# 7 DISCUSSION AND CONCLUSIONS

One of the great challenges for geneticists and molecular biologists is to turn the wealth of sequence data generated by genome sequencing projects into medically relevant functional knowledge. A large part of this is the definition of the function of the many protein coding open reading frames identified in genomes. Observervational techniques such as expression profiling and interaction studies can imply a wealth of information about associations of open reading frames and allow the generation of hypotheses on the basis of these associations. However, direct intervention in the functioning of a gene has long been the geneticist's weapon of choice for defining function in the context of a particular pathway

or process. The availability of genome sequence and the development of high-throughput gene perturbation techniques in mammalian cell culture finally allows the extension of this work into a human system. Using the techniques of RNAi and ORF expression, pathways directly relevant to human disease can now be investigated.

The dream of many a cancer biologist is to turn some property uniquely inherent in the nature of the cancer cell against it. Initially, sensitivity to TRAIL-induced apoptosis seemed to be just such a property, with transformed, but not non-transformed cells being sensitive to apoptosis triggered by TRAIL (Walczak et al. 1999). However, it was soon found that some cancer cell types were resistance to TRAIL-induced apoptosis (Zhang et al. 1999). The mechanisms which distinguish normal cells from transformed cells and sensitive from insensitive transformed cells are not fully understood.

Thus the aims of this work were two fold. Firstly to examine, assess and compare different methods for genome scale gene perturbation and secondly to apply these methods to the identification of novel genes involved in the regulation of sensitivity to TRAIL-induced apoptosis.

The success of screening experiments depends critically on careful selection of assay conditions and selection of gene perturbation technique. Plasmid based hairpin RNAs present an attractive technique for gene knock-down. Such plasmids are an infinite resource, allow for selection of transfectants/transformants, allow introduction into hard-to-transfect cell types, and critically allow the execution of pooled selection type screens. However, the experiments in Chapter 3 suggest that the shRNAs containing within the shRNA library to which the author had access do not elicit phenotypes as strong as those elicited by chemically synthesised siRNAs. One explanation of this is that siRNAs are easier to transfect than shRNAs, reaching higher transfection efficiencies. However, the gap between shRNAs and siRNAs is not closed by using drug selection to increase transfection efficiency, which has more of an effect on the sensitivity of cells transfected with the negative control than it does on cells transfected with an shRNA targeted at knocking down a gene of interest. There is little data in the literature directly comparing shRNAs with siRNAs. The original report of shRNAs functioning in mammalian cell culture demonstrated that for a single sequence targeting Luciferase shRNAs were as effective as siRNAs in a dual Luciferase assay that would normalise for expression from the plasmids (Paddison et al. 2002). It has been reported that chemically synthesised hairpin RNAs outperform chemically synthesised siRNAs. However, the authors did not test for induction of the interferon response, which is a possibility with RNAs of that length (Siolas et al. 2005). One possibility for the difference

between the shRNAs and siRNAs tested here is that the sequences of the shRNA were designed according using different algorithm to the siRNA sequences. Once designed and created, the sequences of shRNAs are more likely to stay stable, while for siRNAs, sequences can be updated as the siRNAs are re-synthesised. This means that siRNAs purchased are more likely to have been designed using more recent algorithms than shRNAs, and therefore are likely to be more effective. Further, with the decreasing price of siRNA synthesis and taking into account the cost of preparing plasmid DNA to transfection quality for the many thousands of constructs required for a screen, the price advantage of shRNAs in terms of reagent cost is less clear. However, the cost of the RNAi library is a fraction of the total cost of conducting a screen, with the total cost of plasticware, cell culture media, transfection reagent and assay materials easily totalling more than the cost of the library. By dramatically reducing the number of experiments necessary to conduct a screen, pool selections have the potential to reduce the total cost of a study. This is where shRNAs could be a real advantage. However, such an approach relies on the assay in question having a very high signal/noise ratio, and therefore a powerful, reliable method for reducing gene expression. Pooled screening also reduces the range of processes that can be studied. Although there are several examples of successful pooled selections using shRNAs (Berns et al. 2004, Hattori et al. 2007, Kolfschoten et al. 2005, Nicke et al. 2005, Paradis et al. 2007, Westbrook et al. 2005), siRNAs will remain the reagent of choice, particularly for those phenotypes where the difference between hits and negatives is quantitative rather than qualitative.

