evaluation.* Toxicol Sci, 2006. **90**(2): p. 296-303.
206. Hussain, S.M., et al., *The interaction of manganese nanoparticles with PC-12 cells induces dopamine depletion.* Toxicol Sci, 2006. **92**(2): p. 456-63.
207. Force, C.o.A., D.o.D.A.P. Needs, and N.R. Council, *A Review of United States Air Force and Department of Defense Aerospace Propulsion Needs*. 2006: The National Academies Press.
208. Hajipour, M.J., et al., *Antibacterial properties of nanoparticles.* Trends Biotechnol, 2012.
209. Roy Choudhury, S. and A. Goswami, *Supramolecular reactive sulphur nanoparticles: a novel and efficient antimicrobial agent.* J Appl Microbiol, 2012.
210. Raj, S., et al., *Nanotechnology in cosmetics: Opportunities and challenges.* J Pharm Bioallied Sci, 2012. **4**(3): p. 186-93.
211. Tavangarian, F. and Y. Li, *Carbon nanostructures as nerve scaffolds for repairing large gaps in severed nerves.* Ceramics International, 2012. **38**(8): p. 6075-6090.
212. Tiwari, M., *Nano cancer therapy strategies.* J Cancer Res Ther, 2012. **8**(1): p. 19-22.
213. Farokhzad, O.C., et al., *Nanoparticle-aptamer bioconjugates: a new approach for targeting prostate cancer cells.* Cancer Res, 2004. **64**(21): p. 7668-72.
214. Gore, A.Y. and G.S. Banker, *Surface chemistry of colloidal silica and a possible application to stabilize aspirin in solid matrixes.* J Pharm Sci, 1979. **68**(2): p. 197-202.

215. Jonat, S., et al., *Investigation of compacted hydrophilic and hydrophobic colloidal silicon dioxides as glidants for pharmaceutical excipients.* Powder Technology, 2004. **141**(1–2): p. 31-43.
216. Bharali, D.J., et al., *Organically modified silica nanoparticles: a nonviral vector for in vivo gene delivery and expression in the brain.* Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11539-44.
217. Bitar, A., et al., *Silica-based nanoparticles for biomedical applications.* Drug Discov Today, 2012.
218. Hellstrand, E., et al., *Complete high-density lipoproteins in nanoparticle corona.* FEBS J, 2009. **276**(12): p. 3372-81.
219. Barlow, P.G., et al., *Serum exposed to nanoparticle carbon black displays increased potential to induce macrophage migration.* Toxicol Lett, 2005. **155**(3): p. 397-401.
220. Duffin, R., et al., *Proinflammogenic effects of low-toxicity and metal nanoparticles in vivo and in vitro: highlighting the role of particle surface area and surface reactivity.* Inhal Toxicol, 2007. **19**(10): p. 849-56.
221. Poland, C.A., et al., *Carbon nanotubes introduced into the abdominal cavity of mice show asbestos-like pathogenicity in a pilot study.* Nat Nanotechnol, 2008. **3**(7): p. 423-8.
222. Vamanu, C.I., et al., *Induction of cell death by TiO2 nanoparticles: studies on a human monoblastoid cell line.* Toxicol In Vitro, 2008. **22**(7): p. 1689-96.
223. Jeng, H.A. and J. Swanson, *Toxicity of metal oxide nanoparticles in mammalian cells.* J Environ Sci Health A Tox Hazard Subst Environ Eng, 2006. **41**(12): p. 2699-711.
224. Dey, S., et al., *Interactions between SIRT1 and AP-1 reveal a mechanistic insight into the growth promoting properties of alumina (Al2O3) nanoparticles in mouse skin epithelial cells.* Carcinogenesis, 2008. **29**(10): p. 1920-9.
225. Thibodeau, M.S., et al., *Silica-induced apoptosis in mouse alveolar macrophages is initiated by lysosomal enzyme activity.* Toxicol Sci, 2004. **80**(1): p. 34-48.
226. Kang, S.J., et al., *Titanium dioxide nanoparticles induce apoptosis through the JNK/p38-caspase-8-Bid pathway in phytohemagglutinin-stimulated human lymphocytes.* Biochem Biophys Res Commun, 2009. **386**(4): p. 682-7.
227. Lao, F., et al., *Fullerene nanoparticles selectively enter oxidation-damaged cerebral microvessel endothelial cells and inhibit JNK-related apoptosis.* ACS Nano, 2009. **3**(11): p. 3358-68.
228. Lao, F., et al., *Fullerene derivatives protect endothelial cells against NO-induced damage.* Nanotechnology, 2009. **20**(22): p. 225103.
229. Xiao, G.G., et al., *Use of Proteomics to Demonstrate a Hierarchical Oxidative Stress Response to Diesel Exhaust Particle Chemicals in a Macrophage Cell Line.* Journal of Biological Chemistry, 2003. **278**(50): p. 50781-50790.
230. Unfried, K., et al., *Carbon nanoparticle-induced lung epithelial cell proliferation is mediated by receptor-dependent Akt activation.* Am J Physiol Lung Cell Mol Physiol, 2008. **294**(2): p. L358-67.
231. Arenz, A., et al., *Gene expression modulation in A549 human lung cells in response to combustion-generated nano-sized particles.* Ann N Y Acad Sci, 2006. **1091**: p. 170-83.
232. Macarron, R., *Critical review of the role of HTS in drug discovery.* Drug Discov Today, 2006. **11**(7-8): p. 277-9.
233. Grimm, S., *The art and design of genetic screens: mammalian culture cells.* Nat Rev Genet, 2004. **5**(3): p. 179-89.
234. Mishra, K.P., et al., *A review of high throughput technology for the screening of natural products.* Biomed Pharmacother, 2008. **62**(2): p. 94-8.
235. Mayr, L.M. and P. Fuerst, *The future of high-throughput screening.* J Biomol Screen, 2008. **13**(6): p. 443-8.

236. Schreiber, S.L., K.C. Nicolaou, and K. Davies, *Diversity-oriented organic synthesis and proteomics. New frontiers for chemistry & biology.* Chem Biol, 2002. **9**(1): p. 1-2.
237. Ota, T., et al., *Complete sequencing and characterization of 21,243 full-length human cDNAs.* Nat Genet, 2004. **36**(1): p. 40-5.
238. Manetta, J.V., et al., *Design and implementation of a particle concentration fluorescence method for the detection of HIV-1 protease inhibitors.* Anal Biochem, 1992. **202**(1): p. 10-5.
239. Okun, I., et al., *Screening for caspase-3 inhibitors: effect of a reducing agent on identified hit chemotypes.* J Biomol Screen, 2006. **11**(6): p. 694-703.
240. Voigt, W., *Sulforhodamine B assay and chemosensitivity.* Methods Mol Med, 2005. **110**: p. 39-48.
241. Smale, S.T., *Chloramphenicol acetyltransferase assay.* Cold Spring Harb Protoc, 2010. **2010**(5): p. pdb prot5422.
242. Zon, L.I. and R.T. Peterson, *In vivo drug discovery in the zebrafish.* Nat Rev Drug Discov, 2005. **4**(1): p. 35-44.
243. Ruike, Y., et al., *Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing.* BMC Genomics, 2010. **11**: p. 137.
244. Berger, M.F., et al., *Melanoma genome sequencing reveals frequent PREX2 mutations.* Nature, 2012. **485**(7399): p. 502-6.
245. Nikolaev, S.I., et al., *Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma.* Nat Genet, 2012. **44**(2): p. 133-9.
246. Barbieri, C.E., et al., *Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer.* Nat Genet, 2012. **44**(6): p. 685-9.
247. Guo, G., et al., *Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma.* Nat Genet, 2012. **44**(1): p. 17-9.
248. Araya, C.L., et al., *Whole-genome sequencing of a laboratory-evolved yeast strain.* BMC Genomics, 2010. **11**: p. 88.
249. Harris, D.R., et al., *Directed evolution of ionizing radiation resistance in Escherichia coli.* J Bacteriol, 2009. **191**(16): p. 5240-52.
250. Vaheri, A. and J.S. Pagano, *Infectious poliovirus RNA: a sensitive method of assay.* Virology, 1965. **27**(3): p. 434-6.
251. Felgner, P.L., et al., *Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure.* Proc Natl Acad Sci U S A, 1987. **84**(21): p. 7413-7.
252. Kievit, F.M., et al., *PEI-PEG-Chitosan Copolymer Coated Iron Oxide Nanoparticles for Safe Gene Delivery: synthesis, complexation, and transfection.* Adv Funct Mater, 2009. **19**(14): p. 2244-2251.
253. Tarunina, M., et al., *Functional genetic screen for genes involved in senescence: role of Tid1, a homologue of the Drosophila tumor suppressor l(2)tid, in senescence and cell survival.* Mol Cell Biol, 2004. **24**(24): p. 10792-801.
254. Wan, D., et al., *Large-scale cDNA transfection screening for genes related to cancer development and progression.* Proc Natl Acad Sci U S A, 2004. **101**(44): p. 15724-9.
255. Whitehurst, A.W., et al., *Synthetic lethal screen identification of chemosensitizer loci in cancer cells.* Nature, 2007. **446**(7137): p. 815-9.
256. Cerone, M.A., et al., *High-throughput RNAi screening reveals novel regulators of telomerase.* Cancer Res, 2011. **71**(9): p. 3328-40.
257. Kaneda, R., et al., *High-throughput screening of genome fragments bound to differentially acetylated histones.* Genes Cells, 2004. **9**(12): p. 1167-74.
258. Lovejoy, C.A., et al., *Functional genomic screens identify CINP as a genome maintenance protein.* Proc Natl Acad Sci U S A, 2009. **106**(46): p. 19304-9.

