Methods for genome-scale gene perturbation studies of the TRAIL-induced apoptosis pathway in mammalian cell culture

Ian Sudbery

Darwin College

University of Cambridge

&

The Wellcome Trust Sanger Institute

This dissertation is submitted for the degree of Doctor of Philosophy

wellcome trust sanger institute

UNIVERSITY OF CAMBRIDGE

Declaration

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements. No part of this work has been submitted for any other degree at this or any University.

This dissertation does not exceed the page limit specified by the Biology Degree Committee.

Ian Sudbery, November 2007.

Acknowledgements

There are so many people without whose help, scientific, intellectual and personal, this thesis would not have been possible, it is difficult to know where to start. I can only begin by thanking my supervisor Dr Ian Dunham for his support and guidance throughout my time as a student in his group and for giving so generously of his time right up to the end and even after his departure from the Institute. Good luck for whatever the future may hold. I must also thank all the members of the late Team 62: Andy, Cat, Caroline, Charmain, Christoph, Dave, Gayle, James, Jamil, Jo, John, Lotte, Mat (keep going Mat, the end is in sight), Owen, and Sarah for their technical and scientific help, and also for making the lab such a great place to work and for making the group a team, especially during what has been a hard time for all concerned. My special thanks to Dr John Collins and Charmain Wright for their help with all things molecular biological and for providing the ORF library and to Dr Dave Beare for his bioinformatics related help.

Thank you to Dr Andrew Fraser, for the original idea for this project, for all the guidance provided over the years and for the regular injections of enthusiasm, and to the other members of my thesis committee, Dr Derek Stemple and Dr Bill Skarnes.

Thank you to Dr Anton Enright and Dr Cei Abreu-Goodger for their scripts and help with analysing and interpreting the seed data and Dr Wolfgang Huber for his pointers on analysing screening data. The shRNA library was very kindly provided by Prof. Greg Hannon and a sample of it was sequenced by Bob Plumb. Miscellaneous sequencing was performed by the faculty small sequencing projects team at the Sanger Institute.

I also wish to thank to the non-scientific staff at the institute particularly Dr Christina Hedberg-Delouka for making sure things ran smoothly, and the guys in purchasing for putting up with my constant urgent requests for things I needed yesterday. Thank you to Mat, Juliette, Harry, Kath, Christine and Dad for extra proof reading.

I am blessed with a great many fantastic friends, too many to mention here. Thank-you to all the Sanger PhD students, to Catherine, Dave, David, Laura, Neil, Richard and to Keltie for being there when I needed you, keeping sane, giving me a kick when it was needed and generally looking after me.

My deepest gratitude to my family for making me who I am today and their support over the last 4 years – my mother Carol, my brother Nick and my father Prof. Pete Sudbery, who has been so much help it's like having another supervisor.

Abstract

Modern techniques, particularly RNA interference, but also the systematic over-expression of full length open-reading frames (ORFs), have promised to allow traditional genetic screening paradigms to be transferred to mammalian cell culture systems in order to study medically relevant pathways and annotate function onto the genome.

TNF Related Apoptosis Inducing Ligand (TRAIL) induces apoptosis in many tumour cells, but not in the majority of normal cells. As such it has generated much excitement as a potential anti-cancer treatment. However, the molecular basis of the regulation of sensitivity to TRAIL is not fully understood. Here an assay for the sensitivity of HeLa cells to TRAIL is used to compare different approaches to RNAi screening. Various tests indicated that RNAi screening for novel TRAIL genes is feasible using siRNAs but not shRNAs.

RNAi screens were carried out using both a library of siRNAs targeting 901 Kinase and Phosphatases and a larger library targeting the "Druggable Genome". Genes having the largest effect on TRAIL sensitivity were rigorously confirmed and controlled for off-target effects using multiple siRNAs and multiple assays. Thus eight novel genes involved in TRAIL-induced apoptosis were identified (Sharpin, MAST4, IKBKE, MAX, IGF1R, PDE11A, INADL and TEGT).

A thorough examination of the seed sequences of high scoring siRNAs revealed that several seed sequences were over-represented in high scoring siRNAs. This suggests that screening may enrich for siRNAs with relevant off-target effects. In addition comparison of these seed sequences to those of natural miRNAs identify four candidate miRNAs which may be involved in regulation of TRAIL-induced apoptosis.

A screen was also carried out to assess the effect of the over-expression of 288 full-length ORFs from chromosome 22. Several clones that have a reproducible effect on the sensitivity of cells to TRAIL were identified, although failure of these genes to have an effect in secondary assays mean that their physiological involvement in the pathway is unknown.