Although at high-throughput there is obviously a large amount of variance in the results of a quantitative assay, the results for particular siRNAs are fairly reliable. This does not, however, translate into a high reliability in the results for particular genes. There are two possible reasons for this. Firstly variance in the efficiency of an siRNA in knocking down the intended target, and secondly a result for an siRNA can be due to off-target effects, rather than due to the effect of knocking down the intended target (or a combination of both). The process of screening is likely to exacerbate both problems. The screening process selects siRNAs which have a large effect on the assay. Therefore, the highest scoring siRNAs are likely to be the ones that have the largest effect on transcripts which affect the assay. These transcripts may be the intended target or other transcripts which affect the assay. It is therefore not surprising that other siRNAs targeting the same transcript are often not as efficient at knocking down the transcript of interest. It is possible that this problem could be reduced by more efficient introduction of siRNAs into the cell. Indeed, conditions in the assay here were not optimised for maximal transcript knockdown, but rather for largest

effect on the assay, which is a combination of conditions that allow knock-down of the transcript in question and the conditions that give maximal sensitivity to TRAIL.

The other problem with using a protocol designed to introduce more siRNA into the cell is that this will increase the severity of phenotypes caused by off-target effects. The off-target effects are generally weaker than the on-target effects (Birmingham et al. 2006, Jackson et al. 2003). Thus the ideal level of knock-down is one where the reduction of the intended target is sufficient to elicit a phenotype, but the level of knock-down of off-target transcripts is not. It often assumed that the finding that a phenotype induced by an siRNA is due to off-target effects is unlucky – the exception rather than the rule. This may be the case for single gene experiments, where siRNAs targeting a gene are chosen on an arbitrary basis rather than on the basis of their performance in an assay against many thousands of others. However, the work here suggests that in the screening situation it is more likely that at least part of the effect of an siRNA is due to its off-target effects than in single gene experiments. Analysis of seed sequences in the high scoring siRNAs such as that performed here or in (Lin et al. 2007) may help to identify suspect seeds and siRNAs containing them. However, it is clear that the seed sequences of an siRNA does not wholly determine its activity in an assay. While other determinants of siRNA off-targeting have been described (Nielsen et al. 2007), it seems likely that the effects of any siRNA are the combination of a large number of weak effects on off-target transcripts and a single, large effect on the intended target.

These conclusions support several practical suggestions for following up RNAi screens. Firstly, a hit should be confirmed with multiple siRNAs, preferably more than two, and further, these siRNAs should not include any siRNAs included in the screen. siRNAs used for confirmation should also be checked to ensure that they do not contain seed sequences that are enriched in high-scoring siRNAs from the screen. It has been report that chemical modification of siRNAs can increase their specificity. The addition of various groups to the nucleotides on the passenger strand of the siRNA prohibits its entry into RISC, and the addition of an O-Methyl group to base 2 of the guide strand may also reduce the number of off-target effects triggered by this strand (Jackson et al. 2006). See the introduction for further details. The use of siRNAs which have been chemically modified could be of use here, particularly if the chemically modified siRNAs with the same sequences as those from the screen failed to elicit the same phenotype, Finally, if at all possible hits should be confirmed using rescue or inhibition through some non-RNAi method. This final suggestion is oft repeated (Echeverri et al. 2006, Sarov, Stewart 2005) and rarely followed, due to the difficulty associated with such experiments. However, there are examples of RNAi

results being confirmed by rescue using either expression of an ORF to rescue siRNAs targeted at the 3' UTR of a gene (Yi et al. 2007) or by using transfection of a mouse BAC to rescue an siRNA directed at a human gene (Kittler et al. 2005).