259. Zitzler, J., et al., *High-throughput functional genomics identifies genes that ameliorate toxicity due to oxidative stress in neuronal HT-22 cells: GFPT2 protects cells against peroxide.* Mol Cell Proteomics, 2004. **3**(8): p. 834-40.
260. Kaelin, W.G., Jr., *The concept of synthetic lethality in the context of anticancer therapy.* Nat Rev Cancer, 2005. **5**(9): p. 689-98.
261. Kenwrick, S., E. Amaya, and N. Papalopulu, *Pilot morpholino screen in Xenopus tropicalis identifies a novel gene involved in head development.* Dev Dyn, 2004. **229**(2): p. 289-99.
262. Yamada, L., et al., *Morpholino-based gene knockdown screen of novel genes with developmental function in Ciona intestinalis.* Development, 2003. **130**(26): p. 6485-95.
263. Stern, H.M. and L.I. Zon, *Cancer genetics and drug discovery in the zebrafish.* Nat Rev Cancer, 2003. **3**(7): p. 533-9.
264. Amatruda, J.F., et al., *Zebrafish as a cancer model system.* Cancer Cell, 2002. **1**(3): p. 229-31.
265. Langenau, D.M., et al., *Myc-induced T cell leukemia in transgenic zebrafish.* Science, 2003. **299**(5608): p. 887-90.
266. Kalev-Zylinska, M.L., et al., *Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis.* Development, 2002. **129**(8): p. 2015-30.
267. Li, L. and J.E. Dowling, *A dominant form of inherited retinal degeneration caused by a non-photoreceptor cell-specific mutation.* Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11645-50.
268. Albayrak, T. and S. Grimm, *A high-throughput screen for single gene activities: isolation of apoptosis inducers.* Biochem Biophys Res Commun, 2003. **304**(4): p. 772-6.
269. Park, K.M., et al., *Identification of novel regulators of apoptosis using a high-throughput cell-based screen.* Mol Cells, 2007. **23**(2): p. 170-4.
270. Ziauddin, J. and D.M. Sabatini, *Microarrays of cells expressing defined cDNAs.* Nature, 2001. **411**(6833): p. 107-10.
271. Palmer, E.L., A.D. Miller, and T.C. Freeman, *Identification and characterisation of human apoptosis inducing proteins using cell-based transfection microarrays and expression analysis.* BMC Genomics, 2006. **7**: p. 145.
272. Mannherz, O., et al., *Functional screening for proapoptotic genes by reverse transfection cell array technology.* Genomics, 2006. **87**(5): p. 665-72.
273. Johnson, P.H., et al., *Multiplex gene expression analysis for high-throughput drug discovery: screening and analysis of compounds affecting genes overexpressed in cancer cells.* Mol Cancer Ther, 2002. **1**(14): p. 1293-304.
274. Tian, H., et al., *A high throughput drug screen based on fluorescence resonance energy transfer (FRET) for anticancer activity of compounds from herbal medicine.* Br J Pharmacol, 2007. **150**(3): p. 321-34.
275. Gupta, P.B., et al., *Identification of selective inhibitors of cancer stem cells by high-throughput screening.* Cell, 2009. **138**(4): p. 645-59.
276. Chew, S.K., et al., *Genome-wide silencing in Drosophila captures conserved apoptotic effectors.* Nature, 2009. **460**(7251): p. 123-7.
277. *Entering the golden age of lab automation.* Biotechniques, 2010. **48**(1): p. 11.
278. Beall, A., *Microbiology lab automation arrives.* MLO Med Lab Obs, 2008. **40**(7): p. 48, 50.
279. Chapman, T., *Lab automation and robotics: Automation on the move.* Nature, 2003. **421**(6923): p. 661, 663, 665-6.
280. Blow, N., *Lab automation: tales along the road to automation.* Nat Methods, 2008. **5**(1): p. 109-12.
281. Burt, S.M., T.J. Carter, and L.J. Kricka, *Thermal characteristics of microtitre plates used in immunological assays.* J Immunol Methods, 1979. **31**(3-4): p. 231-6.
282. Lundholt, B.K., K.M. Scudder, and L. Pagliaro, *A simple technique for reducing edge effect in cell-based assays.* J Biomol Screen, 2003. **8**(5): p. 566-70.