In conclusion, genome-scale systematic gene perturbation studies are powerful tools for annotation of gene function, and for isolating novel genes in medically relevant pathways, but they must be used with care and an awareness of their possible pitfalls.

Table of Contents

Full results of screens are available on the included CD or at

http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_Kinase/
http://www.sanger.ac.uk/HGP/Chr22/RNAi/TRAIL_DG/
http://www.sanger.ac.uk/HGP/Chr22/ORFScreen.txt

List of Figures

List of Tables

Abbreviations

5-FU	5-Fluorouracil
6mer	Hexamer
7mer-A1	Heptamer matching bases 2-7 of a mature miRNA with an additional A at position 1
7mer-m8	Heptamer matching bases 2-8 of a mature miRNA
8mer	Octamer
ADP	Adenosine DiPhosphate
ATP	Adenosine TriPhosphate
bp	Base Pair(s)
BSA	Bovine Serum Albumin
Casp8	Caspase-8
cDNA	Complementary DNA
cFLIP	Cellular FLICE Inhibitory Protein (also known as CFLAR)
CMV	CytoMegaloVirus
CNS	Central Nervous System
CNS-DCs	Central Nervous System DCs
COSMIC	Catalogue Of Somatic Mutations In Cancer
DAPI	4',6-DiAmidino-2-PhenylIndole
DcR1	Decoy Receptor 1 (also known as TNFRSF10C or TRAIL-R3)
DcR2	Decoy Receptor 2 (also known as TNFRSF10D or TRAIL-R4)
DCs	Dendritic Cells
DISC	Death Inducing Signalling Complex
DNA	Deoxyribose Nucleic Acid
DR4	Death Receptor 4 (also known as TNFRSF10A or TRAIL-R1
DR5	Death Receptor 5 (also known as TNFRSF10B or TRAIL-R2)
dsRBP	Double Strand RNA Binding Protein
dsRNA	Double-Stranded RNA
EDTA	EthyleneDiamineTetra-acetic Acid
EMCV	EncephaloMyoCarditis Virus
esiRNA	endoribonuclease-prepared siRNA
EST	Expressed Sequence Tag

FACS	Fluorescent Activated Cell Sorting
FADD	FAS Associated Death Domain protein
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FITC	Fluorescein IsoThioCyanate
Flp/FRT	FLiPase/Flipase Recombination Target
FWER	FamilyWise Error Rate
GDP	Guanosine DiPhosphate
GFP	Green Fluorescent Protein
GNF	Genomics Institute of the Novartis Research Foundation
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GTP	Guanosine TriPosphate
H_2O	Water
H_2O_2	Hydrogen Peroxide
IL-2	Interleukin-2
kb	KiloBase
LB	Luria-Bertani Broth
Log	Logarithm
LTR	Long Terminal Repeat
MAD	Median Absolute Deviation
MAPK	Mitogen Activated Protein Kinase
miRNA	MIcro RNA
mRNA	Messenger RNA
NF-$\varkappa$B	Nuclear Factor-kappa B
NK (cells)	Natural Killer (cells)
NoT	Not Transfected
NPI	Normalised Percentage Inhibition
nt	NucleoTide
Oligo	Oligonucleotide
ORF	Open Reading Frame
ORFeome	The totality of all ORFs in an organism
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction

PGK	PhosphoGlycerate Kinase
PKC	Protein Kinase C
pSM2	pSHAG-MAGIC2
PTGS	Post Transcriptional Gene Silencing
qPCR	Quantitative PCR
qRT-PCR	Quantitative Reverse Transcription PCR
RdRP	RNA Dependent RNA Polymerase
Rep 1	Replicate 1
Rep 2	Replicate 2
RISC	RNA-Induced Silencing Complex
RLC	RISC Loading Complex
RLU	Relative Luminescent Units
RNA	Ribose Nucleic Acid
RNAi	RNA Interference
rpm	Revolutions Per Minute
RT-PCR	Reverse Transcription PCR
SCF	Skp, Cullin, F-box
shRNA	Short Hairpin RNA
$shRNA^{mir}$	Short hairpin RNA with micro RNA based design
siRNA	Small Interfering RNA
ssRNA	Single-Stranded RNA
TNF	Tumor Necrosis Factor
TRAIL	TNF-Related Apoptosis Inducing Factor
UTR	UnTranslated Region
UV	UltraViolet (radiation)
cAMP	Cyclic adenosine monophosphate
cGMP	Cyclic guanosine monophosphate