RNAi screening is often thought of as a "reverse genetics" technique: each gene is methodically tested in turn and its effect on the process of interest recorded. However, RNAi screens are far from 100% sensitive. The overlap between the screen of kinases and phosphatases presented here and another screen of the kinome is minimal (Aza-Blanc et al. 2003), and the druggable genome screen failed to have several highly important genes amongst its top hits: while, for example, library siRNAs targeting Caspase 8 may have scored highly in the screen, siRNAs targeting PDE11A, a gene which is surely less important in the TRAIL-induced apoptosis pathway, scored more highly. This, taken together with the fact that results from an siRNA cannot be guaranteed to be solely related to the intended target, suggest that RNAi screening is analogous more to traditional or forward genetics. Instead of providing a quantitative readout of the involvement of each and every gene in a process, they can identify a set of genes with a confirmed qualitative effect. As such RNAi screens are powerful tools, which can identify novel genes involved in a pathway, but should not be regarded as "saturating". This is also the case for RNAi screens in model organisms. A genome wide screen of the *C. elegans* genome for embryonic and post embryonic phenotypes found identified 63% genes with a known phenotype (Kamath et al. 2003). This is similar to the proportion of genes previously associated with TRAIL-induced apoptosis that could be confirmed in assay development experiments here, although higher than the number of such genes that performed well in the screens themselves. The situation is even more pronounced in model organism cell culture systems. Two independent screens for regulators of the JAK/STAT signalling pathway found 91 and 121 hits respectively. The overlap between these two sets of hits was only six genes (Baeg, Zhou & Perrimon 2005, Müller et al. 2005)

This is highlighted by the results of an siRNA screen for TRAIL regulators published in the last few days of preparation of this thesis, during the composing of this discussion (Ovcharenko et al. 2007). In this study a fluorescent caspase-3 activity assay, whereby cleavage of a caspase-3 substrate leads to an increase in fluorescence was used as a primary assay. The screen included 3 siRNAs targeting each gene, and the screen was performed in triplicate. Caspase-3 activity levels were not normalised to the pre-treatment viability of the cells, so differences in Caspase-3 activity levels could be due to difference in cell growth. Genes were selected as candidate hits if the average score of two siRNAs were above an arbitrary threshold. The presence or absence of positive controls is not reported, results from

| Screen | Gene | Score | | Average Survival | | Rank | |
|---|---|---|---|---|---|---|---|
| | | siRNA A | siRNA B | siRNA A | siRNA B | siRNA A | siRNA B |
| **Druggable Genome** | | | | | | | |
| | PTGS1 | -0.01 | -3.50 | 86.5% | 9.0% | 4514 | 10851 |
| | CLCN3 | -0.063 | -1.07 | 29.4% | 24.0% | 7013 | 8290 |
| **Kinase and Phosphatase** | | | | | | | |
| | CDK2 | 1.1 | 0.99 | 45.0% | 30.1% | 167 | 197 |
| | CDK4 | 0.13 | -0.56 | 37.3% | 9.60% | 586 | 1117 |
| | CDK9 | -0.33 | -0.78 | 12.6% | 13.8% | 950 | 1315 |
| | IRAK4 | 0.31 | -0.15 | 21.20% | 14.30% | 466 | 804 |
| | MAP3K6 | -0.62 | -0.83 | 9.8% | 14.9% | 1261 | 1459 |
| | MAP3K8 | 0.67 | 0.52 | 28.70% | 30.90% | 294 | 382 |
| | PAK2 | 0.27 | -0.41 | 22.5% | 10.6% | 487 | 996 |
| | PAK1 | -0.74 | -1.00 | 13.3% | 8.6% | 1273 | 1502 |
| | HK1 | -1.24 | -1.45 | 1.6% | 0% | 1679 | 1783 |

