6 AN OVEREXPRESSION SCREEN OF ORFS ON CHROMOSOME 22

In the preceding chapters the assay developed at the start of this thesis was applied in two RNAi screens. The benefits and drawbacks to this approach were discussed, particularly with relevance to the reliability and specificity of the results obtained. In an RNAi screen, the involvement of a gene in a process is assessed by studying the effect of reducing the level of its transcript. This is equivalent to studying loss-of-function mutations (with null/hypomorphic alleles of the gene in question) in traditional genetics. The role of a gene

in a process can also be implied by studying the effect of increasing the levels of a transcript. In traditional genetics, this is equivalent to studying gain-of-function mutations (with hypermorphic alleles of the gene in question). Prior to the use of RNAi this was the only form of genetic screening available in the non-sexual, diploid, mammalian cell culture system. Random mutagenesis can be used to screen for gain-of-function mutants as they tend to be dominant. A more common approach is to introduce libraries of cDNA clones generated by the reverse transcription of RNA, isolated from tissue or cell lines into cell lines and selecting for clones that induce the desired phenotype. Once the clone is identified it can be sequenced to identify its source. Indeed one group has taken this approach to identifying regulators of the TRAIL-induced apoptosis pathway, identifying two known apoptosis inhibitors from a library of HepG2 cDNAs (Burns, El-Deiry 2001). However, such strategies usually involve, to a greater or lesser extent, some amount of pooling of clones and therefore will suffer from the problems with such a strategy outlined earlier in this thesis. Most importantly, the clone in question must induce a very large effect compared to the background level in order to be identified in a selective screen. Furthermore, such cDNA libraries are generally of an unknown complexity, with the make up being dependent on the complement of transcripts expressed in the source material. They will also contain truncated and mis-spliced transcripts. The availability of high-quality annotations of genomes allows for the construction of libraries of clones that contain one, sequence verified, clone for each full-length open reading frame (ORF) in a genome (the ORFeome,(Brasch, Hartley & Vidal 2004)). Such collections may be used to conduct reverse genetic screens for the effect of over-expression of these ORFs, in a one-well-one gene manner.

Here a pilot of such an approach is described. Plasmids driving the over expression of each of 288 full length ORFs from chromosome 22 are introduced individually into HeLa cells, and their effect on TRAIL-induced apoptosis assessed.

6.1 The Chromosome 22 ORF collection

The Chromosome 22 ORF collection used in this chapter is a collection of clones corresponding to 288 of the 398 (72%) predicted full-length open reading frames identified on chromosome 22 (Collins et al. 2003, Collins et al. 2004). The ORFs were cloned using an annotation driven approach. Many large scale ORF cloning efforts are based on sequencing full-length cDNA collections (which include 3' and 5' UTRs), selecting a clone to represent each gene and sub-cloning the ORF. An annotation driven approach involves creating PCR primers directed to each annotated ORF, amplifying the ORFs from cDNA pools and

cloning them into a sequencing vector. Clones containing the ORF are then sequenced and compared to the sequence predicted from the annotations. Clones with base changes that are not previously reported Single Nucleotide Polymorphisms (SNP), and clones that are splice variants which do not maintain the reading frame of the ORF are rejected. This leads to a single, sequence verified clone for each ORF (Figure 6.1). ORFs in the chromosome 22 ORF collection were cloned both with and without stop codons and then sub-cloned into holding and expression vectors (with either a C or N terminal T7 tag) compatible with Invitrogen's Gateway cloning technology.

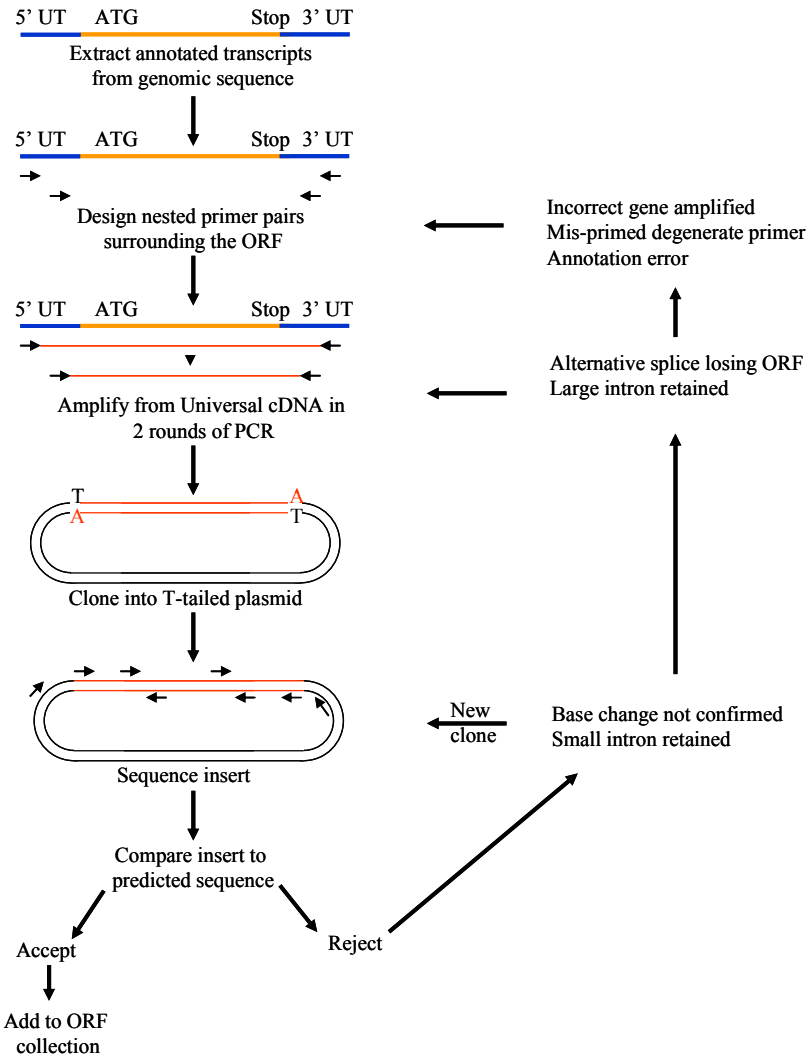

Figure 6.1 An annotation driven ORF cloning strategy (from Collins et al. 2004)

ORFs are tagged at both C and N terminals in separate constructs to control for effects of the tag on the function or localisation of the resulting protein. It was been reported that tagging ORFs at the C terminal is generally more reliable than tagging at the N terminal, presumably because tagging at the N terminal would mask any localisation sequences at this terminus (Palmer, Freeman 2004, Simpson et al. 2000). However, others have found that

tagging at the N terminal is at least as good as or better than tagging at the C terminal (J. Collins, personal communication).

6.1.1 Gateway cloning

The Gateway cloning system utilises modified enzymes from bacteriophage λ to clone fragments and allow the movement of fragments between vectors using homologous recombination, without restriction digestion, purification and ligation. Inserts are generated by PCR amplification using primers tagged with a bacteriophage λ recombination site (attB). A modified enzyme from bacteriophage λ (BP clonase) is then used to recombine the attB sites with the bacteriophage λ recombination site attP in a donor vector to form an "entry" clone. The recombination leads to the attB and attP sites forming composite attL and attR sites, with the attL sites flanking the insert. This recombination removes a ccdB gene from the donor vector and allows selection against unrecombined vectors in bacterial strains in which the ccdB gene is toxic. Inserts can be sub-cloned into another vector which contains two attR sites flanking a ccdB gene, known as a destination vector, in a second recombination reaction using the LR clonase enzyme.

The chromosome 22 ORF collections is available cloned into the pGEM holding vector, as entry clones cloned into the donor vector pDONR223, and cloned into an expression vector based on pCDNA3, with a T7 epitope fused at either to C or N terminal and expression of the ORF driven from the CMV promoter. The expression vectors are arrayed in six 96-well plates, three with C terminal tags (plates 1,3 and 5) and three with N terminal tags (plates 2,4 and 6).

6.2 Design and execution of screen

DNA was successfully prepared from 555 of the 576 clones which comprise the 288 chromosome 22 ORF set cloned in expression vectors, tagged at both the C and N terminals. While several genes exist for which it is known that overexpression leads to a reduction in TRAIL sensitivity (e.g. the gene for the anti-apoptotic protein cFLIP, Inhibitor proteins such as XIAP, cIAP1 or cIAP, or indeed the TRAIL decoy receptors), no such clones are present in this ORF set, and are not easily obtainable in the correct format. Due to a lack of the time required to clone and tests such a construct, the screen was conducted without a positive control. Choice of such a negative control is difficult. The standard control in such experiments would be an empty vector. However, it is unclear if this is a suitable negative control, as it does not control for the effects of producing large amounts of

protein. Another option is to select a protein which is not expected to be involved in the process. One option would be a non-native protein such as GFP or Luciferase. This would not control for the effects of large amounts of endogenous protein being produced however. Another option is to choose a native protein which is not predicted to be involved. In order to select such a protein as truly not having an effect on TRAIL-induced cytotoxicity a panel of such proteins could be tested. Since the majority of ORFs are not expected to affect the sensitivity of cells to TRAIL-induced cytotoxicity, measuring against the baseline effect seen in the majority of clones is effectively equivalent to testing a panel of potential negative controls which includes all the constructs in the library. .

DNA prepared from the 555 chromosome 22 ORFs containing clones was transfected in duplicate into HeLa cells (passaged four times since defrosting) and assayed for sensitivity to 0.5µg/ml TRAIL using alamarBlue. Plates were processed in batches of three plates per experiment.

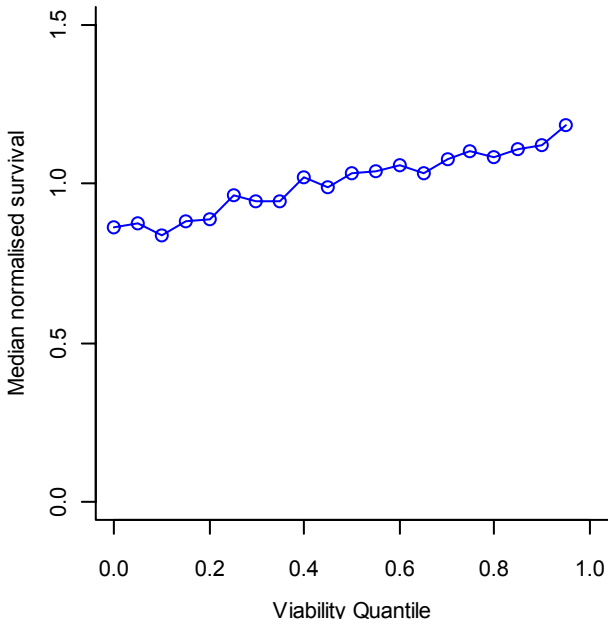

Figure 6.2 Relationship between pre-treatment viability and sensitivity to TRAIL-induced cytotoxicity
Viability of cells in each well prior to treatment was normalized to plate median viabilities. Normalised viabilities were divided into 20 quantiles. The median normalized post treatment survival was calculated for wells in each of these quantiltes.