283. Irshad, S., et al., *Isolation of ORCTL3 in a novel genetic screen for tumor-specific apoptosis inducers.* Cell Death Differ, 2009. **16**(6): p. 890-8.
284. Lin, B., et al., *Determining signalling nodes for apoptosis by a genetic high-throughput screen.* PLoS One, 2011. **6**(9): p. e25023.
285. Grimm, S. and V. Kachel, *Robotic high-throughput assay for isolating apoptosis-inducing genes.* Biotechniques, 2002. **32**(3): p. 670-2, 674-7.
286. Kachel, V., G. Sindelar, and S. Grimm, *High-throughput isolation of ultra-pure plasmid DNA by a robotic system.* BMC Biotechnol, 2006. **6**: p. 9.
287. Maruyama, K. and S. Sugano, *Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides.* Gene, 1994. **138**(1-2): p. 171-4.
288. Suzuki, Y., et al., *Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library.* Gene, 1997. **200**(1-2): p. 149-56.
289. Takebe, Y., et al., *SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat.* Mol Cell Biol, 1988. **8**(1): p. 466-72.
290. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
291. Chung, C.T., S.L. Niemela, and R.H. Miller, *One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution.* Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2172-5.
292. van Die, I.M., H.E. Bergmans, and W.P. Hoekstra, *Transformation in Escherichia coli: studies on the role of the heat shock in induction of competence.* J Gen Microbiol, 1983. **129**(3): p. 663-70.
293. Ohnuma, K., et al., *Sorting of cells of the same size, shape, and cell cycle stage for a single cell level assay without staining.* BMC Cell Biol, 2006. **7**: p. 25.
294. Raymond, C., et al., *A simplified polyethylenimine-mediated transfection process for large-scale and high-throughput applications.* Methods, 2011. **55**(1): p. 44-51.
295. Ito, T., N. Iida-Tanaka, and Y. Koyama, *Efficient in vivo gene transfection by stable DNA/PEI complexes coated by hyaluronic acid.* J Drug Target, 2008. **16**(4): p. 276-81.
296. Miller, O.J., et al., *High-resolution dose-response screening using droplet-based microfluidics.* Proc Natl Acad Sci U S A, 2012. **109**(2): p. 378-83.
297. Kim, J.S., et al., *Antimicrobial effects of silver nanoparticles.* Nanomedicine, 2007. **3**(1): p. 95-101.
298. Weir, A., et al., *Titanium dioxide nanoparticles in food and personal care products.* Environ Sci Technol, 2012. **46**(4): p. 2242-50.
299. He, Q., et al., *An anticancer drug delivery system based on surfactant-templated mesoporous silica nanoparticles.* Biomaterials, 2010. **31**(12): p. 3335-46.
300. Mamaeva, V., et al., *Mesoporous silica nanoparticles as drug delivery systems for targeted inhibition of Notch signaling in cancer.* Mol Ther, 2011. **19**(8): p. 1538-46.
301. Shaw, T., et al., *True photoallergy to sunscreens is rare despite popular belief.* Dermatitis, 2010. **21**(4): p. 185-98.
302. Shavandi, Z., T. Ghazanfari, and K.N. Moghaddam, *In vitro toxicity of silver nanoparticles on murine peritoneal macrophages.* Immunopharmacol Immunotoxicol, 2011. **33**(1): p. 135-40.
303. Haase, A., et al., *Toxicity of silver nanoparticles in human macrophages: uptake, intracellular distribution and cellular responses.* Journal of Physics: Conference Series, 2011. **304**(1): p. 012030.
304. Ghosh, M., et al., *In vitro and in vivo genotoxicity of silver nanoparticles.* Mutat Res, 2012.
305. Kim, S. and D.Y. Ryu, *Silver nanoparticle-induced oxidative stress, genotoxicity and apoptosis in cultured cells and animal tissues.* J Appl Toxicol, 2012.
306. Huang, X. and A. Madan, *CAP3: A DNA sequence assembly program.* Genome Res, 1999. **9**(9): p. 868-77.

307. Hubbard, T.J., et al., *Ensembl 2009.* Nucleic Acids Res, 2009. **37**(Database issue): p. D690-7.
308. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
309. Zdobnov, E.M. and R. Apweiler, *InterProScan--an integration platform for the signature-recognition methods in InterPro.* Bioinformatics, 2001. **17**(9): p. 847-8.
310. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
311. Mi, H., et al., *PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium.* Nucleic Acids Res, 2010. **38**(Database issue): p. D204-10.
312. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Res, 2003. **13**(9): p. 2129-41.
313. Wee, L.J., T.W. Tan, and S. Ranganathan, *CASVM: web server for SVM-based prediction of caspase substrates cleavage sites.* Bioinformatics, 2007. **23**(23): p. 3241-3.
314. Lamarca, V., et al., *Exposure of any of two proapoptotic domains of presenilin 1-associated protein/mitochondrial carrier homolog 1 on the surface of mitochondria is sufficient for induction of apoptosis in a Bax/Bak-independent manner.* Eur J Cell Biol, 2008. **87**(5): p. 325-34.
315. Mao, G., et al., *Both the N-terminal fragment and the protein-protein interaction domain (PDZ domain) are required for the pro-apoptotic activity of presenilin-associated protein PSAP.* Biochim Biophys Acta, 2008. **1780**(4): p. 696-708.
316. Xu, X., et al., *The novel presenilin-1-associated protein is a proapoptotic mitochondrial protein.* J Biol Chem, 2002. **277**(50): p. 48913-22.
317. Fleischer, A. and A. Rebollo, *Induction of p53-independent apoptosis by the BH3-only protein ITM2Bs.* FEBS Lett, 2004. **557**(1-3): p. 283-7.
318. Fleischer, A., V. Ayllon, and A. Rebollo, *ITM2BS regulates apoptosis by inducing loss of mitochondrial membrane potential.* Eur J Immunol, 2002. **32**(12): p. 3498-505.
319. Lv, B.F., et al., *Protein tyrosine phosphatase interacting protein 51 (PTPIP51) is a novel mitochondria protein with an N-terminal mitochondrial targeting sequence and induces apoptosis.* Apoptosis, 2006. **11**(9): p. 1489-501.
320. Stenzinger, A., et al., *Cell and molecular biology of the novel protein tyrosine-phosphatase-interacting protein 51.* Int Rev Cell Mol Biol, 2009. **275**: p. 183-246.
321. De Vos, K.J., et al., *VAPB interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis.* Hum Mol Genet, 2012. **21**(6): p. 1299-311.
322. Kirkegaard, T. and M. Jaattela, *Lysosomal involvement in cell death and cancer.* Biochim Biophys Acta, 2009. **1793**(4): p. 746-54.
323. Fu, Y., H. Sies, and X.G. Lei, *Opposite roles of selenium-dependent glutathione peroxidase-1 in superoxide generator diquat- and peroxynitrite-induced apoptosis and signaling.* J Biol Chem, 2001. **276**(46): p. 43004-9.
324. Prendergast, G.C., *Actin' up: RhoB in cancer and apoptosis.* Nat Rev Cancer, 2001. **1**(2): p. 162-8.
325. Suen, D.F., K.L. Norris, and R.J. Youle, *Mitochondrial dynamics and apoptosis.* Genes Dev, 2008. **22**(12): p. 1577-90.
326. Ferri, K.F. and G. Kroemer, *Organelle-specific initiation of cell death pathways.* Nat Cell Biol, 2001. **3**(11): p. E255-63.
327. Bratton, S.B. and G.M. Cohen, *Apoptotic death sensor: an organelle's alter ego?* Trends Pharmacol Sci, 2001. **22**(6): p. 306-15.
328. Dodonov, Y.S.D., Y. A. , *Robust measures of central Tendency: weighting as a possible alternative to trimming in response-time data analysis. .* Psikhologicheskie Issledovaniya, 2011. **5**(19).

329. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
330. Petretto, E., et al., *Heritability and tissue specificity of expression quantitative trait loci.* PLoS Genet, 2006. **2**(10): p. e172.
331. Hornik, K., *The R FAQ.* http://CRAN.R-project.org/doc/FAQ/R-FAQ.html, 2012.
332. JD, S., *A direct approach to false discovery rates.* Journal of the Royal Statistical Society, Series B, 2002(64): p. 479-498.
333. Evan, G.I., et al., *Induction of apoptosis in fibroblasts by c-myc protein.* Cell, 1992. **69**(1): p. 119-28.
334. Cohen, G.M., *Caspases: the executioners of apoptosis.* Biochem J, 1997. **326 (Pt 1)**: p. 1-16.
335. Mudunuri, U., et al., *bioDBnet: the biological database network.* Bioinformatics, 2009. **25**(4): p. 555-6.
336. Huang da, W., et al., *Extracting biological meaning from large gene lists with DAVID.* Curr Protoc Bioinformatics, 2009. **Chapter 13**: p. Unit 13 11.
337. Barber, G.N., *Host defense, viruses and apoptosis.* Cell Death Differ, 2001. **8**(2): p. 113-26.
338. Stohs, S.J. and D. Bagchi, *Oxidative mechanisms in the toxicity of metal ions.* Free Radic Biol Med, 1995. **18**(2): p. 321-36.
339. Egorova, K.S., O.M. Olenkina, and L.V. Olenina, *Lysine methylation of nonhistone proteins is a way to regulate their stability and function.* Biochemistry (Mosc), 2010. **75**(5): p. 535-48.
340. Stallcup, M.R., *Role of protein methylation in chromatin remodeling and transcriptional regulation.* Oncogene, 2001. **20**(24): p. 3014-20.
341. Xie, W. and R.B. Denman, *Protein methylation and stress granules: posttranslational remodeler or innocent bystander?* Mol Biol Int, 2011. **2011**: p. 137459.
342. Sarkar, S. and M. Mandal, *Growth factor receptors and apoptosis regulators: signaling pathways, prognosis, chemosensitivity and treatment outcomes of breast cancer.* Breast Cancer (Auckl), 2009. **3**: p. 47-60.
343. Hattori, Y., et al., *A selective estrogen receptor modulator inhibits tumor necrosis factor-alpha-induced apoptosis through the ERK1/2 signaling pathway in human chondrocytes.* Biochem Biophys Res Commun, 2012. **421**(3): p. 418-24.
344. Razandi, M., A. Pedram, and E.R. Levin, *Plasma membrane estrogen receptors signal to antiapoptosis in breast cancer.* Mol Endocrinol, 2000. **14**(9): p. 1434-47.
345. Stevens, F.C., *Calmodulin: an introduction.* Can J Biochem Cell Biol, 1983. **61**(8): p. 906-10.
346. Chin, D. and A.R. Means, *Calmodulin: a prototypical calcium sensor.* Trends Cell Biol, 2000. **10**(8): p. 322-8.
347. Olofsson, M.H., et al., *Charting calcium-regulated apoptosis pathways using chemical biology: role of calmodulin kinase II.* BMC Chem Biol, 2008. **8**: p. 2.
348. Bader, G.D., D. Betel, and C.W. Hogue, *BIND: the Biomolecular Interaction Network Database.* Nucleic Acids Res, 2003. **31**(1): p. 248-50.
349. Chatr-aryamontri, A., et al., *MINT: the Molecular INTeraction database.* Nucleic Acids Res, 2007. **35**(Database issue): p. D572-4.
350. Licata, L., et al., *MINT, the molecular interaction database: 2012 update.* Nucleic Acids Res, 2012. **40**(Database issue): p. D857-61.
351. Croft, D., et al., *Reactome: a database of reactions, pathways and biological processes.* Nucleic Acids Res, 2011. **39**(Database issue): p. D691-7.
352. Bos, J.L., *ras oncogenes in human cancer: a review.* Cancer Res, 1989. **49**(17): p. 4682-9.
353. Kandpal, R.P., *Rho GTPase activating proteins in cancer phenotypes.* Curr Protein Pept Sci, 2006. **7**(4): p. 355-65.
354. Gilmore, T.D., *The Re1/NF-kappa B/I kappa B signal transduction pathway and cancer.* Cancer Treat Res, 2003. **115**: p. 241-65.
355. Hankins, G.R., et al., *Identification of the deleted in liver cancer 1 gene, DLC1, as a candidate meningioma tumor suppressor.* Neurosurgery, 2008. **63**(4): p. 771-80; discussion 780-1.

356. Jonson, L., et al., *Molecular composition of IMP1 ribonucleoprotein granules.* Mol Cell Proteomics, 2007. **6**(5): p. 798-811.
357. Colombo, E., et al., *Nucleophosmin regulates the stability and transcriptional activity of p53.* Nat Cell Biol, 2002. **4**(7): p. 529-33.
358. Falini, B., et al., *Translocations and mutations involving the nucleophosmin (NPM1) gene in lymphomas and leukemias.* Haematologica, 2007. **92**(4): p. 519-32.
359. Grisendi, S., et al., *Nucleophosmin and cancer.* Nat Rev Cancer, 2006. **6**(7): p. 493-505.
360. Jurica, M.S., et al., *Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis.* RNA, 2002. **8**(4): p. 426-39.
361. Hunter, S., et al., *InterPro: the integrative protein signature database.* Nucleic Acids Res, 2009. **37**(Database issue): p. D211-5.
362. Letunic, I., T. Doerks, and P. Bork, *SMART 7: recent updates to the protein domain annotation resource.* Nucleic Acids Res, 2012. **40**(Database issue): p. D302-5.
363. Schultz, J., et al., *SMART: a web-based tool for the study of genetically mobile domains.* Nucleic Acids Res, 2000. **28**(1): p. 231-4.
364. Wu, C.H., et al., *The Protein Information Resource.* Nucleic Acids Res, 2003. **31**(1): p. 345-7.
365. Walmsley, A.R., et al., *Sugar transporters from bacteria, parasites and mammals: structure-activity relationships.* Trends Biochem Sci, 1998. **23**(12): p. 476-81.
366. Freemont, P.S., *The RING finger. A novel protein sequence motif related to the zinc finger.* Ann N Y Acad Sci, 1993. **684**: p. 174-92.
367. Joazeiro, C.A. and A.M. Weissman, *RING finger proteins: mediators of ubiquitin ligase activity.* Cell, 2000. **102**(5): p. 549-52.
368. Teichmann, S.A. and C. Chothia, *Immunoglobulin superfamily proteins in Caenorhabditis elegans.* J Mol Biol, 2000. **296**(5): p. 1367-83.
369. He, Q.Y., et al., *G8: a novel domain associated with polycystic kidney disease and non-syndromic hearing loss.* Bioinformatics, 2006. **22**(18): p. 2189-91.
370. Li, D. and R. Roberts, *WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases.* Cell Mol Life Sci, 2001. **58**(14): p. 2085-97.
371. Stec, I., et al., *The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation?* FEBS Lett, 2000. **473**(1): p. 1-5.
372. Ponting, C.P., *Tudor domains in proteins that interact with RNA.* Trends Biochem Sci, 1997. **22**(2): p. 51-2.
373. Mallery, D.L., et al., *Antibodies mediate intracellular immunity through tripartite motif-containing 21 (TRIM21).* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 19985-90.
374. Sabe, H., et al., *ArfGAP family proteins in cell adhesion, migration and tumor invasion.* Curr Opin Cell Biol, 2006. **18**(5): p. 558-64.
375. Han, W. and P. Christen, *cis-Effect of DnaJ on DnaK in ternary complexes with chimeric DnaK/DnaJ-binding peptides.* FEBS Lett, 2004. **563**(1-3): p. 146-50.
376. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nat Genet, 2003. **34**(3): p. 267-73.
377. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
378. Gomez del Arco, P., et al., *A role for the p38 MAP kinase pathway in the nuclear shuttling of NFATp.* J Biol Chem, 2000. **275**(18): p. 13872-8.
379. Olson, E.N. and J.D. Molkentin, *Prevention of cardiac hypertrophy by calcineurin inhibition: hope or hype?* Circ Res, 1999. **84**(6): p. 623-32.
380. Cadwell, K., et al., *A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells.* Nature, 2008. **456**(7219): p. 259-63.