**Table 7-1 Performance of siRNAs targeting hits from Ovcharenko *et al* screen in screens performed here**

The score, average raw survival and rank in the complete ranking is shown for siRNAs targeting hits from the Ovcharenko *et al* screen that were included in the screens reported in this work . Rank represent position out of 11162 siRNAs for the druggable genome screen and 1785 siRNAs for the kinase and phosphatase screen. The genes ALG2 and LRRFIP1 are reported as hits by Ovcharenko are not included in the library used in this work

negative controls are also not reported but are used to normalise data. Candidates were validated by performing the same assay on the same siRNAs and applying a t-test to the results. The authors claim to use a 1% p-value threshold, with no multiple testing corrections, but several of the genes they claim as hits do not make this threshold. Again, the authors to not report results from any positive controls. No measurement of the effect of the siRNAs on the level of transcript or protein is undertaken. No secondary assay is undertaken. The authors report seven known and thirteen novel genes which when knocked down reduced sensitivity to TRAIL, although only 5 of the known and 8 of the novel genes meet their own criteria for a hit. None of these genes overlap with the set of genes reported here, and they confirmed only one of the novel genes in a separate screen of kinases (Aza-Blanc et al. 2003). The performance of their hit genes in the screens conducted here is shown in Table 7-1. They do report that in their hands, knock-down of IGF1R reduces sensitivity to TRAIL, but does not meet their criteria for statistical significance. Differences between the two screens include the cell line used (MDA-MB-453 cells mammary carcinoma cells vs. the HeLa cervical carcinoma cells used here and in Aza-Blanc *et al*) and the assay used (TRAIL-induced Caspase-3 activity vs. alamarBlue survival assay used here an in Aza-Blanc *et al*). However, it is still significant that none of the genes isolated here were also isolated in this similar screen.

In many ways overexpression screens suffer from the same technical reliability vs. biological reliability issue as RNAi screens. The effects of 3 of the 4 hits from the

overexpression screen tested were confirmed to have an effect on post TRAIL treatment viability (although in at least one case this effect disappears when the effect of overexpression on un-treated cells is accounted for). However, the difficulty in the interpretation of the biological significance of these hits is not connected with associating the technical results with the gene being overexpressed, but what it means, biologically, that overexpressing these genes leads to this phenotype. Although the hits from the overexpression screen of chromosome 22 genes could be shown to reliably affect TRAIL-induced cytotoxicity, no effect on TRAIL-induced caspase activity, or sensitivity to other apoptosis inducers could be shown. This could be due to a question of sensitivity, as one of the hits RBX1, had previously been shown to have an effect on levels of Caspase-3 (Tan et al. 2006). Indeed, that two of the three hits could be connected to regulation of apoptosis suggests that overexpression screening can provide relevant results. Even so, it is unclear what the true meaning of an overexpression phenotype is. If a protein is expressed at a higher level than is found in an *in vivo* setting or in an environment which it is not normally expressed, is its effect necessarily indicative of its *in vivo* function? Many of the early experiments which suggested that the TRAIL decoy receptors were involved in regulation of TRAIL sensitivity used overexpression of these receptors (Degli-Esposti et al. 1997a, Degli-Esposti et al. 1997b, Emery et al. 1998, Pan et al. 1997a). Later studies found minimal correlation between decoy receptor expression and sensitivity to TRAIL –induced apoptosis (Ganten et al. 2004, Kim et al. 2000, Zhang et al. 1999).

In the case of the overexpression screen there was a clear distinction between those ORFs that did cause a change in TRAIL-inducted cytotoxicity (hits) and those that did not. This distinction was less clear in the case of the RNAi screens, with each siRNA having a quantitative effect on the percentage of cells which survived treatment with TRAIL. It is unclear whether this is due variability in the efficiency of the siRNAs knocking down their targets, due to the different complement of off-target effects caused by each of the siRNAs, or whether it is due to differing effects of knocking-down each gene on the sensitivity of cell to TRAIL-induced apoptosis. Indeed, Friedman and Perrimon have suggested that the sorts of continuous distributions seen here in the screen results imply a network model of signalling, where each phenotypic output is not the result of a defined number of genes, but each gene has a smaller or larger contribution to signalling of the network (Friedman, Perrimon 2007). They also argue that such distributions are unlikely to be the product of off-target effects since they are also observed in genetic screens with a quantitative output. The hits identified in the screens here act in seemingly disparate pathways, with members of the