It has been previously observed that the density of cells at time of treatment affects the sensitivity of cells to TRAIL-induced apoptosis. In the previous chapter, the relationship between pre-treatment viability and post-treatment survival was examined to establish a pre-treatment viability cut off, below which results would be removed from further analysis (Figure 5.2). This examination showed a sharp drop in post-treatment survival in wells which scored in the bottom 20% for pre-treatment viability. For the results from the ORF screen no such sharp drop was observed (Figure 6.2). There is an increase in normalized

survival with increasing pre-treatment viability across the whole range of pre-treatment viability, although the range of median normalized survival is much smaller than is the case for kinase and phosphatase RNAi screen. The relationship observed can not be removed by applying a cut-off to the data. In order to remove those data points where the pre-treatment survival is so low as to cause problems for data analysis the wells with the 5% lowest pre-treatment viability were removed from further analysis.

Data was normalised using the median survival for each plate. Figure 6.3 shows the effects of this normalization. The difference between different batches of plates processed on different days, using different batches of cells, can be clearly seen in the unnormalised data (Figure 6.3a), while no such differences are apparent in the normalised data (Figure 6.3b).

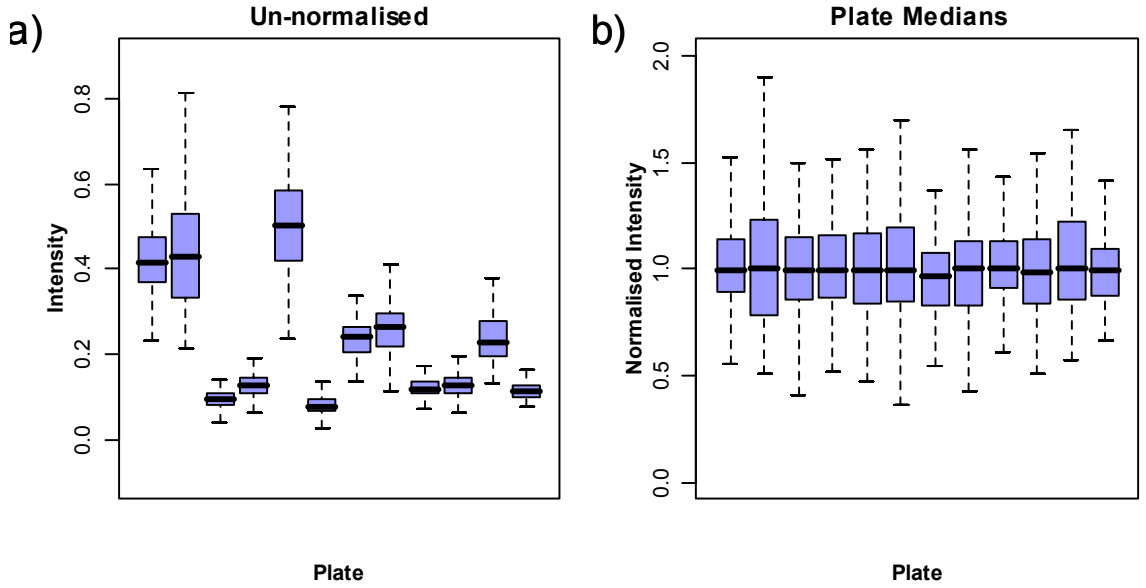

Figure 6.3 Normalisation of data from over-expression screen of Chromosome 22 ORFs
a) Boxplot of raw survival data from screen on a per plate basis. b) Boxplot of survival data normalized to plate median survival on a per plate basis

The distribution of data from previous screens was shown to be non-normal. A link was demonstrated between mean survival and standard deviation between replicates. Log transformation of the data did not abolish this relationship, instead inverting it. In the case of the data from the siRNA screen of the druggable genome, while log transformed data showed a relationship between mean survival and standard deviation between replicates, the relationship was weaker than for non-transformed data. The mean survival rank of data from the over-expression screen was plotted against the standard deviation between replicates for both untransformed and log transformed data (Figure 6.4). A relationship was observed between mean survival rank and standard deviation in untransformed data (Figure 6.4a). This relationship is less pronounced in log transformed data (Figure 6.4b).

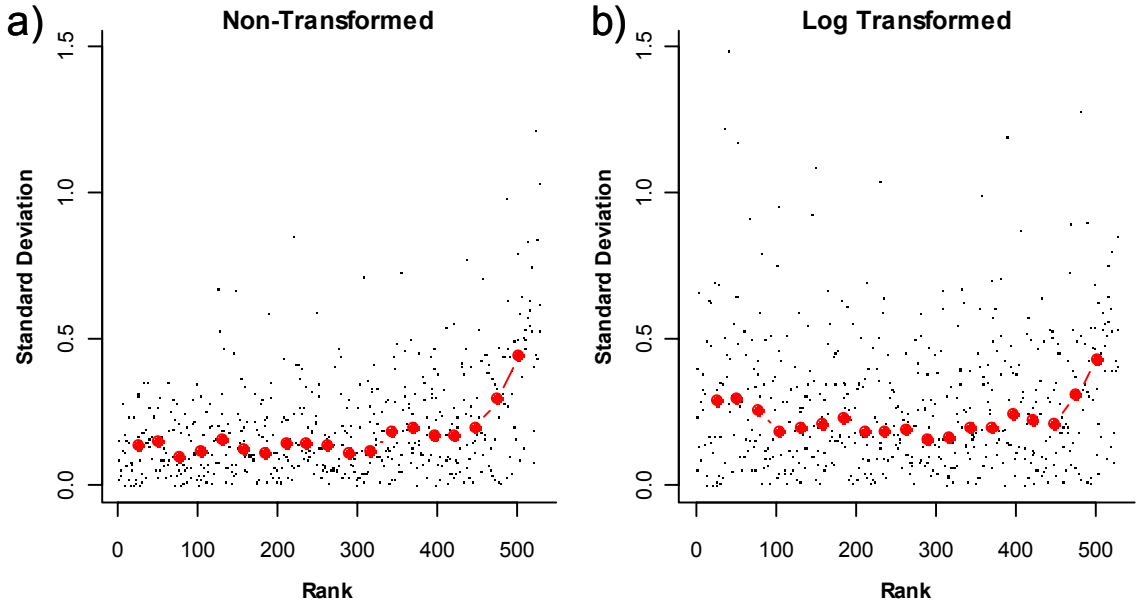

Figure 6.4 Relationship between rank of mean and standard deviation between replicates of Chromosome 22 ORF expression screen

For each expression clone the mean of the normalized data for each replicate was calculated. The rank of this mean was then plotted against the standard deviation between the replicates for a) Non-transformed data and b) Log transformed data. The red line in each plot represents the running median standard deviation.

Data were analysed using the R/Bioconductor package cellHTS, first excluding wells with a low pre-treatment viability, and then median normalizing plates with a log transformation, and using the minimum of replicates as a summary function.

6.3 Screen Results

Without controls, an analysis of the quality of the screen is restricted. There is very little correlation between the two replicates of the screen (Figure 6.5a). The correlation coefficient is 0.32 (r^2 = 0.10). This means that variation caused by random variation between replicates is greater than the variation caused by the effect of the expression of the ORF on the survival of TRAIL-treated cells. This could be due to a large amount of random variation. Alternatively it could suggest that the majority of the ORFs have little or no effect of the survival of cell treated with TRAIL. The correlation between the N and C terminal tagged version of the same ORF is slightly lower (Figure 6.5b), with a correlation coefficient of 0.22 (r^2 = 0.048). This could suggest that there is a real difference between ORFs tagged at different ends, or alternatively that again, the effects of random variation are higher than effects caused by the expression of the ORFs. One reason for the lack of effect could be that the ORFs are not expressed from the transfected constructs. This is unlikely however as transfection of the constructs from the same preparations into COS cells, immuno-fluorescent staining showed expression of 73% of genes (J. Collins, manuscript in

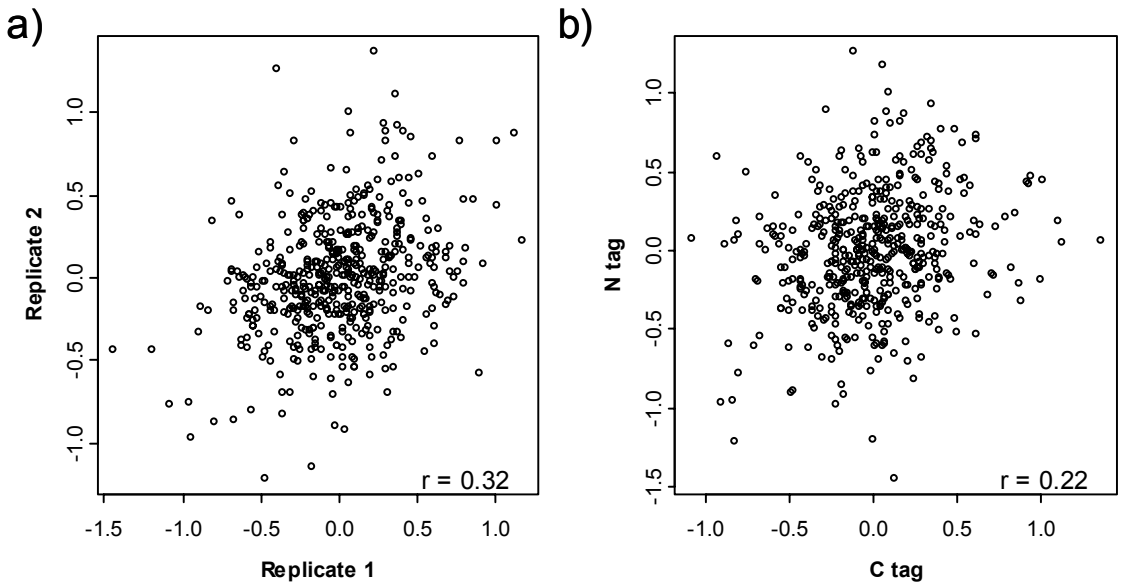

Figure 6.5 Correlations in data from chromosome 22 ORF expression screen
a) Normalized survival from replicate 1 plotted against normalised survival from replicate 2. b) Plot showing normalised survival of the two constructs expressing the same ORF, tagged at either the C or N terminal. The Pearson's correlation co-efficient is shown in the bottom right corner of each plot.

preparation).