381. Gauldie, J., et al., *Interferon beta 2/B-cell stimulatory factor type 2 shares identity with monocyte-derived hepatocyte-stimulating factor and regulates the major acute phase protein response in liver cells.* Proc Natl Acad Sci U S A, 1987. **84**(20): p. 7251-5.
382. Biffl, W.L., et al., *Interleukin-6 suppression of neutrophil apoptosis is neutrophil concentration dependent.* J Leukoc Biol, 1995. **58**(5): p. 582-4.
383. Regis, G., et al., *IL-6, but not IFN-gamma, triggers apoptosis and inhibits in vivo growth of human malignant T cells on STAT3 silencing.* Leukemia, 2009. **23**(11): p. 2102-8.
384. Bild, A.H., et al., *Oncogenic pathway signatures in human cancers as a guide to targeted therapies.* Nature, 2006. **439**(7074): p. 353-7.
385. Vogt, P.K. and T.J. Bos, *jun: oncogene and transcription factor.* Adv Cancer Res, 1990. **55**: p. 1-35.
386. Moreau-Gachelin, F., et al., *Spi-1 oncogene activation in Rauscher and Friend murine virus-induced acute erythroleukemias.* Leukemia, 1990. **4**(1): p. 20-3.
387. Warde-Farley, D., et al., *The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W214-20.
388. Smoot, M.E., et al., *Cytoscape 2.8: new features for data integration and network visualization.* Bioinformatics, 2011. **27**(3): p. 431-2.
389. Michell, R.H., et al., *The stimulation of inositol lipid metabolism that accompanies calcium mobilization in stimulated cells: defined characteristics and unanswered questions.* Philos Trans R Soc Lond B Biol Sci, 1981. **296**(1080): p. 123-38.
390. Suzuki, K., et al., *Identification and characterization of a novel human phosphatidylinositol 4-kinase.* DNA Res, 1997. **4**(4): p. 273-80.
391. Heilmeyer, L.M., Jr., et al., *Mammalian phosphatidylinositol 4-kinases.* IUBMB Life, 2003. **55**(2): p. 59-65.
392. Nakashima, S., *Protein kinase C alpha (PKC alpha): regulation and biological function.* J Biochem, 2002. **132**(5): p. 669-75.
393. Myers, M.P., et al., *P-TEN, the tumor suppressor from human chromosome 10q23, is a dual-specificity phosphatase.* Proc Natl Acad Sci U S A, 1997. **94**(17): p. 9052-7.
394. Maehama, T. and J.E. Dixon, *The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate.* J Biol Chem, 1998. **273**(22): p. 13375-8.
395. Myers, M.P., et al., *The lipid phosphatase activity of PTEN is critical for its tumor supressor function.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13513-8.
396. Georgescu, M.M., et al., *The tumor-suppressor activity of PTEN is regulated by its carboxyl-terminal region.* Proc Natl Acad Sci U S A, 1999. **96**(18): p. 10182-7.
397. Wang, X., et al., *NEDD4-1 is a proto-oncogenic ubiquitin ligase for PTEN.* Cell, 2007. **128**(1): p. 129-39.
398. Van Themsche, C., et al., *X-linked inhibitor of apoptosis protein (XIAP) regulates PTEN ubiquitination, content, and compartmentalization.* J Biol Chem, 2009. **284**(31): p. 20462-6.
399. Schaap, D., et al., *Purification, cDNA-cloning and expression of human diacylglycerol kinase.* FEBS Lett, 1990. **275**(1-2): p. 151-8.
400. Lykidis, A., et al., *The role of CDP-diacylglycerol synthetase and phosphatidylinositol synthase activity levels in the regulation of cellular phosphatidylinositol content.* J Biol Chem, 1997. **272**(52): p. 33402-9.
401. Tan, K.M., et al., *The Caenorhabditis elegans sex-determining protein FEM-2 and its human homologue, hFEM-2, are Ca2+/calmodulin-dependent protein kinase phosphatases that promote apoptosis.* J Biol Chem, 2001. **276**(47): p. 44193-202.
402. Krishnan, N., et al., *H2S-Induced sulfhydration of the phosphatase PTP1B and its role in the endoplasmic reticulum stress response.* Sci Signal, 2011. **4**(203): p. ra86.

403. Walter, P. and D. Ron, *The unfolded protein response: from stress pathway to homeostatic regulation.* Science, 2011. **334**(6059): p. 1081-6.
404. Graham, K. and M.F. Olson, *The ras signalling pathway as a target in cancer therapy.* Recent Results Cancer Res, 2007. **172**: p. 125-53.
405. Yap, T.A., et al., *Targeting the PI3K-AKT-mTOR pathway: progress, pitfalls, and promises.* Curr Opin Pharmacol, 2008. **8**(4): p. 393-412.
406. Yu, Y., et al., *NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas.* Proc Natl Acad Sci U S A, 1999. **96**(1): p. 214-9.
407. Rodriguez-Viciana, P., et al., *A phosphatase holoenzyme comprised of Shoc2/Sur8 and the catalytic subunit of PP1 functions as an M-Ras effector to modulate Raf activity.* Mol Cell, 2006. **22**(2): p. 217-30.
408. Scott, J.W., et al., *CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations.* J Clin Invest, 2004. **113**(2): p. 274-84.
409. Maekawa, M., et al., *Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase.* Science, 1999. **285**(5429): p. 895-8.
410. Matsuda, A., et al., *Large-scale identification and characterization of human genes that activate NF-kappaB and MAPK signaling pathways.* Oncogene, 2003. **22**(21): p. 3307-18.
411. Sheikh, M.S. and Y. Huang, *Death receptor activation complexes: it takes two to activate TNF receptor 1.* Cell Cycle, 2003. **2**(6): p. 550-2.
412. Calleros, L., et al., *RhoA and p38 MAPK mediate apoptosis induced by cellular cholesterol depletion.* Apoptosis, 2006. **11**(7): p. 1161-73.
413. Neisch, A.L., et al., *Rho1 regulates apoptosis via activation of the JNK signaling pathway at the plasma membrane.* J Cell Biol, 2010. **189**(2): p. 311-23.
414. Fransson, A., A. Ruusala, and P. Aspenstrom, *Atypical Rho GTPases have roles in mitochondrial homeostasis and apoptosis.* J Biol Chem, 2003. **278**(8): p. 6495-502.
415. Fransson, S., A. Ruusala, and P. Aspenstrom, *The atypical Rho GTPases Miro-1 and Miro-2 have essential roles in mitochondrial trafficking.* Biochem Biophys Res Commun, 2006. **344**(2): p. 500-10.
416. Naora, H., et al., *Altered cellular responses by varying expression of a ribosomal protein gene: sequential coordination of enhancement and suppression of ribosomal protein S3a gene expression induces apoptosis.* J Cell Biol, 1998. **141**(3): p. 741-53.
417. Dalby, K.N., et al., *Identification of regulatory phosphorylation sites in mitogen-activated protein kinase (MAPK)-activated protein kinase-1a/p90rsk that are inducible by MAPK.* J Biol Chem, 1998. **273**(3): p. 1496-505.
418. Buck, M., et al., *C/EBPbeta phosphorylation by RSK creates a functional XEXD caspase inhibitory box critical for cell survival.* Mol Cell, 2001. **8**(4): p. 807-16.
419. Anjum, R., et al., *The tumor suppressor DAP kinase is a target of RSK-mediated survival signaling.* Curr Biol, 2005. **15**(19): p. 1762-7.
420. Nishimoto, S. and E. Nishida, *MAPK signalling: ERK5 versus ERK1/2.* EMBO Rep, 2006. **7**(8): p. 782-6.
421. Steinmetz, M.O., et al., *Op18/stathmin caps a kinked protofilament-like tubulin tetramer.* EMBO J, 2000. **19**(4): p. 572-80.
422. Redeker, V., et al., *Probing the native structure of stathmin and its interaction domains with tubulin. Combined use of limited proteolysis, size exclusion chromatography, and mass spectrometry.* J Biol Chem, 2000. **275**(10): p. 6841-9.
423. Dalton, S. and R. Treisman, *Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element.* Cell, 1994. **76**(2): p. 411.
424. Shipley, J., et al., *Mapping of the human SAP1 (SRF accessory protein 1) gene and SAP2, a gene encoding a related protein, to chromosomal bands 1q32 and 12q23, respectively.* Genomics, 1994. **23**(3): p. 710-1.