MYC, NF-κB (IKBKE), tyrosine kinase (IGF1R) and cAMP (PDE11A) pathways identified as well as genes seemingly connected to none of the canonical signalling pathways. Indeed, an RNAi screen for genes involved in regulating Caspase-3 activity in *Drosophila* cells has recently implicated genes encoding metabolic enzymes in the regulation of sensitivity to apoptosis inducing stimuli (Yi et al. 2007). This suggests that rather than the sensitivity of cells TRAIL being the output of a single linear signalling pathway it rather depends on the complex interactions between many of the cell's signalling modules.

Several of these modules are known to have both pro- and anti-apoptotic effects. In many cases, where a gene was isolated here with a known role in control of apoptosis, the role is either anti-apoptotic with the role described here being pro-apoptotic (e.g. TEGT); pro-apoptotic with the role described here being anti-apoptotic (e.g. AIFM3); or both pro- and anti-apoptotic (e.g. IGF1R or IKBKE). IGF1R signals through AKT, which can have both positive and negative effects on TRAIL-induced apoptosis (Chen et al. 2001, Pugazhenthi et al. 2000, Remacle-Bonnet et al. 2005, Thakkar et al. 2001). Therefore the effect of IGF1R signalling must depend on the balance of these signals as determined by other parts of the network. The same reasoning applies to the activity of NF-κB possibly through the balance of the action of the c-Rel and RelA subunits (Ravi et al. 2001). TRAIL itself has both pro- and anti-apoptotic activity through its action on NF-κB and so transformation must imply either an inhibition of this pro-survival pathway, a strengthening of the pro-apoptotic pathway, or a change in balance of NF-κB outputs such that NF-κB activity becomes pro-apoptotic. This all suggests that rather than there being a single mechanism through which cells become sensitive to TRAIL on transformation, any of the alterations to the cells signalling network which led to transformation also lead to a promotion of TRAIL-induced apoptosis.

Given that there may be multiple routes to TRAIL resistance/sensitivity it will be of interest to determine if the same genes are involved in regulation of TRAIL-induced apoptosis in cell lines other than the HeLa cells used here. That is, do the same perturbations of the signalling networks in one cell line lead to the same changes in phenotypic output as in another cell line. One interpretation of the lack of overlap between the screens reported here and a screen for regulators in a breast carcinoma cell line suggests that this isn't the case, although another interpretation is that RNAi screens are not saturating (Ovcharenko et al. 2007).

One overlooked aspect of RNAi screening is that although RNAi screening is analogous to screening hypomorphic mutations in protein coding genes, it is also analogous

to overexpression screening of miRNAs. miRNAs are key regulators of many aspects of cell biology and can control the expression of many genes co-ordinately. It is not unreasonable to suppose therefore that miRNAs are involved in the sensitivity of cells to apoptosis. Indeed, a report published during the preparation of this discussion reports a list of miRNAs, overexpression of which changes the sensitivity of cells to TRAIL-induced Caspase-3 activity (Ovcharenko et al. 2007). Included in this list are miR-145 and miR-155, miRNAs which were identified in chapter 5 as containing seed sequences which were either over-represented in the hit siRNAs, or enriched in high scoring siRNAs in general. This supports the idea that seed analysis can isolate relevant and interesting phenomena. It is also reported that miR-26a alters the level of non-TRAIL-induced Caspase-3 activity. miR-26a contains the seed sequence ACTTGA, which is also found in four of the hit siRNAs, and is enriched in the high-scoring siRNAs in general. It also appears multiple times in the 3' UTRs of DR4, DR5 and BIS. Thus it is possible that siRNAs containing this seed are acting as a miRNA and knocking down the same transcripts as miR-26a. This would also implicate miR-26a in the regulation of Trail-induced cytotoxicity.