The distribution of scores from the screen is roughly normal (Figure 6.6a and Figure 6.6b). Examination of the distribution of well scores across and between plates reveals no obvious position dependent effects (Figure 6.6c).

ORFs were ranked according to their score in the screen. A portion of this ranking is shown in Table 6-1. Examination of the quartile-quartile plot of the data, which plots the actually quartile of a datum point against the theoretical quartile were the data from a normal distribution, shows four points clearly score higher than would be expected if the data were normally distributed (blue points, Figure 6.6b). These points correspond to the four highest scoring clones, those expressing C-terminal tagged RBX1 and AIFM3 and the N-terminal tagged LIMK2 and MTMR3 (Table 6-1).

6.4 Confirmation of Hit genes

The four ORFs which scored significantly higher than other ORFs in the screen (RBX1, AIFM3, LIMK2 and MTMR3) were selected for confirmation. In each case only one of the two clones containing each ORF scored highly in the screen, although the clone containing the second RBX1 clone also appeared in the top ten clones (Table 6-1). This could be due to interference from the T7 epitope tag when at one end of the ORF, but not the other. To avoid interference from the tag, the ORFs were transferred into an expression vector containing no tag.

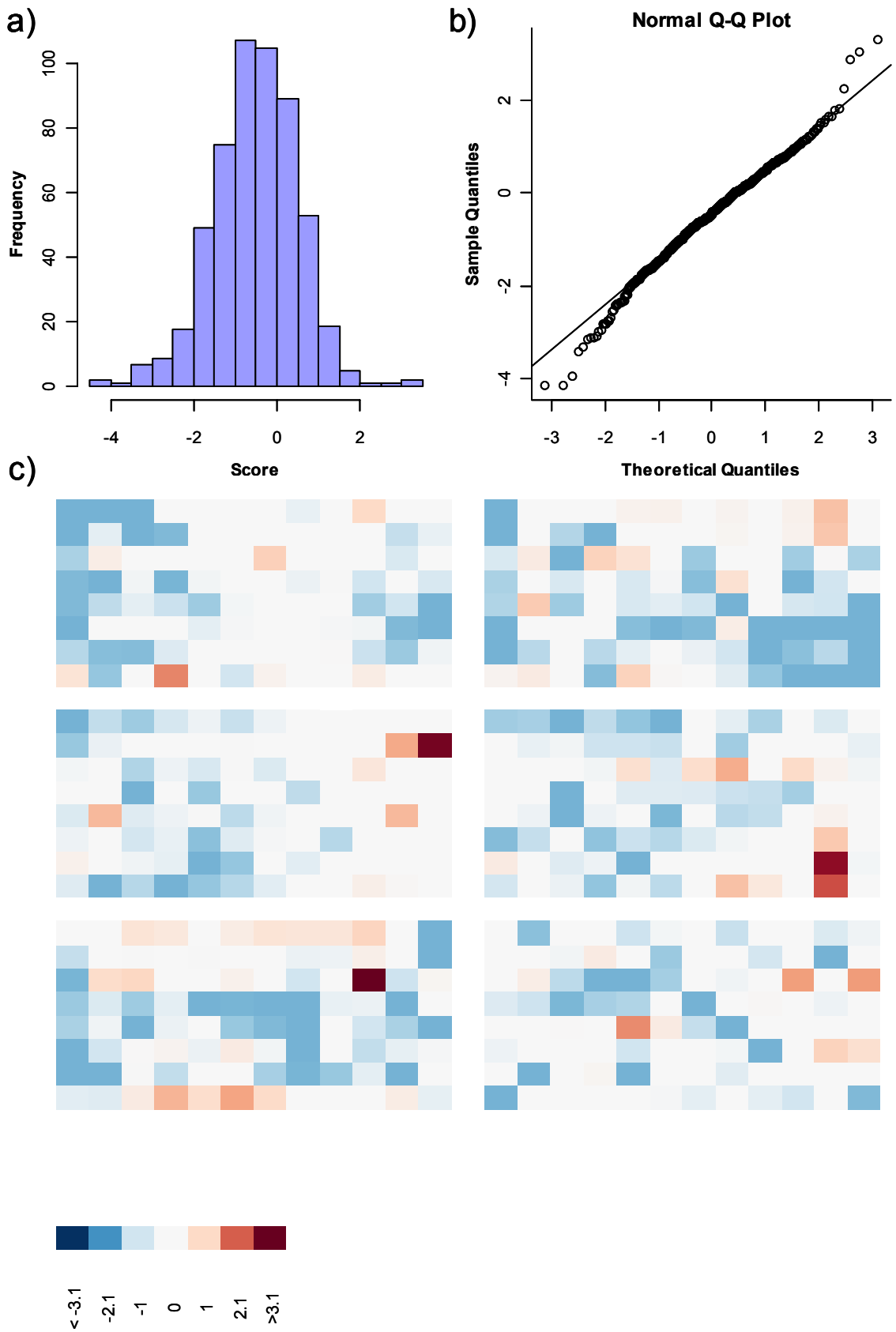

Figure 6.6 Scores of clones from screen of chromosome 22 ORFs
a) Histogram showing distribution of scores. b) A normal Quartile-Quartile plot of scores. The actual quantile of a data point is plotted against the theoretical quartile of that point if the data were normally distributed. If the data were perfectly normally distributed all data would fall on the line shown. c) A heat map showing the scores of each well from the screen. High scoring wells are shown in red, low scoring wells in blue.

Gene ID	*Gene Symbol*	*Tag*	*Description*	*Normalised Survival*		***score***
				Rep 1	*Rep 2*	
9978	RBX1	C	Ring-box 1	NA	0.948	**3.26**
150209	AIFM3	C	Apoptosis-inducing factor, mitochondrion-associated, 3	1.129	0.876	**3.01**
3985	LIMK2	N	LIM domain kinase 2	1.007	0.824	**2.83**
8897	MTMR3	N	Myotubularin related protein 3	0.771	0.825	**2.22**
129138	ANKRD54	C	Ankyrin repeat domain 54	0.625	NA	**1.8**
10478	SLC25A17	N	Solute carrier family 25 (mitochondrial carrier; peroxisomal membrane protein, 34kDa), member 17	0.606	0.738	**1.74**
10740	RFPL1S	N	Ret finger protein-like 1 antisense	0.856	0.474	**1.63**
9978	RBX1	N	Ring-box 1	0.794	0.468	**1.61**
23765	IL17RA	C	Interleukin 17 receptor A	0.541	0.46	**1.56**
150280	HORMAD2	C	HORMA domain containing 2	1.012	0.436	**1.5**
9514	GAL3ST1	N	Galactose-3-O-sulfotransferase 1	0.513	0.626	**1.48**
758	MPPED1	C	Metallophosphoesterase domain containing 1	0.491	NA	**1.41**
2130	EWSR1	C	Ewing sarcoma breakpoint region 1	0.469	0.851	**1.35**
66035	SLC2A11	C	Solute carrier family 2 (facilitated glucose transporter), member 11	NA	0.392	**1.35**
468	ATF4	N	Activating transcription factor 4 (tax-responsive enhancer element B67)	0.448	0.449	**1.29**
57591	MKL1	N	Megakaryoblastic leukemia (translocation) 1	0.444	0.598	**1.28**
84133	ZNRF3	N	Zinc and ring finger 3	0.559	0.355	**1.22**
51512	GTSE1	N	G-2 and S-phase expressed 1	0.415	0.884	**1.2**
1399	CRKL	N	V-crk sarcoma virus CT10 oncogene homolog (avian)-like	0.415	0.506	**1.19**
150290	DUSP18	C	Dual specificity phosphatase 18	0.395	0.346	**1.14**

Table 6-1 Top scoring clones form chromosome 22 ORF over-expression screen
Table shows the 20 highest scoring clones from the over-expression screen of ORFs from the chromosome 22 ORF collection. Gene ID is the EntrezGene ID for the gene and Tag indicates the terminal at which the ORF is tagged with a T7 epitope. Complete ranking can be found on included CD or online at http://www.sanger.ac.uk/HGP/Chr22/ORFScreen.txt

No Gateway compatible expression vectors are available that do not contain a tag fused to either terminal of the ORF. To allow analysis of the effect of expression of hit ORFs free from tag dependent effects a Gateway compatible expression vector containing no tag was created. First the gateway cassette was removed from the pCDNA3.GW.V5N vector by digestion, relegation and transformation into a ccdB sensitive strain to select against gateway cassette containing plasmids. The V5 tag was then removed by digestion. The gateway cassette was then reintroduced and the vector transformed into a ccdB insensitive strain. Multiple colonies resulting from the transformation were cultured and plasmid DNA prepared. The orientation of the gateway cassette within these clones was verified by restriction digestion. This new vector was named the pcDNA3.GW.NoTag vector.

Gateway recombination was used to transfer ORF inserts identified in the screen from the entry clones into the pCDNA3.GW.NoTag vector. The identities of the ORFs were confirmed by DNA sequencing.

To confirm the effect of over-expression of the 'hit' ORFs on sensitivity of cells to TRAIL-induced cytotoxicity, cells were transfected with pCDNA3.GW.NoTag constructs containing each of the ORFs and tested for sensitivity to a range of TRAIL concentrations (Figure 6.7). While in the screen, the majority of constructs which had no effect on TRAIL sensitivity could be used as a negative control, this is not possible here, since all the clones could be expected to score highly. Therefore, despite the issues raised earlier, the empty vector pCDNA3.T7 was used as a negative control. In addition the ORF of the gene PICK1 was also transferred into the NoTag vector. PICK1 had a low score (-2..34) in the screen and serves as a second negative control. The pSM2.shCasp8.2 construct, which expresses a hairpin targeting Caspase-8, was used as a positive control.