425. Miura, K., et al., *A case of Kallmann syndrome carrying a missense mutation in alternatively spliced exon 8A encoding the immunoglobulin-like domain IIIb of fibroblast growth factor receptor 1.* Hum Reprod, 2010. **25**(4): p. 1076-80.
426. Cross, M.J., et al., *The Shb adaptor protein binds to tyrosine 766 in the FGFR-1 and regulates the Ras/MEK/MAPK pathway via FRS2 phosphorylation in endothelial cells.* Mol Biol Cell, 2002. **13**(8): p. 2881-93.
427. Ballester, R., et al., *The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins.* Cell, 1990. **63**(4): p. 851-9.
428. Larsson, P.K., H.E. Claesson, and B.P. Kennedy, *Multiple splice variants of the human calcium-independent phospholipase A2 and their effect on enzyme activity.* J Biol Chem, 1998. **273**(1): p. 207-14.
429. Malhotra, A., et al., *Role of calcium-independent phospholipase A2 in the pathogenesis of Barth syndrome.* Proc Natl Acad Sci U S A, 2009. **106**(7): p. 2337-41.
430. Shapira, S., et al., *The tumor suppressor neurofibromin confers sensitivity to apoptosis by Ras-dependent and Ras-independent pathways.* Cell Death Differ, 2007. **14**(5): p. 895-906.
431. Balsinde, J., R. Perez, and M.A. Balboa, *Calcium-independent phospholipase A2 and apoptosis.* Biochim Biophys Acta, 2006. **1761**(11): p. 1344-50.
432. Hediger, M.A., et al., *The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteinsIntroduction.* Pflugers Arch, 2004. **447**(5): p. 465-8.
433. Huang, Y. and W. Sadee, *Membrane transporters and channels in chemoresistance and -sensitivity of tumor cells.* Cancer Lett, 2006. **239**(2): p. 168-82.
434. Crabtree, G.R. and E.N. Olson, *NFAT signaling: choreographing the social lives of cells.* Cell, 2002. **109 Suppl**: p. S67-79.
435. Fiedler, B., et al., *Inhibition of calcineurin-NFAT hypertrophy signaling by cGMP-dependent protein kinase type I in cardiac myocytes.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11363-8.
436. Johnson, E.N., et al., *NFATc1 mediates vascular endothelial growth factor-induced proliferation of human pulmonary valve endothelial cells.* J Biol Chem, 2003. **278**(3): p. 1686-92.
437. Flockhart, R.J., et al., *NFAT regulates induction of COX-2 and apoptosis of keratinocytes in response to ultraviolet radiation exposure.* FASEB J, 2008. **22**(12): p. 4218-27.
438. Jayanthi, S., et al., *Calcineurin/NFAT-induced up-regulation of the Fas ligand/Fas death pathway is involved in methamphetamine-induced neuronal apoptosis.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 868-73.
439. Berthold, J., K. Schenkova, and F. Rivero, *Rho GTPases of the RhoBTB subfamily and tumorigenesis.* Acta Pharmacol Sin, 2008. **29**(3): p. 285-95.
440. Kawai, T., H. Sanjo, and S. Akira, *Duet is a novel serine/threonine kinase with Dbl-Homology (DH) and Pleckstrin-Homology (PH) domains.* Gene, 1999. **227**(2): p. 249-55.
441. Hiramoto-Yamaki, N., et al., *Ephexin4 and EphA2 mediate cell migration through a RhoG-dependent mechanism.* J Cell Biol, 2010. **190**(3): p. 461-77.
442. Oliver, A.W., et al., *The HPV16 E6 binding protein Tip-1 interacts with ARHGEF16, which activates Cdc42.* Br J Cancer, 2011. **104**(2): p. 324-31.
443. Adams, R.H., et al., *Roles of ephrinB ligands and EphB receptors in cardiovascular development: demarcation of arterial/venous domains, vascular morphogenesis, and sprouting angiogenesis.* Genes Dev, 1999. **13**(3): p. 295-306.
444. Wu, X., et al., *Evidence for regulation of the PTEN tumor suppressor by a membrane-localized multi-PDZ domain containing scaffold protein MAGI-2.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4233-8.
445. Gao, Y., et al., *Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain.* J Biol Chem, 2001. **276**(13): p. 9838-45.

446. Small, D.H., S.S. Mok, and J.C. Bornstein, *Alzheimer's disease and Abeta toxicity: from top to bottom.* Nat Rev Neurosci, 2001. **2**(8): p. 595-8.
447. Annaert, W. and B. De Strooper, *A cell biological perspective on Alzheimer's disease.* Annu Rev Cell Dev Biol, 2002. **18**: p. 25-51.
448. Miranda, L., et al., *Isolation of the human PC6 gene encoding the putative host protease for HIV-1 gp160 processing in CD4+ T lymphocytes.* Proc Natl Acad Sci U S A, 1996. **93**(15): p. 7695-700.
449. Mori, K., et al., *A novel human PACE4 isoform, PACE4E is an active processing protease containing a hydrophobic cluster at the carboxy terminus.* J Biochem, 1997. **121**(5): p. 941-8.
450. Iverson, H.A., et al., *Identification and structural determination of the M(3) muscarinic acetylcholine receptor basolateral sorting signal.* J Biol Chem, 2005. **280**(26): p. 24568-75.
451. Williams, M.E., et al., *Structure and functional expression of alpha 1, alpha 2, and beta subunits of a novel human neuronal calcium channel subtype.* Neuron, 1992. **8**(1): p. 71-84.
452. Ohno, S., et al., *Four genes for the calpain family locate on four distinct human chromosomes.* Cytogenet Cell Genet, 1990. **53**(4): p. 225-9.
453. Benayoun, B., et al., *NF-kappaB-dependent expression of the antiapoptotic factor c-FLIP is regulated by calpain 3, the protein involved in limb-girdle muscular dystrophy type 2A.* FASEB J, 2008. **22**(5): p. 1521-9.
454. Raevskaya, N.M., et al., *Structural organization of the human complexin 2 gene (CPLX2) and aspects of its functional activity.* Gene, 2005. **359**: p. 127-37.
455. DiProspero, N.A., et al., *Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements.* J Neurocytol, 2004. **33**(5): p. 517-33.
456. Tai, G., et al., *Participation of the syntaxin 5/Ykt6/GS28/GS15 SNARE complex in transport from the early/recycling endosome to the trans-Golgi network.* Mol Biol Cell, 2004. **15**(9): p. 4011-22.
457. Kajiwara, Y., et al., *FE65 binds Teashirt, inhibiting expression of the primate-specific caspase-4.* PLoS One, 2009. **4**(4): p. e5071.
458. LeBoeuf, M., et al., *Hdac1 and Hdac2 act redundantly to control p63 and p53 functions in epidermal progenitor cells.* Dev Cell, 2010. **19**(6): p. 807-18.
459. Tavaria, M., et al., *Localization of the gene encoding the human heat shock cognate protein, HSP73, to chromosome 11.* Genomics, 1995. **29**(1): p. 266-8.
460. Yahata, T., et al., *The MSG1 non-DNA-binding transactivator binds to the p300/CBP coactivators, enhancing their functional link to the Smad transcription factors.* J Biol Chem, 2000. **275**(12): p. 8825-34.
461. Takayama, S., et al., *Expression and location of Hsp70/Hsc-binding anti-apoptotic protein BAG-1 and its variants in normal tissues and tumor cell lines.* Cancer Res, 1998. **58**(14): p. 3116-31.
462. Luo, S. and D.C. Rubinsztein, *Huntingtin promotes cell survival by preventing Pak2 cleavage.* J Cell Sci, 2009. **122**(Pt 6): p. 875-85.
463. Volinia, S., et al., *A human phosphatidylinositol 3-kinase complex related to the yeast Vps34p-Vps15p protein sorting system.* EMBO J, 1995. **14**(14): p. 3339-48.
464. Meredith, J.E., Jr. and M.A. Schwartz, *Integrins, adhesion and apoptosis.* Trends Cell Biol, 1997. **7**(4): p. 146-50.
465. Zhang, Y. and B. Zhang, *TRAIL resistance of breast cancer cells is associated with constitutive endocytosis of death receptors 4 and 5.* Mol Cancer Res, 2008. **6**(12): p. 1861-71.
466. Rao, P.K., et al., *Distinct roles for miR-1 and miR-133a in the proliferation and differentiation of rhabdomyosarcoma cells.* FASEB J, 2010. **24**(9): p. 3427-37.
467. Nasser, M.W., et al., *Down-regulation of micro-RNA-1 (miR-1) in lung cancer. Suppression of tumorigenic property of lung cancer cells and their sensitization to doxorubicin-induced apoptosis by miR-1.* J Biol Chem, 2008. **283**(48): p. 33394-405.