## 7.1 Future directions

In addition to the confirmed hits reported here, each of the siRNA screens identified a number of "unconfirmed" hits. These genes were targeted by only one siRNA which significantly altered the sensitivity of cells to TRAIL-induced cytotoxicity or Caspase-3/7 activity, further siRNAs did not significantly alter the sensitivity of cells to TRAIL, but neither did they knock-down the targeted transcript to the same extent. Results from further siRNAs could help either to categorise these genes as hits or identify that the effects of the original siRNAs were due to off-target regulation. Results from further siRNAs could also increase confidence in several of the genes here classified as confirmed hits. Sharpin, MAST4, IKBKE, INADL and TEGT are all genes for which questions as to their role in TRAIL-induced apoptosis remain. For Sharpin and MAST4 siRNA targeting them gave contradictory results in caspase activity assays, while siRNAs targeting IKBKE give contradictory results in experiment examining the effects of knockdown on sensitivity to a range of apoptosis inducers. Either one or both siRNAs targeting INADL and TEGT contain suspected seed sequences – confidence would be increased by results from siRNAs not containing these seed sequences.

As stated above, the final confirmation of an siRNA result is given by rescue of the RNAi induced phenotype. This can be achieved using siRNAs directed against the UTR

sequence of the gene in question, and rescuing the phenotype by expressing the gene's open reading frame from a plasmid construct without the UTR sequence. This is made easier by the availability of ORF clones for a large portion of the genome. It will also be of interest to see if the result of knocking down these genes in HeLa cells can be replicated in other cell types.

Ultimately, confirmed hits simply form a list of genes. Our knowledge of a process is only really increased when the role of these genes in the process is understood. Although studying the literature can help to generate hypotheses about how these genes are involved in TRAIL-induced apoptosis, only direct experimentation can confirm or refute these hypotheses. A good place to start in a functional exploration of the effects of these genes is their effect on the expression levels of known direct actors in TRAIL-induced apoptosis. Obvious candidates are the death receptors, the cFLIP DISC inhibitor and Mcl-2, all of which have been previously reported to be the endpoints of pathways regulating TRAIL-induced apoptosis (Wang et al. 2004/5, Ricci et al. 2004, Ricci et al. 2007). In the case of the death receptors, it is important that it is the surface expression rather than the bulk protein or RNA level that is measured (Ren et al. 2004).

Some of the hypotheses for the action of hits suggested by the literature make specific testable predictions. For example, if PDE11A functions to inhibit AKT induced Bcl-2 expression through CREB (Pugazhenthi et al. 2000), then PDE11A knock-down should increase CREB-mediated transcriptional activation, which could be measured using a reporter system.

Little is known about the effects of miRNAs on apoptotic pathways, with the first reports beginning to be published about connections between miRNAs and the TRAIL pathway (Ovcharenko et al. 2007). The involvement in the regulation of TRAIL sensitivity of two of the four miRNAs sharing one of the seed matches, highlighted as suspicious by an analysis of seeds in high scoring siRNAs, has already been confirmed (Ovcharenko et al. 2007). Transfection of artificial miRNA mimics with the same sequence as miR-26a and miR-384 could determine if these miRNAs regulate TRAIL sensitivity in HeLa and other cell lines. It would also be of interest to see if the expression level of these miRNAs in clinical isolates is correlated with the sensitivity of these isolates to TRAIL-induced apoptosis.

Finally, the seed analysis performed in this work is, to the best of the author's knowledge, the most thorough investigation of seed sequences in the hit list of an siRNA screen ever performed. However, while this analysis allows the generation of many hypotheses, it does not prove any of them. Possibly the easiest hypothesis to investigate

would be that siRNAs with seeds from the "enriched" seed set have, on average, higher numbers of off-target effects than other siRNAs. Expression profiles could be generated for cells transfected with siRNAs containing these seeds and compared with expression profiles for cells transfected with siRNAs targeting the same genes which did not contain suspect seed sequences. Indeed, such profiles might also help to identify the off-target effects of import for TRAIL-induced apoptosis. On a smaller scale, the effect of siRNAs containing suspect seed sequences on known regulators of TRAIL-induced cytotoxicity could be studied using qRT-PCR.