Transfection of the positive control, pSM2.shCasp8.2 caused a modest increase in survival compared to the empty pCDNA3.T7 vector at all concentrations of TRAIL tested, with 19% of pSM2.shCasp8.2 transfected cells surviving treatment with 1μg/ml TRAIL compared with 11% of vector transfected cells. This compares with an increase in survival from 29% to 40% seen previously (Figure 3.11). Transfection of clones expressing PICK1 has survivals very similar to the empty vector (12% of PICK1 transfected surviving treatment with 1μg/ml TRAIL compared with 11% of vector transfected cells, Figure 6.1e). Transfection of clones expressing RBX1 leads to a decrease in sensitivity to TRAIL at all concentrations tested (Figure 6.7a), with 30% of cells surviving treatment with 1μg/ml

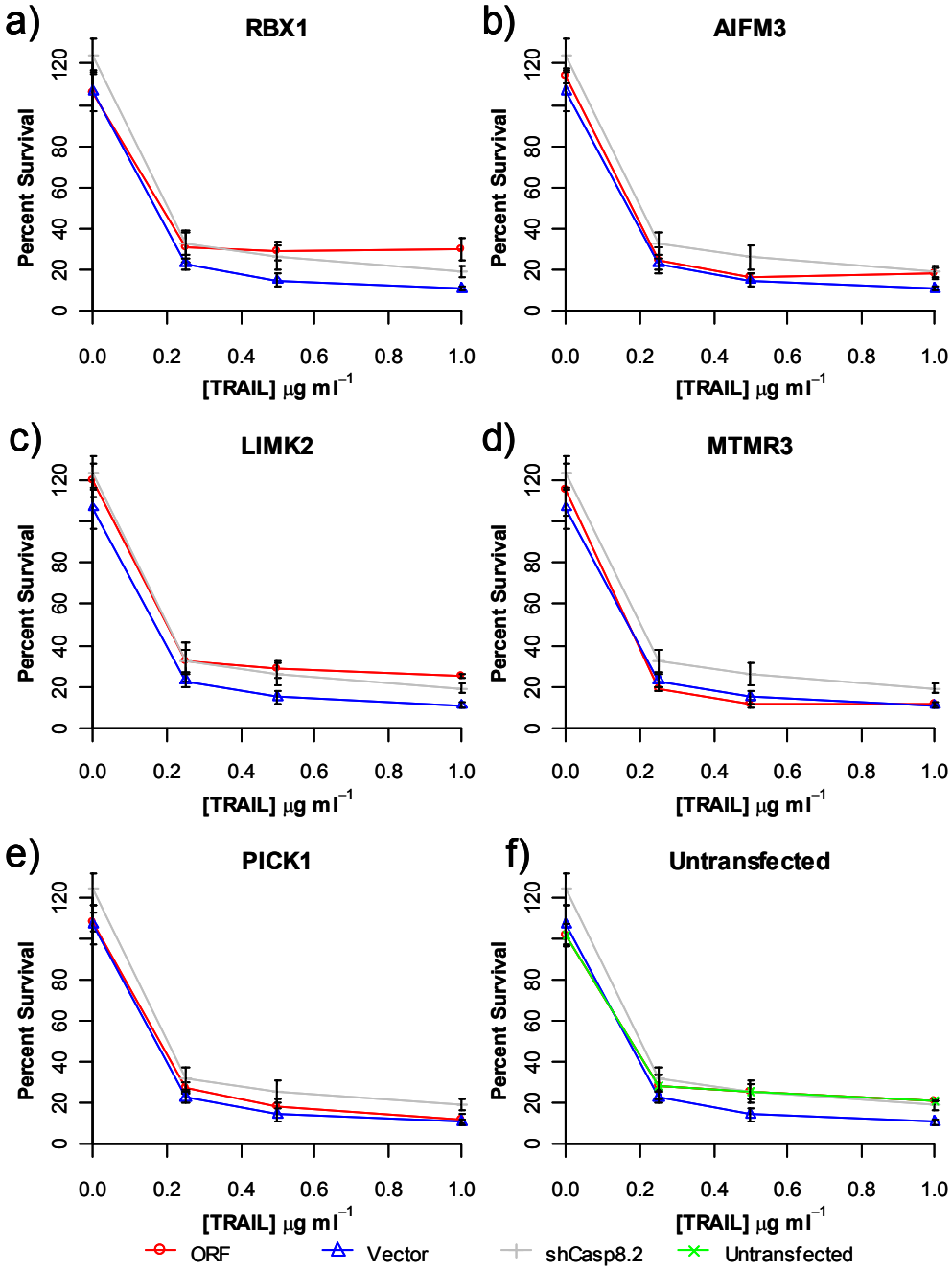

Figure 6.7 Effect of transfection of candidate hit ORFs on sensitivity to TRAIL-induced cytotoxicity
Cells were transfected with either expression clones for a) RBX1, b) AIFM3, c) LIMK2, d) MTMR3, e) PICK1, pSM2.shCasp8.2 as a positive control, or empty pCDNA3.T7 vector as a negative control, or f) mock transfected. Viability was assessed 48 hours later and cells were treated with the concentration of TRAIL indicated for 24 hours and viability reassessed. Error bars represent 1 standard deviation, n=3

TRAIL compared to 11% of cells transfected with vector. Similarly transfection of cells with the construct expressing LIMK2 caused a reduction in sensitivity at all concentration compared to transfection with an empty vector, with 25% and 11% of cells transfected with LIMK2 or empty vector respectively surviving treatment with 1μg/ml TRAIL (Figure 6.7c). A difference between cells transfected with the construct expressing AIFM3 and empty vector was not seen at lower concentrations of TRAIL, however more AIFM3 expressing

cells survived treatment with 1μg/ml than cells transfected with empty vector (18% compared with 11%, Figure 6.7b). In all cases these differences at 1μg/ml TRAIL are significant at the 5% significance level (p values calculated by Bonferroni corrected Student's t-test on log transformed data). No difference was seen between pCDNA3.MTMR3 transfected cells and vector transfected cells at any concentration of TRAIL (Figure 6.7d). As was seen in previous chapters, mock transfected cells were less sensitivity than negative control transfected cells at all concentrations of TRAIL tested (Figure 6.7f).

Worryingly transfection of several of the clones increases the survival of untreated cell as well as TRAIL treated cells when compared to control transfected cells (although in no individual case is this difference significant at the 5% level). This raises the possibility that the effect of the clones on the survival of cells compared to the negative control could be due to a growth advantage, or some other, non-TRAIL specific effect. In order to address this possibility, the data were renormalized by dividing all survival values by the survival of untreated cells, thus expressing survivals as a proportion the untreated cells which survived treatment (Figure 6.8). With this new normalisation, the difference between RBX1 and LIMK2 transfected cells and negative control transfected remain statistically significant at the 5% level. However, the difference between normalised survival for AIFL transfected cells and normalised survival for negative control transfected cells is not significant at the 5% level (survival of AIFL transfected cells treated with 1μg/ml TRAIL was 15% of that in untreated cells, compared with 10% for negative control transfected cells).

These data show that transfection with constructs expressing three of the four ORFs identified from the screen do reduce the sensitivity of cells to TRAIL-induced cytotoxicity. The finding that untransfected cells are also less sensitive to TRAIL-induced cytotoxicity raises the possibility that the effect seen for these ORFs might not be due to the over expression of the particular ORF. Specifically it is possible that some property of these constructs is preventing their entry into cells, preventing the expression of the ORF from the construct or killing transfected cells, such that at the time of treatment, the cells are not expressing the ORF being tested. To rule out the possibility that transfection of these constructs was killing the transfected cells, the pre-treatment viabilities of cells transfected with each construct were compared. There was no difference between the viability of cells transfected with any of the pCDNA3.ORF constructs compared to cells transfected with the empty vector pCDNA3.T7, suggesting that the difference was not due differences in toxicities between transfections with these different constructs (data not shown). A small, but highly statistically significant difference was seen between the viabilities of mock-

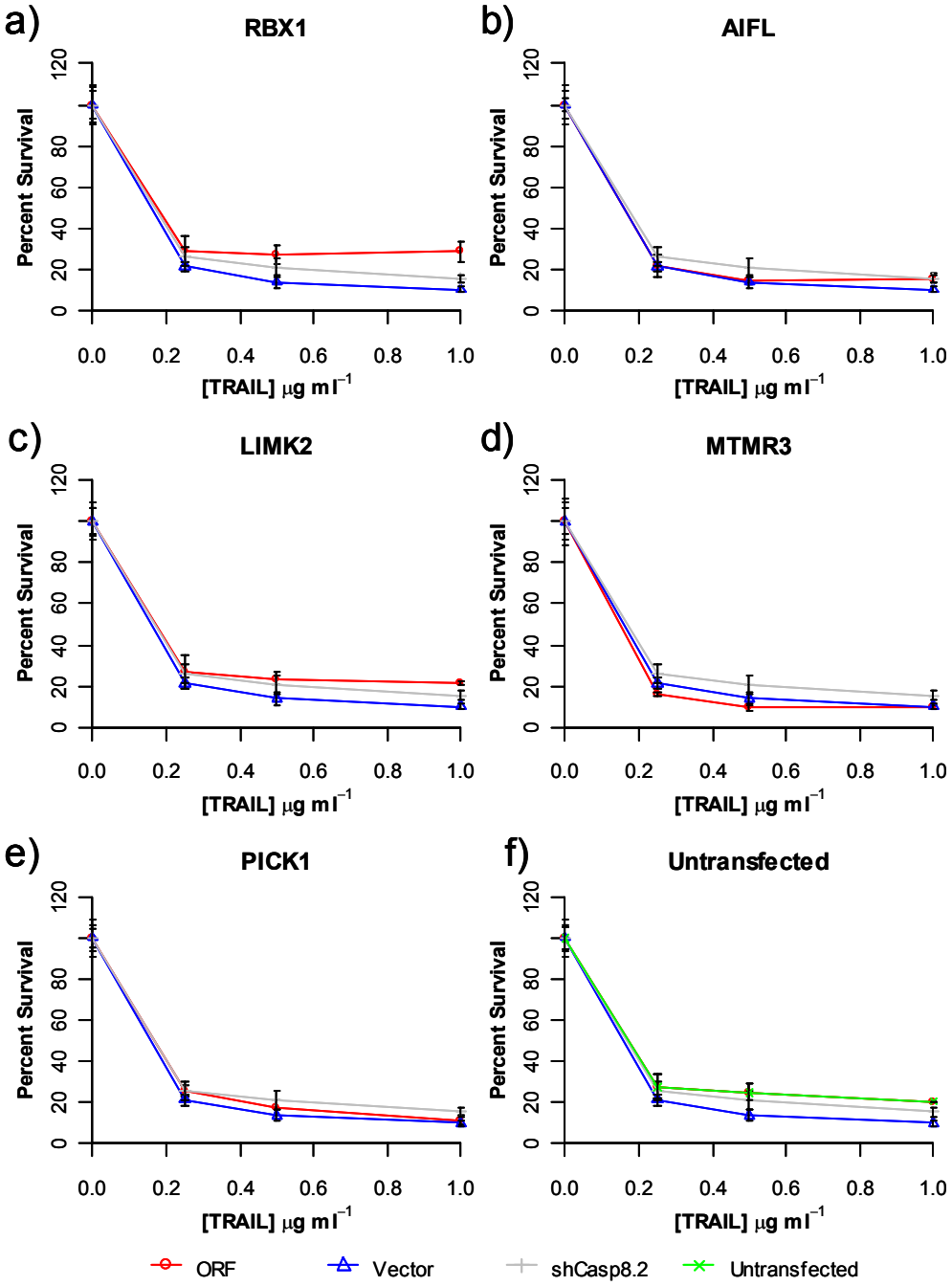

Figure 6.8 Renormalised effect of transfection of candidate hit ORFs on sensitivity to TRAIL induced cytotoxicity

Experiment was carried out as described for Figure 6.7. Data is presented as survival relative to untreated cells.

transfected and negative controls transfected cells (viability of mock transfected cells was 110% that of cells transfected with the negative control, p = 0.001 by Students' t-test, n = 12). Thus difference in the TRAIL sensitivity of cells transfected with hit ORF expressing constructs and negative control transfected cells is not due to the constructs killing transfected cells. It is possible that the increased survival of mock-transfected cells compared to non-transfected cells could be due to the toxicity of transfection, as it was shown that sensitivity to TRAIL is linked to pre-treatment cell number (Figure 6.2).