468. Chiyomaru, T., et al., *Functional role of LASP1 in cell viability and its regulation by microRNAs in bladder cancer.* Urol Oncol, 2012. **30**(4): p. 434-43.
469. Pitson, S.M., et al., *Human sphingosine kinase: purification, molecular cloning and characterization of the native and recombinant enzymes.* Biochem J, 2000. **350 Pt 2**: p. 429-41.
470. Melville, M.W., et al., *The cellular inhibitor of the PKR protein kinase, P58(IPK), is an influenza virus-activated co-chaperone that modulates heat shock protein 70 activity.* J Biol Chem, 1999. **274**(6): p. 3797-803.
471. Zhu, X.D., et al., *Caveolae-dependent endocytosis is required for class A macrophage scavenger receptor-mediated apoptosis in macrophages.* J Biol Chem, 2011. **286**(10): p. 8231-9.
472. Chitnis, N.S., et al., *Modulation of iridovirus-induced apoptosis by endocytosis, early expression, JNK, and apical caspase.* Virology, 2008. **370**(2): p. 333-42.
473. Sanes, J.R. and J.W. Lichtman, *Induction, assembly, maturation and maintenance of a postsynaptic apparatus.* Nat Rev Neurosci, 2001. **2**(11): p. 791-805.
474. Tse, W.T., et al., *A new spectrin, beta IV, has a major truncated isoform that associates with promyelocytic leukemia protein nuclear bodies and the nuclear matrix.* J Biol Chem, 2001. **276**(26): p. 23974-85.
475. Zhang, D., et al., *Neuregulin-3 (NRG3): a novel neural tissue-enriched protein that binds and activates ErbB4.* Proc Natl Acad Sci U S A, 1997. **94**(18): p. 9562-7.
476. Olayioye, M.A., et al., *ErbB receptor-induced activation of stat transcription factors is mediated by Src tyrosine kinases.* J Biol Chem, 1999. **274**(24): p. 17209-18.
477. Schrantz, N., et al., *Mechanism of p21-activated kinase 6-mediated inhibition of androgen receptor signaling.* J Biol Chem, 2004. **279**(3): p. 1922-31.
478. Zhang, M., et al., *Inhibition of p21-activated kinase 6 (PAK6) increases radiosensitivity of prostate cancer cells.* Prostate, 2010. **70**(8): p. 807-16.
479. Schultz, D., et al., *Cloning, chromosomal localization, and functional expression of the alpha 1 subunit of the L-type voltage-dependent calcium channel from normal human heart.* Proc Natl Acad Sci U S A, 1993. **90**(13): p. 6228-32.
480. Petersen, L.G., et al., *IL-1beta-induced pro-apoptotic signalling is facilitated by NCAM/FGF receptor signalling and inhibited by the C3d ligand in the INS-1E rat beta cell line.* Diabetologia, 2006. **49**(8): p. 1864-75.
481. Stephanou, A. and D.S. Latchman, *STAT-1: a novel regulator of apoptosis.* Int J Exp Pathol, 2003. **84**(6): p. 239-44.
482. Matsuda, M., et al., *Two species of human CRK cDNA encode proteins with distinct biological activities.* Mol Cell Biol, 1992. **12**(8): p. 3482-9.
483. Turner, C.E., *Paxillin and focal adhesion signalling.* Nat Cell Biol, 2000. **2**(12): p. E231-6.
484. Hannigan, G., A.A. Troussard, and S. Dedhar, *Integrin-linked kinase: a cancer therapeutic target unique among its ILK.* Nat Rev Cancer, 2005. **5**(1): p. 51-63.
485. Younes, M.N., et al., *Integrin-linked kinase is a potential therapeutic target for anaplastic thyroid cancer.* Mol Cancer Ther, 2005. **4**(8): p. 1146-56.
486. Gourlay, C.W. and K.R. Ayscough, *The actin cytoskeleton: a key regulator of apoptosis and ageing?* Nat Rev Mol Cell Biol, 2005. **6**(7): p. 583-9.
487. Dadgostar, H., et al., *Cooperation of multiple signaling pathways in CD40-regulated gene expression in B lymphocytes.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1497-502.
488. Pype, S., et al., *TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-kappa B activation.* J Biol Chem, 2000. **275**(24): p. 18586-93.
489. MacMillan-Crow, L.A. and J.A. Thompson, *Tyrosine modifications and inactivation of active site manganese superoxide dismutase mutant (Y34F) by peroxynitrite.* Arch Biochem Biophys, 1999. **366**(1): p. 82-8.

490. Mauro, C., et al., *ABIN-1 binds to NEMO/IKKgamma and co-operates with A20 in inhibiting NF-kappaB.* J Biol Chem, 2006. **281**(27): p. 18482-8.
491. Wuyts, A., et al., *Isolation of the CXC chemokines ENA-78, GRO alpha and GRO gamma from tumor cells and leukocytes reveals NH2-terminal heterogeneity. Functional comparison of different natural isoforms.* Eur J Biochem, 1999. **260**(2): p. 421-9.
492. Yarden, Y. and B.Z. Shilo, *SnapShot: EGFR signaling pathway.* Cell, 2007. **131**(5): p. 1018.
493. Baselga, J. and S.M. Swain, *Novel anticancer targets: revisiting ERBB2 and discovering ERBB3.* Nat Rev Cancer, 2009. **9**(7): p. 463-75.
494. Nagan, N. and R.A. Zoeller, *Plasmalogens: biosynthesis and functions.* Prog Lipid Res, 2001. **40**(3): p. 199-229.
495. Esposti, M.D., *Lipids, cardiolipin and apoptosis: a greasy licence to kill.* Cell Death Differ, 2002. **9**(3): p. 234-6.
496. Vernimmen, D., et al., *Identification of HTF (HER2 transcription factor) as an AP-2 (activator protein-2) transcription factor and contribution of the HTF binding site to ERBB2 gene overexpression.* Biochem J, 2003. **370**(Pt 1): p. 323-9.
497. Zhou, J., et al., *Duplication of the IGFBP-2 gene in teleost fish: protein structure and functionality conservation and gene expression divergence.* PLoS One, 2008. **3**(12): p. e3926.
498. Maher, B., *ENCODE: The human encyclopaedia.* Nature, 2012. **489**(7414): p. 46-8.
499. Gellman, S.H. and D.N. Woolfson, *Mini-proteins Trp the light fantastic.* Nat Struct Biol, 2002. **9**(6): p. 408-10.
500. Gardner, P.P., et al., *Rfam: Wikipedia, clans and the "decimal" release.* Nucleic Acids Res, 2011. **39**(Database issue): p. D141-5.
501. Amaral, P.P., et al., *Complex architecture and regulated expression of the Sox2ot locus during vertebrate development.* RNA, 2009. **15**(11): p. 2013-27.
502. Makhanova, N., et al., *Salt-sensitive blood pressure in mice with increased expression of aldosterone synthase.* Hypertension, 2008. **51**(1): p. 134-40.
503. Chen, Z.L., et al., *Proteolytic fragments of laminin promote excitotoxic neurodegeneration by up-regulation of the KA1 subunit of the kainate receptor.* J Cell Biol, 2008. **183**(7): p. 1299-1313.
504. Stevens, S.W., et al., *Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from Saccharomyces cerevisiae.* RNA, 2001. **7**(11): p. 1543-53.
505. Liao, J., et al., *Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer.* Mol Cancer, 2010. **9**: p. 198.
506. Wassarman, K.M., et al., *Identification of novel small RNAs using comparative genomics and microarrays.* Genes Dev, 2001. **15**(13): p. 1637-51.
507. Liu, Y., et al., *Cloning of two candidate tumor suppressor genes within a 10 kb region on chromosome 13q14, frequently deleted in chronic lymphocytic leukemia.* Oncogene, 1997. **15**(20): p. 2463-73.
508. Weinberg, Z., et al., *Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes.* Genome Biol, 2010. **11**(3): p. R31.
509. Rhee, S.G., H.Z. Chae, and K. Kim, *Peroxiredoxins: a historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling.* Free Radic Biol Med, 2005. **38**(12): p. 1543-52.
510. Omer, A.D., et al., *Homologs of small nucleolar RNAs in Archaea.* Science, 2000. **288**(5465): p. 517-22.
511. Heppner Goss, K., et al., *Attenuated APC alleles produce functional protein from internal translation initiation.* Proc Natl Acad Sci U S A, 2002. **99**(12): p. 8161-6.
512. Venkova-Canova, T., M. Patek, and J. Nesvera, *Control of rep gene expression in plasmid pGA1 from Corynebacterium glutamicum.* J Bacteriol, 2003. **185**(8): p. 2402-9.
513. Young, R.P., et al., *Chromosome 4q31 locus in COPD is also associated with lung cancer.* Eur Respir J, 2010. **36**(6): p. 1375-82.