Another way in which off-targeted transcripts could be identified, would be to investigate the effect of these siRNAs on genes predicted to be off-targets, and also the effect of intentional knock-down of these predicted targets on TRAIL-induced apoptosis. One method for predicting these targets was outlined in Chapter 5. However, very recently, more sophisticated algorithms for predicting the off-target effects of siRNAs have been suggested. Nielsen *et al* describe an algorithm based on the number and length of seed matches, plus the AU content, and conservation, of the sequence surrounding the matches, which allowed the prediction of the off-target effects of several siRNAs (Nielsen et al. 2007). An implementation of this algorithm could be developed to predict the most common off-target effects between siRNAs containing suspect seeds.

The hypothesis that the effect of siRNAs targeting hit genes from the screen is a combination of off and on target effects could also be tested. Four siRNAs could be designed, two targeting a gene known to be involved in the TRAIL-induced apoptosis pathway (gene A), and two targeting a gene known not to be involved (gene B). One siRNAs targeting each of the genes would contain a suspect seed sequence and one would not. If the siRNAs targeting each gene were of similar efficiencies, the hypothesis would predict that the siRNA targeting gene A which contains the suspect seed would have the largest effect on TRAIL-induced apoptosis, follow by the siRNA targeting gene A which does not contain the suspect seed and the siRNA targeting gene B which does. Finally the siRNA targeting gene B which does not contain the seed would be predicted to have minimal effect.

Several of these suggestions are simple and should not take long to perform, in particular the testing of additional siRNAs against hits, the titration of the amount of siRNA used, and the testing of miRNA mimics. These were not performed due to severe time limitations towards the end of this project. Other suggestions are more open ended, such as avenues of investigation into the biological relevance of the selected hits, and the microarray experiments aimed at further understanding the importance of hit and enriched seeds, both

of which constitute separate projects in their own right. Finally, while it would be possible to repeat the overexpression screen and follow-up experiments using a positive control and better negative controls, the negative nature of the results suggest that this would not be a good use of time and resources.

## 7.2 Conclusion

This work has shown that, at least for screening in the TRAIL-induced apoptosis system, siRNAs are a more powerful tool for functional screening than shRNAs. It has shown that RNAi screening identifies siRNAs that have a reproducible effect on the process of interest. Although RNAi screening is less reliable at identifying genes than siRNAs which are involved in the process of interest, genes with a confirmed effect on the process can be identified. Six novel genes connected with TRAIL-induced apoptosis were identified in this way, along with three genes with a known effect. These genes were from distinct pathways.

This work has also suggests that overexpression screens of cloned ORFs can also be used to identify constructs with a reproducible effect on the TRAIL-induced cytotoxicity assay, although the biological significance of these hits is unknown. Thus the technique did not prove useful for furthering understanding of the pathway.

Seed analysis has shown that the process of siRNA screening can enrich for siRNAs that contain certain seed sequences. Although not demonstrated, the obvious hypothesis is that these seed sequences specify relevant off-target effects. This seed analysis also identified four miRNAs with a possible role in TRAIL-induced apoptosis. The involvement of two of these was confirmed elsewhere experimentally. This shows that a knock-down by an siRNA should also potentially be regarded as an overexpression of a miRNA.

In conclusion, genome-scale systematic gene perturbation studies are powerful tools for annotation of gene function, and in isolating novel genes in medically relevant pathways, but they must be used with care and an awareness of their possible pitfalls. Care should be taken interpreting the function of a gene based solely on isolation in an overexpression screens. Like RNAi screens in model organisms, siRNA screens in mammalian cells do not isolate all of the genes known to be associated with a process and there can be limited overlap between genes isolated in different screens in the same pathway. It also appears that off-target effects may be more prevalent than may have been appreciated in the past. Application of the recommendations outlined above to future screens will help to increase confidence in the results. It should also be remembered that small RNAs have a natural role in the cell and that transfection of an siRNA could be equivalent to overexpression of a

miRNA. However, awareness of this can allow candidate miRNAs connected with the pathway to be identified and turn this potential problem into an advantage