These data show that transfection with constructs expressing three of the four ORFs identified from the screen do reduce the sensitivity of cells to TRAIL-induced cytotoxicity. The finding that untransfected cells are also less sensitive to TRAIL-induced cytotoxicity raises the possibility that the effect seen for these ORFs might not be due to the over expression of the particular ORF. Specifically it is possible that some property of these constructs is preventing their entry into cells, preventing the expression of the ORF from the construct or killing transfected cells, such that at the time of treatment, the cells are not expressing the ORF being tested. To rule out the possibility that transfection of these constructs was killing the transfected cells, the pre-treatment viabilities of cells transfected with each construct were compared. There was no difference between the viability of cells transfected with any of the pCDNA3.ORF constructs compared to cells transfected with the empty vector pCDNA3.T7, suggesting that the difference was not due differences in toxicities between transfections with these different constructs (data not shown). A small, but highly statistically significant difference was seen between the viabilities of mock-transfected and negative controls transfected cells (viability of mock transfected cells was 110% that of cells transfected with the negative control, $p = 0.001$ by Students' t-test, $n = 12$). Thus difference in the TRAIL sensitivity of cells transfected with hit ORF expressing constructs and negative control transfected cells is not due to the constructs killing transfected cells. It is possible that the increased survival of mock-transfected cells compared to non-transfected cells could be due to the toxicity of transfection, as it was shown that sensitivity to TRAIL is linked to pre-treatment cell number (Figure 6.2).

To ensure that transfection with the ORF expressing clones is driving expression of these ORFs at higher than normal levels, the expression of the ORFs in cells transfected with both fresh preparations of the T7 tagged constructs used in the screen and the NoTag constructs was measured using qRT-PCR (Figure 6.9). Unfortunately, despite multiple attempts, DNA could not be prepared for the N-terminal tagged MTMR3 ORF. Since the cloned ORFs do not contain introns, it is not possible to design oligonucleotide primers that will amplify from mRNA, but not DNA. Therefore, to control for the presence of plasmid DNA from the transfection, RNA was prepared in a mock reverse-transfection reaction. qPCR on this RT- sample would amplify from DNA only, therefore allowing an estimate of the proportion of the increase in RNA expression measured in the reverse-transcribed samples (RT+) that is due to contamination of the RNA extraction with DNA. The majority of the samples were shown to be contaminated with DNA. However, DNA contamination made up only a small part of the increase in product amplified when transfected with ORF

expression construct. For example, there was an 87 fold increase in amplification of AIFM3 in the RT- sample prepared from pCDNA.AIFM3.T7C transfected cells compared to cDNA from the negative control sample. However there is a 1,257 fold increase in amplification of AIFM3 in the RT+ sample prepared from the same transfection compared to the negative control sample.

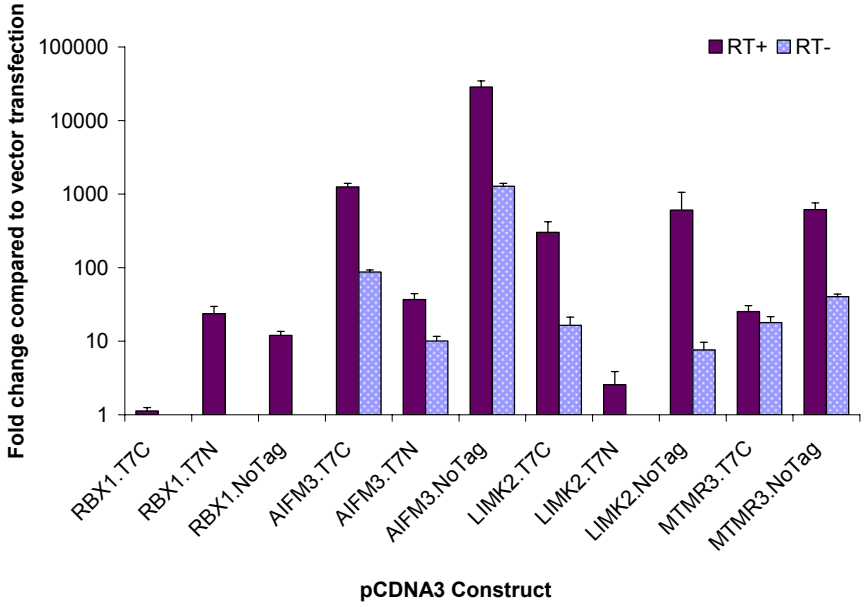

Figure 6.9 Expression of hit ORFs in cells transfected with ORF expressing constructs
RNA was isolated from cells transfected with constructs which express the 'hit' ORFs, either tagged with T7 epitope at the C or N terminal or not tagged, or pCDNA3.T7 as a negative control. RNA was either reverse transcribed to generate cDNA (RT+), or mock reverse transcribed in a reaction lacking reverse transcriptase (RT-). qPCR was performed on the samples using oligonucleotide primers designed to amplify either from the ORF or from GAPDH or ACTB. Expression levels were calculated using a variation of the Pfaffl method to allow normalisation to multiple housekeeping genes (Hellemans et al. 2007). Levels of GAPDH and ACTB in RT+ samples were used to normalize all samples. Results are presented as fold increase compared to levels in the RT+ sample from the negative control transfected cells. Error bars represent 1 standard error of the mean. n = 3.

In all cases, with the exception of the RBX1.T7N and MTMR3.T7C expressing constructs, transfection with ORF expressing constructs led to an increase in the level of the transcript for that ORF. The largest increase was the increase in the levels of AIFM3 transcript in pCDNA3.AIFM3.NoTag transfected cells, which showed a greater than 28,000 fold increase in transcript levels compared to negative control transfected levels. Such a huge increase in transcript levels suggests that levels in negative control transfected cells are very low. By contrast, transfection with pCDNA3.LIMK2.T7N causes a 2.5 fold increase in transcript levels compared to negative control transfected cells.

These data show that transfection with constructs expressing three of the four candidate hit ORFs (RBX1, AIFM3 and LIMK2) does lead to a reduction in sensitivity to TRAIL, and that this reduction in sensitivity is accompanied by an increase in transcript

levels.

6.5 Characterisation of hit ORFs

6.5.1 Effect on TRAIL-induced Caspase activity

The effect of three of the ORFs identified in the screen of ORFs from chromosome 22 on TRAIL-induced cytotoxicity was confirmed. To investigate the involvement of these ORFs in TRAIL-induced apoptosis, and also map the position in the pathway at which they function, the effect of transfection of constructs expressing these ORFs on TRAIL-induced activation of Caspases was measured using luminescencent caspase assays (Figure 6.10).

Treatment of negative control transfected cells with TRAIL induces a 2.5-fold increase in Capase-8 activity (Figure 6.10a), a 3.0-fold increase in Caspase-9 activity (Figure 6.10b) and a 19% increase in Caspases-3/7 activity (Figure 6.10c). Transfection with pSM2.shCasp8.2 reduces the level of TRAIL-induced Caspase-8 and Caspase-9 activity to 78% of that in negative control transfected cells treated with TRAIL and Caspases-3/7 activity to 67% of that in negative control transfected cells. Note that this reduction in Caspase-3/7 reduces levels of caspase activity in pSM2.shCasp8.2 transfected cells to below that seen in untreated negative control transfected cells, but not below that seen in untreated pSM2.shCasp8.2 transfected cells.

Transfection of none of the pCDNA3.ORF constructs reduced the level of TRAIL-induced activity of Caspase-8, Caspase-9 or Caspases-3/7. The level of Caspase activity measured in non-TRAIL treated samples is similar for Caspases-8 and Caspase-9 irrespective of the construct with which the cells were transfected. The level of Caspase-3/7 activity measured in untreated samples does vary depending on the construct transfected. However, since knock-down of Caspase-8 does not reduce levels of Caspase-8 activity measured in untreated samples, it is possible that these levels are assay background rather than a measure of some level of caspase activity in the absence of an activator.