514. Liston, P., et al., *Suppression of apoptosis in mammalian cells by NAIP and a related family of IAP genes.* Nature, 1996. **379**(6563): p. 349-53.
515. Maier, J.K., et al., *The neuronal apoptosis inhibitory protein is a direct inhibitor of caspases 3 and 7.* J Neurosci, 2002. **22**(6): p. 2035-43.
516. Koonin, E.V. and L. Aravind, *The NACHT family - a new group of predicted NTPases implicated in apoptosis and MHC transcription activation.* Trends Biochem Sci, 2000. **25**(5): p. 223-4.
517. Derrien, T., et al., *The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression.* Genome Res, 2012. **22**(9): p. 1775-89.
518. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data.* Nucleic Acids Res, 2011. **39**(Database issue): p. D152-7.
519. Kusenda, B., et al., *MicroRNA biogenesis, functionality and cancer relevance.* Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2006. **150**(2): p. 205-15.
520. Huang, H.Y., et al., *RegRNA: an integrated web server for identifying regulatory RNA motifs and elements.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W429-34.
521. Jacobsen, A., et al., *miRMaid: a unified programming interface for microRNA data resources.* BMC Bioinformatics, 2010. **11**: p. 29.
522. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.
523. Garcia, D.M., et al., *Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs.* Nat Struct Mol Biol, 2011. **18**(10): p. 1139-46.
524. Reshmi, G., et al., *Identification and analysis of novel microRNAs from fragile sites of human cervical cancer: computational and experimental approach.* Genomics, 2011. **97**(6): p. 333-40.
525. Meiri, E., et al., *Discovery of microRNAs and other small RNAs in solid tumors.* Nucleic Acids Res, 2010. **38**(18): p. 6234-46.
526. Stark, M.S., et al., *Characterization of the Melanoma miRNAome by Deep Sequencing.* PLoS One, 2010. **5**(3): p. e9685.
527. Lui, W.O., et al., *Patterns of known and novel small RNAs in human cervical cancer.* Cancer Res, 2007. **67**(13): p. 6031-43.
528. Cummins, J.M., et al., *The colorectal microRNAome.* Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3687-92.
529. Wang, F., et al., *A systematic survey of mini-proteins in bacteria and archaea.* PLoS One, 2008. **3**(12): p. e4027.
530. Moran, V.A., R.J. Perera, and A.M. Khalil, *Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs.* Nucleic Acids Res, 2012. **40**(14): p. 6391-400.
531. Levine, B., N. Mizushima, and H.W. Virgin, *Autophagy in immunity and inflammation.* Nature, 2011. **469**(7330): p. 323-35.
532. Mathew, R., V. Karantza-Wadsworth, and E. White, *Role of autophagy in cancer.* Nat Rev Cancer, 2007. **7**(12): p. 961-7.
533. Rosenfeldt, M.T. and K.M. Ryan, *The multiple roles of autophagy in cancer.* Carcinogenesis, 2011. **32**(7): p. 955-63.
534. Moreira, P.I., et al., *Autophagy in Alzheimer's disease.* Expert Rev Neurother, 2010. **10**(7): p. 1209-18.
535. Eng, C.H., et al., *Ammonia derived from glutaminolysis is a diffusible regulator of autophagy.* Sci Signal, 2010. **3**(119): p. ra31.
536. Tanida, I., T. Ueno, and E. Kominami, *Human light chain 3/MAP1LC3B is cleaved at its carboxyl-terminal Met121 to expose Gly120 for lipidation and targeting to autophagosomal membranes.* J Biol Chem, 2004. **279**(46): p. 47704-10.
537. Tanida, I., et al., *HsAtg4B/HsApg4B/autophagin-1 cleaves the carboxyl termini of three human Atg8 homologues and delipidates microtubule-associated protein light chain 3- and*

GABAA receptor-associated protein-phospholipid conjugates. J Biol Chem, 2004. **279**(35): p. 36268-76.
538. Langley, K.E., et al., *Molecular basis of beta-galactosidase alpha-complementation.* Proc Natl Acad Sci U S A, 1975. **72**(4): p. 1254-7.
539. Olson, K.R. and R.M. Eglen, *Beta galactosidase complementation: a cell-based luminescent assay platform for drug discovery.* Assay Drug Dev Technol, 2007. **5**(1): p. 137-44.
540. Ko, Y.H., et al., *Glutamine fuels a vicious cycle of autophagy in the tumor stroma and oxidative mitochondrial metabolism in epithelial cancer cells: Implications for preventing chemotherapy resistance.* Cancer Biol Ther, 2011. **12**(12).
541. Sakiyama, T., et al., *Glutamine increases autophagy under Basal and stressed conditions in intestinal epithelial cells.* Gastroenterology, 2009. **136**(3): p. 924-32.
542. Tannous, B.A., *Gaussia luciferase reporter assay for monitoring biological processes in culture and in vivo.* Nat Protoc, 2009. **4**(4): p. 582-91.
543. Ketteler, R. and B. Seed, *Quantitation of autophagy by luciferase release assay.* Autophagy, 2008. **4**(6): p. 801-6.
544. Bjorkoy, G., et al., *Monitoring autophagic degradation of p62/SQSTM1.* Methods Enzymol, 2009. **452**: p. 181-97.
545. El Debs, B., et al., *Functional single-cell hybridoma screening using droplet-based microfluidics.* Proc Natl Acad Sci U S A, 2012. **109**(29): p. 11570-5.
546. Wlodkowic, D. and Z. Darzynkiewicz, *Microfluidics: Emerging prospects for anti-cancer drug screening.* World J Clin Oncol, 2010. **1**(1): p. 18-23.
547. Rohde, C., et al., *High-throughput in vivo genetic and drug screening using femtosecond laser nano-surgery, and microfluidics.* Conf Proc IEEE Eng Med Biol Soc, 2008. **2008**: p. 2642.
548. Ohlfest, J.R., A.B. Freese, and D.A. Largaespada, *Nonviral vectors for cancer gene therapy: prospects for integrating vectors and combination therapies.* Curr Gene Ther, 2005. **5**(6): p. 629-41.
549. Sasaki, T., et al., *[New combination therapies for gastrointestinal cancer].* Gan To Kagaku Ryoho, 2000. **27**(3): p. 348-55.
550. Andersen, M.H., et al., *Cancer treatment: the combination of vaccination with other therapies.* Cancer Immunol Immunother, 2008. **57**(11): p. 1735-43.

Word Count: 83,879

Page Count: 268