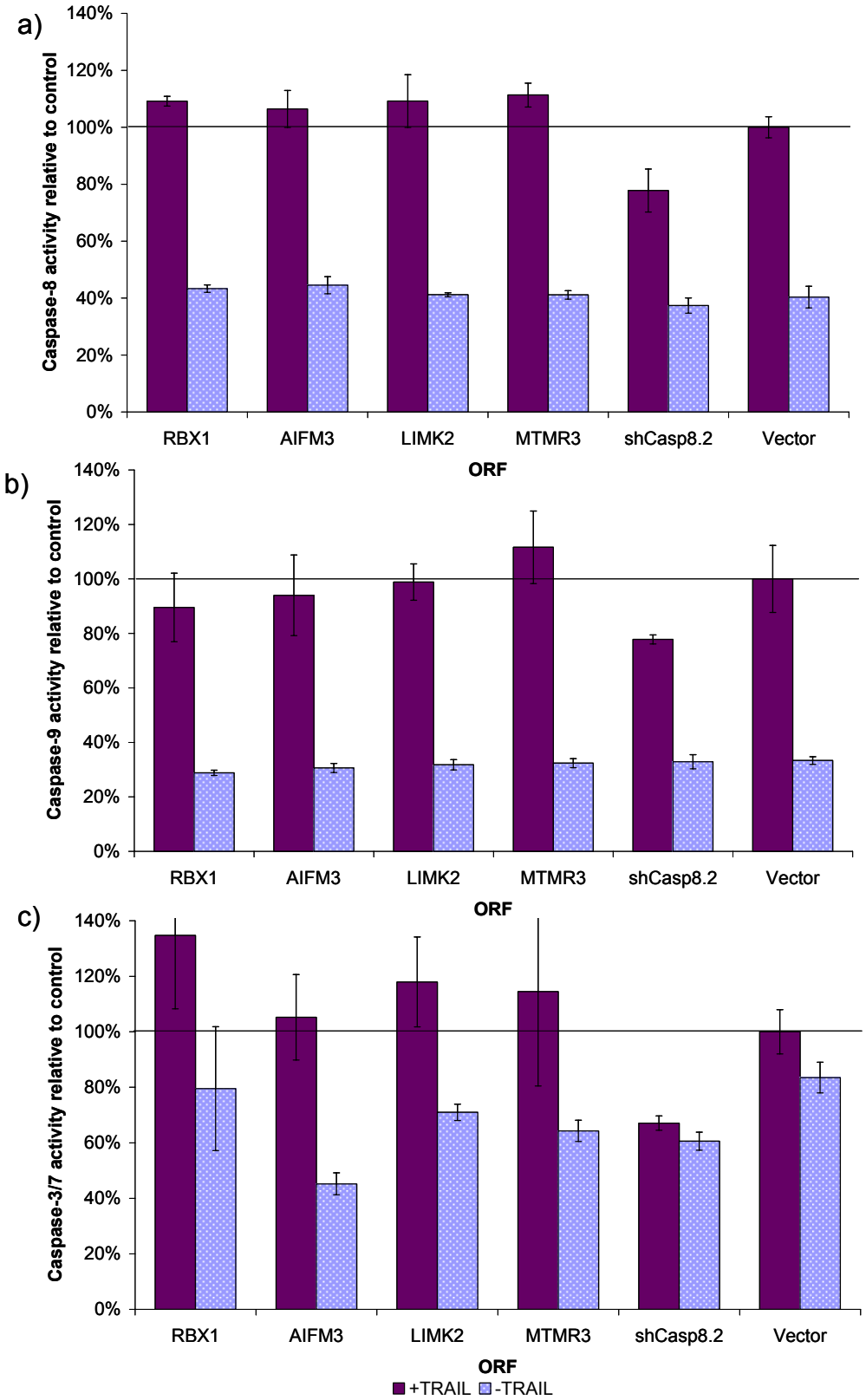

Figure 6.10 Effect of over-expression of hit ORFs on TRAIL-induced Caspase activity
Cells were transfected with pCDNA3.NoTag constructs expressing one of the hit ORFs or empty pCDNA3.T7 vector. 48 hours after transfection cells were treated with 0.5µg/ml TRAIL for 6 hours. a) Caspase-8, b) Caspase-9 or c) Caspase-3/7 activity was measured using Promega Caspase-Glo luminescent caspase assays. Resluts are expressed as percentage of caspase activity in TRAIL treated negative control transfected cells. In each case solid line represents 100% of control. Error bars represent 1 standard deviation, n = 3.

The assay used in the screen measures the effects of TRAIL-induced cytotoxicity rather than TRAIL-induced apoptosis. A key characteristic of apoptosis is the induction of Caspases. Evidence that expression of ORFs that reduce the level of TRAIL-induced cytotoxicity also reduce the levels of TRAIL-induced caspase activity would allow the conclusion that expression of such ORFs reduces TRAIL-induced apoptosis. No evidence was found for the ORFs identified from the screen and so no such conclusion can be drawn. One conclusion that could be drawn is that overexpression of these ORFs is affecting sensitivity of cells to TRAIL-induced cytotoxicity in a manner unconnected to apoptosis, although it must be noted that the data presented here do not necessarily demonstrate that.

6.5.2 Effect on sensitivity to other apoptosis inducing conditions

If expression of hit ORFs does in fact affect the sensitivity of cells to TRAIL-induced apoptosis, it could do at so several levels. They could affect the sensitivity of cells to ligand induced apoptosis (or purely TRAIL-induced apoptosis), or they could affect the sensitivity of cells to apoptosis in general. If this were the case it would be expected that expression of these ORFs would reduce the sensitivity of cells to non-ligand induced apoptosis as well as ligand induced apoptosis.

To test the sensitivity of cells to apoptosis induced by ligands other than TRAIL, the level of cytotoxicity induced by treatment with a range of concentrations of FAS was measured. Unfortunately minimal cell death was observed when negative control transfected cells were treated with any of the concentrations of FAS ligand tested (data not shown).

Exposure to H_2O_2 and UV radiation are two treatments with induce apoptosis through the intrinsic pathway, via oxidative stress and DNA damage respectively. In order to determine if the ORFs identified in the screen affect non-ligand induced apoptosis the effect of expression of hit ORFs on sensitivity to cytotoxicity triggered by these treatments was assessed (Figure 6.11 and Figure 6.12).

Treatment of cells with H_2O_2 caused cytotoxicity approximately proportional to the concentration of H_2O_2, with treatment with 100μM of H_2O_2 killing all of the cells and treatment with 50μM leading to a 36% survival rate. This was true for both negative control transfected and untransfected cells (Figure 6.11e). Expression of none of the hit ORFs changed the sensitivity of cells to H_2O_2 at any concentration tested (Figure 6.11a-d). The lack of a positive control in this experiment makes it difficult to assess the meaning of these results, as it is not possible to know if the experiment would show a clear effect for genes involved in the oxidative stress. An appropriate positive control would be a protein known to

be an inhibitor of the mitochondrial (intrinsic) apoptosis pathway, such as Bcl-2 or XIAP.

Exposure to UV radiation caused cytotoxicity to a large portion of cells, with 25% of negative control transfected cells surviving a 200 Jm^{-2} dose (Figure 6.12). Unfortunately, as was seen in similar experiments testing the effect of transfection of hit siRNAs from the kinase and phosphatase screen on the sensitivity of cells to UV irradiation (Figure 4.13), cells apparently not exposed to UV radiation showed some cytotoxicity in repsonse to mock treatment with 73% of cells surviving the mock treatment (Figure 6.12). The transfection of the negative control had no effect on the sensitivity of cells to UV irradiation compared with mock transfected cells (Figure 6.12e). The expression of MTMR3 has no effect on the sensitivity of cells to UV induced cytotoxicity at any dose tested (Figure 6.12d). Expression of RBX1, AIFM3 and LIMK2 had no effect at higher doses (100 and 200 Jm^{-2}) compared to the negative control (Figure 6.12a-c). A difference was observed at the lowest dose, with 75%, 78% and 84% of RBX1, AIFM3 and LIMK2 expressing cells respectively surviving exposure to 50 Jm^{-2} UV radiation, while 64% of negative controls cells survived. However, this difference was replicated in untreated cells, with 86% of RBX1 expressing cells, and 85% of both AIFM3 and LIMK2 expressing cells surviving the mock transfection, while 73% of negative control cells survived. If the supposedly untreated cells are receiving some dose of UV radiation due to leakage in the experimental protocol, this difference could be due a protective effect from the over-expression of the ORFs. However, if the cytotoxicity is due to some other artefact of the experiment, then the difference at 50 Jm^{-2} could be due to the same artefact that is causing the difference in mock treated cells. Again, the difficulty in interpreting the results is compounded by the absence of a positive control, which would show if a difference at higher doses of UV would be expected. This means it is not possible to draw any conclusion from these data other than it does not demonstrate a protective effect of expression of the hit ORFs on UV induced cytotoxicity.

6.6 Discussion and conclusions

In previous chapters an assay for the effect of gene perturbation on sensitivity to TRAIL-induced apoptosis was established. This assay was used to compare methods for siRNA-mediated screening and two RNAi screens assessing the effect of gene knockdown on sensitivity of cells to TRAIL were executed. In this chapter the assay developed was applied to an overexpression screen of 288 high-quality full length ORFs from chromosome 22. As with the RNAi screen of kinases and phosphatase, the screen presented here serves two functions. Firstly it serves as a gene discovery experiment. Secondly it serves as an

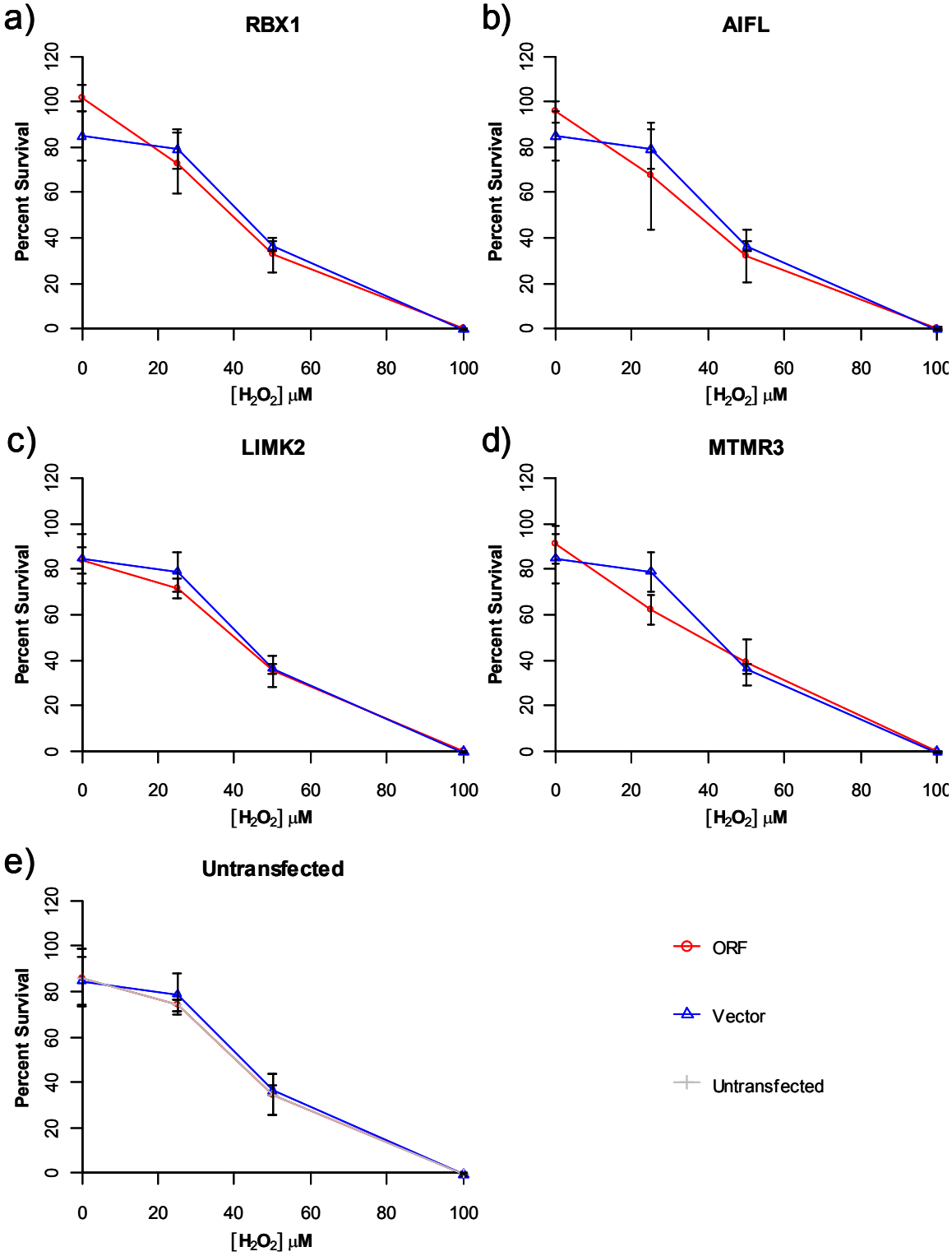

Figure 6.11 Effect of expression of hit ORFs on sensitivity of H_2O_2 induced cytotoxicity
HeLa cells were transfected with constructs expressing a) RBX1, b) AIFM3, c)LIMK2, or d)MTMR3 ORFs, , an empty pCDNA3.T7 vector, or e) mock transfected. 48 hours after transfection viability of cells was assessed and cells were treated with the concentration of H_2O_2 indicated. Viability was reassessed 24 hours later. After a further 24 hours viability was reassessed. Error bars represent 1 standard deviation, n = 3.

investigation of the usefulness of such gene-by-gene ORF overexpression experiments.

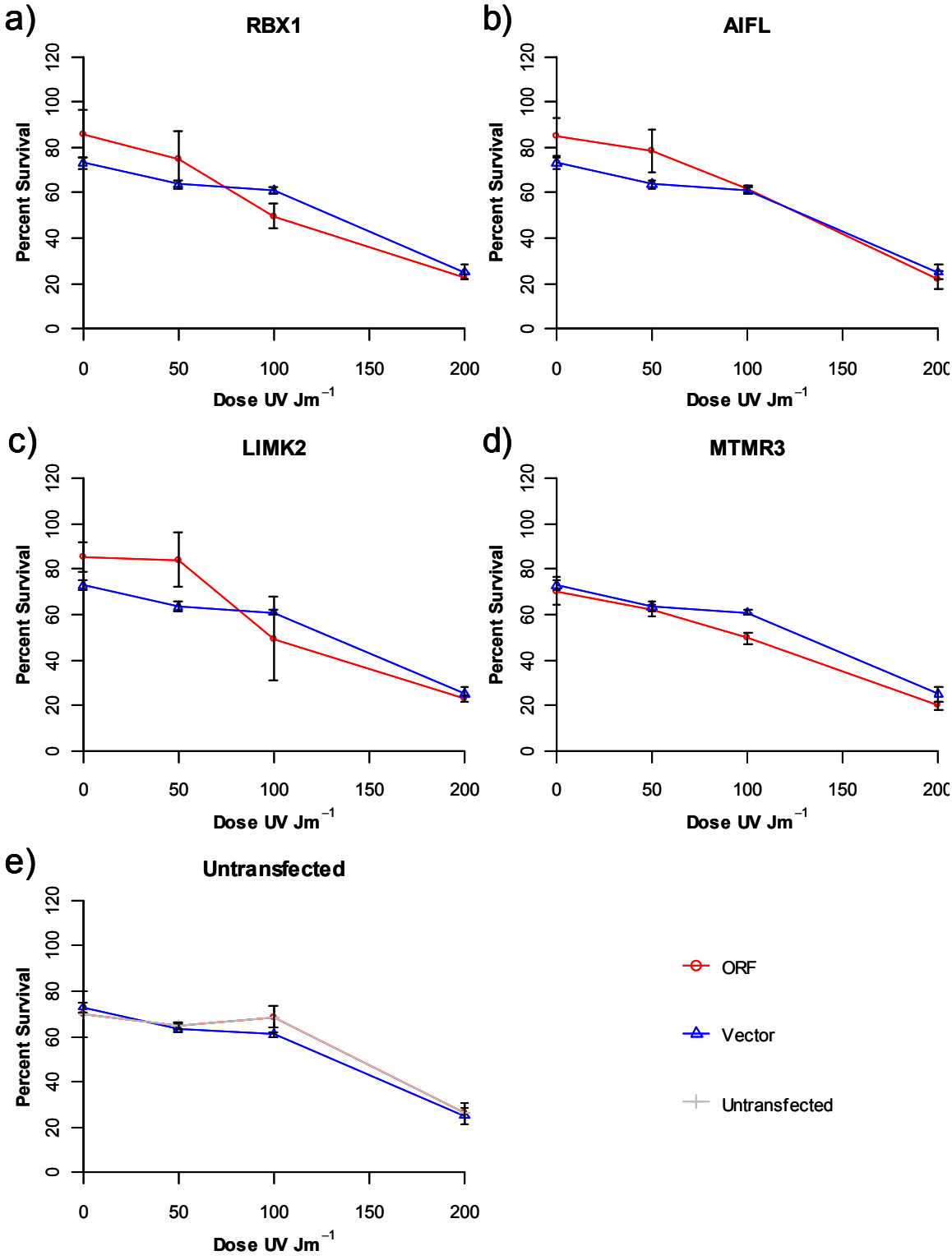

Figure 6.12 Effects of expression of hit ORFs on sensitivity of cells to UV irradiation
HeLa cells were transfected with constructs expressing the a) RBX1, b) AIFM3, c) LIMK2 or d) MTMR3 open reading frames, the empty vector pCDNA3.T7 or e) mock transfected. 48 hours after transfection viability of cells was assessed and cell were exposed to the dose of UV irradiation indicated. Viability was reassessed 24 hours later. Error bars represent 1 standard deviation, n=3

An expression screen of 288 ORFs from chromosome 22 was undertaken to identify

ORFs that could protect cells from TRAIL-induced apoptosis. Four constructs were identified from the screen whose effects clearly deviated from the distribution of effects for the majority of the constructs (Figure 6.6b). These four constructs expressed four different ORFs. When transferred to an untagged expression vector, expression of three of these four ORFs continued to show a significant effect on TRAIL-induced cytotoxicity (Figure 6.7), although only in two cases was this difference still significant when differences in the viabilities of untreated samples were taken into consideration. It was also shown that transfection with these constructs did lead to an increase in transcript levels (Figure 6.9).

Experiments examining the role of these genes in apoptosis failed to show any effect of expression of these ORFs on TRAIL-induced caspase activity or any effect on cytotoxicity caused by the non-ligand apoptosis induction by H_2O_2 or UV irradiation. However, the lack of appropriate controls makes the results of these experiments difficult to interpret. Without a positive control, it is impossible to say if the assays were sensitive enough to detect a change any chance in sensitivity to these apoptosis inducing conditions. Further, had a difference been detected, it would have been difficult to determine if the results were purely due to the over-expression of protein, since the negative control did not express any protein, although possibly in this situation MTMR3, which showed no activity in the TRIAL assay, could have served as a control for this.

6.6.1 The screen

In the previously described siRNA screens a correlation between the rank of the mean normalized survival for an siRNA and the standard deviation between the replicates was observed. The relationship between the rank of the mean normalised survival for an overexpression construct and the standard deviation between the replicates was examined. It was observed that constructs with a higher mean normalised survival tended to have a high standard deviation. Log transformation of the data reduced the strength of this relationship (Figure 6.3). This was supported by the finding that the majority of the scores from the screen calculated using log transformed data are normally distributed (Figure 6.6b).

The normalised survival observed for any given data point is affected by two factors. The first is the biological effect of overexpression of the ORF expressed by the construct transfected into the cells. The second is the random variation in the system. The low correlation seen between replicates of the same construct suggest that in the majority of cases here the random variation in the system dominates the biological effect of ORF expression (Figure 6.5a). The similarly low correlation between constructs expressing the

same ORF tagged at opposite termini could be due to differing biological effects of expressing the ORF tagged at one terminus compared to the other (Figure 6.5b). For example, the T7 tag could interfere with localization signals when fused to one terminus, leading to mis-localization of the protein. Alternatively, the bulk of the tag could interfere with protein folding. However, given the lack of correlation between replicates of the same construct, it seems likely that the lack of correlation here is also largely due to random variation. This lack of biological variation could suggest two things. Firstly it could suggest that the level of overexpression of the ORFs caused by transfection with the constructs is insufficient to trigger an effect that is greater than the level of random variation. In this way, the screening method outlined here would share the same problems as the shRNA screening method explored previously: a low/variable transfection efficiency and insufficient transcription from transfected constructs. Alternatively, it could be due to a genuine lack of ORFs which, when over-expressed, cause a reduction in sensitivity to TRAIL-induced apoptosis. Unfortunately in the absence of multiple repeats of negative and positive controls separate the random variation from the biological variation. Given that a TRAIL sensitive cell has a complete and functional pathway, it is necessarily the case that there will exist genes which when knocked down will disable this pathway. While genes may exist, which when overexpressed, actively inhibit the pathway, the existence of such gene is not necessary. Therefore, an idealised knock-down screen can always be expected to identify genes involved in the pathway, while this is not necessarily the case for an overexpression screen.

Given the conclusion that the random variation is dominating the biological variation, it is logical to assume that the largely normal distribution of scores resulting from the analysis of screening data represents random rather than biologically relevant variation (Figure 6.6a). The normality of this distribution supports the use of log transformation in analysis of survival values. The hits selected for confirmation are found outside this normal distribution, suggesting that the higher survival observed is due to biological, as well as random variation (Figure 6.6b). This conclusion is supported by the finding that three of the four ORFs selected for confirmation showed a significant effect on TRAIL sensitivity when expressed without a tag. The one clone that did not show a significant effect on TRAIL sensitivity was the lowest scoring of the four ORFs selected for confirmation (Table 6-1). However, since these results were not compared to the results of retesting a random selection of genes, it is not possible to definitively conclude that the screen performed better than random for selecting genes, the overexpression of which has an effect on TRAIL-induced cytotoxicity.

It was shown that transfection of constructs expressing ORFs identified in the screen led to an increase in ORF transcript levels (Figure 6.9) and that in three of the four cases transfection lead to an increased survival after treatment with TRAIL (Figure 6.7) compared to empty vector or a ORF which did not score highly in the screen, although in at least one case it is possible that this difference is due to a increase in cell viability in the absence of TRAIL. However, the biological significance of this finding is unclear. The experiments aimed at characterising the ORFs identified as hits from the screen and whose effect was confirmed in the following experiments failed to show that overexpression of these ORFs changed the activity of any of the caspases (Figure 6.10) or that expression of these ORFs altered the sensitivity of cells to any of the non-ligand inducers of apoptosis tested (Figure 6.11 and Figure 6.12). It is important to note that this lack of evidence for involvement of hit ORFs in regulating caspase activity or non-ligand induced cytotoxicity is not evidence for the lack of involvement, particularly given the lack of good controls. Such a situation was also observed when characterising the effect of siRNAs targeting Sharpin and MAST4. In each case one of the two siRNAs targeting these genes could be shown to have an effect on the sensitivity of cells to TRAIL, but not on caspase activity. In each case the siRNA inducing the smaller change in sensitivity to TRAIL-induced cytotoxicity was the siRNA which failed to induce a difference in TRAIL-induced caspase activity. One possible explanation is simply that the assays for caspase activity and assays for the effect of non-ligand apoptosis inducers are less sensitive than the assay for effects on TRAIL-induced cytotoxicity. This seems unlikely in the case of measurements of caspase activity, as transfection of pSM2.shCasp8.2 had a similar or smaller effect on TRAIL-induced cytotoxicity, but a significant effect on caspase activity. The lack of a positive control in the measurement of sensitivity to non-ligand inducers makes it difficult to draw conclusions as to the sensitivity of the assays measuring the sensitivity of cells under these conditions.

Caspase-dependent apoptosis is only one of several forms of programmed cell death, with others including caspase-independent apoptosis, autophagy and programmed necrosis (Reviewed: Assuncao Guimaraes, Linden 2004). There are several reports of TRAIL triggering caspase-independent cell death (Holler et al. 2000, Thon et al. 2006). It is possible that the ORFs identified in this screen are affecting some aspect of this caspase-independent cell death pathway. Alternatively these genes could be involved in the apoptotic pathway downstream of Caspase-3. In both cases it seems unlikely that all three genes identified in a screen should affect sensitivity to TRAIL in this way.

The final possibility is that the finding that overexpression of these ORFs reduces

sensitivity of cells to TRAIL-induced cytotoxicity has no biological significance in terms of the natural functioning of the TRAIL-induced apoptosis pathway. The result could be merely an artefact of the screening system, for example, simply due to the amount of overexpression and the unpredictable effects this has on the function of the cell. Indeed all overexpression studies suffer from similar problem. That a reduction in the level of some gene in a network affects the output of this network gives an indication of the natural function of the network. However, an effect on the network output of introducing a novel factor into the network does not necessarily say anything about the natural function of the network.

6.6.2 The Hits

The above discussion not withstanding, three ORFs were identified that when overexpressed, did lead to a reduction of the sensitivity of cells to TRAIL-induced cytotoxicity (although possibly not TRAIL-induced caspase-dependent apoptosis).

6.6.2.1 RBX1

RBX1 is a RING-finger protein and member of the SCF E3 ubiquitin ligase complex (Ohta et al. 1999). Ubiquitin ligase complexes catalyse the addition of ubiquitin to their targets, marking them for degradation by the proteasome. Proteasome inhibitors are known to sensitise resistant cells to TRAIL-mediated apoptosis (Ganten et al. 2005). RBX1 has been shown to catalyse the ubiquitination of IkB, which would lead to an activation of NF-κB(Ohta et al. 1999). RBX1 has also been shown to bind procaspase-3 leading to its ubiquitination and degradation. Overexpression of RBX1 leads to a reduction of the steady state levels of procaspase-3 and its knock-down leads to a sensitization of cell to TRAIL-induced apoptosis (Tan et al. 2006).

If overexpression of RBX1 protects against TRAIL-induced apoptosis, as shown here (Figure 6.7), via the increasing the ubiquitination of pro-caspase-3 and therefore reducing its levels, it is unclear why no decrease in TRAIL-induced Caspase-3 activity was observed, and why RBX1 overexpression did not offer protection against non-ligand inducers of apoptosis as no change in Caspase-3/7 activity was seen (Figure 6.10,Figure 6.11 and Figure 6.12). If this is indeed the case it suggests a defect in the follow up experiments, despite the fact that the positive control showed an effect. Possibly this suggests that the positive control used in the caspase activity experiments was unsuitable. Despite these discrepancies the finding that one of the genes identified in the screen has previously been associated with TRAIL-induced apoptosis lends confidence to the idea that ORF

overexpression screening can be used to identify genes involved in TRAIL-induced apoptosis.

6.6.2.2 AIFM3

The AIFM3 gene encodes the Apoptosis Inducing Factor; Mitochondrial associated 3 protein (also known as AIF like). AIF is released from the mitochondria along with cytochrome c and DIABLO upon death signalling. Overexpression of AIF triggers a caspase-independent form of apoptosis (Joza et al. 2001, Moubarak et al.). However, the role of AIF in apoptosis control is more complicated, since in some cell types down regulation of AIF also sensitizes cells to apoptosis induction by cellular stress, but not apoptosis inducing ligands. The protection afforded to cells by AIF is dependent on the pyr_redox domain of the protein, which is responsible for its reactive oxygen species (ROS) generating NADH oxidase function (Urbano et al. 2005). AIFM3 is 35% similar to AIF with that similarity mostly residing in the region homologous to the pyr_redox domain. Overexpression of AIFM3 has been shown to induce apoptosis in a cytochrome c and Caspase-3 dependent manner. However, this apoptosis induction was dependent on the Riseke domain of the protein, which isn't found in AIF (Xie et al. 2005). No anti-apoptosis role for AIFM3 has been reported, although since it contains the domain of AIF which is responsible for that protein's anti-apoptosis activity, it is possible that AIFM3 also has both pro- and anti-apoptotic activity.

6.6.2.3 LIMK2

LIM kinases are a family of kinases which regulate actin cytoskeleton dynamics in response to several stimuli, particularly the Rho effector kinase ROCK (Reviewed: Scott, Olson 2007). Other functions for LIMK have been demonstrated including LIMK2 activation of cyclin A1 (Croft, Olson 2006) and involvement in the spindle assembly checkpoint (Sumi et al. 2006). Overexpression of LIMK2 has been reported to lead to membrane blebbing reminiscent of that seen in apoptosis (Amano et al. 2001). LIMK1 but not LIMK2 contains a Caspase-3 target sequence and has been reported to be a target of Caspase-3 (Tomiyoshi et al. 2004), but there are no reports of either LIMK being involved in regulation of apoptosis.

Examination of the expression patterns reported by the GNF expression atlas does not support a correlation between LIMK2 and TRAIL sensitivity. The two cell lines from the NC160 panel of cell lines showing the strongest expression of LIMK2 are COLO205 and

SK-MEL-28 (http://symatlas.gnf.org) , which are respectively very sensitive and insensitive to TRAIL-induced apoptosis (Bae et al. 2007, Lippa et al. 2007).

6.6.3 Conclusions

The aims of this screen were two-fold, firstly to identify novel genes involved in the TRAIL-induced apoptosis, and secondly to assess the general usefulness of the approach taken to identifying genes involved in biological pathways.

Constructs, representing each of the 288 ORFs fused to a T7 epitope at both the C and N terminals, were transfected into cells in duplicate and the sensitivity of transfected cells to TRAIL measured. The data was normalised to the median of each plate, log transformed and standardized. The minimum of the two repeats was taken as the score for each construct.

A low correlation between the normalized survivals of the two replicates demonstrates that in the majority of cases the random variation in the system is greater than any biological effect. This shows that either overexpression of most ORFs has no biological effect, or that the system is not sensitive enough to pick any genuine biological signal present. This is in contrast to the results from RNAi screens where a strong correlation exists between the two replicates performed, suggesting that many siRNAs have a varying degree of influence on the sensitivity of the cell to TRAIL-induced cytotoxicity, although this could simply to a measure of the number and strength of off-target effects elicited by the siRNA in question.

Transfection of 4 constructs caused an effect which was significantly outside the distribution of effects of the other constructs, although in the absence of a positive control in the screen it is impossible to comment on the size of these effects. The ORFs from these constructs were transferred to an expression vector which did not fuse a T7 epitope to the ORF. Transfection of three of these four constructs led to a significant reduction in sensitivity to TRAIL-induced cytotoxicity, although in one case this was not significant when pre-treatment viability was taken into account. However, the biological interpretation of these results is difficult since transfection of these constructs did not change the level of TRAIL-induced caspase activation for any of the caspases. Transfection of none of these constructs altered the sensitivity to the non-ligand apoptosis inducers H_2O_2 and UV radiation. It is unclear in these cases whether the characterisation experiments are faulty or if the original results were an artefact of the assay with no biological significance, or if the

ORFs identified are involved in the some pathway parallel to the caspase-dependent apoptosis pathway or act down-stream of the caspase cascade. This is particularly true given the lack of good controls in these experiments.

One of the genes identified, RBX1, a member of the ubiquitin ligase complex SCF, has been previously shown to be involved in the regulation apoptosis through regulation of the stability, and therefore steady state level of, Caspase-3. This shows that the screening paradigm presented here is capable of identifying true regulators of the apoptosis pathways. This finding suggests that the characterisation experiments were not sensitive enough to capture the role of this gene in the regulation of Caspase-3 activity or apoptosis inducing agents other than TRAIL. The second gene, AIFM3, is a member of the apoptosis inducing factor family. AIF has been shown to be have both pro and anti-apoptotic activity. Over-expression to AIFM3 has been shown to induce apoptosis, but is homologous to AIF in the domain that is responsible for AIFs anti-apoptotic activity. The third gene, LIMK2, is a kinase involved in controlling remodelling of the actin cytoskeleton. It has no previously reported affect on regulation of apoptosis. That a plausible story can be constructed for the involvement of two of the three hit genes suggests that the system maybe capable of identifying genes involved in the process. However, the construction of a plausible story for the involvement of these genes does not constitute evidence of their involvement.

Taken together these results suggest that ORF-by-ORF over-expression screening can identify ORFs that affect TRAIL-induced cytotoxicity, although the level of noise is high and it is impossible to say if the identification of hit is better than selecting genes at random. . They show that interpretation of the biological significance of these results can be difficult, including whether these hits are involved in TRAIL-induced apoptosis, or simply TRAIL-induced cytotoxicity. Even if a role in the pathway may be defined for the overexpressed ORFs *in vitro*, it is unclear whether this role has any relevance to the pathway *in vivo*