5
Zebrafish: At the Nexus of Functional and Chemical Genomics

ANDREW DODD[1], DAVID R. GREENWOOD[2], ANDREW L. MILLER[3], SARAH E. WEBB[3], STEPHEN P. CHAMBERS[1], BRENT R. COPP[4] AND DONALD R. LOVE[1]*

[1]*School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1001, New Zealand,* [2]*Gene Technologies, HortResearch, Mt. Albert Research Centre, Private Bag 92169, Auckland 1003, New Zealand,* [3]*Department of Biology, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, PRC and* [4]*Department of Chemistry, University of Auckland, Private Bag 92019, Auckland 1001, New Zealand*

Introduction

The study of biological pathways and processes implicated in human disease is the driving force behind a significant amount of biological research that is undertaken today. These studies are based on the expectation that a detailed knowledge of the course and pathogenesis of disease will lead to improved treatment, or at least preventative measures. This expectation has been realized to a large extent for a range of nutritional and infectious diseases. Significant attention is now being focused in the area of genetic disease. The difficulty in developing cures or control approaches for genetic disease highlights the extreme complexity of disease processes and the difficulty in restoring homeostasis in a disease situation (Murphy, 1985).

The above statement can be illustrated in *Figure 5.1*, where three distinct states can be proposed: homeostasis, diseased, and treated. The fundamental problem is our ability to understand and unravel the processes that are disrupted or involved in disease, and the effects of treatments because we are generally dealing with a level of complexity outside the range of current techniques and understanding. The constant development of platforms for disease analysis is essential for us to realize any goal of

*To whom correspondence may be addressed (d.love@auckland.ac.nz)

Abbreviations: DGC, dystrophin glycoprotein complex; DIGE, differential in-gel electrophoresis; DMD, Duchenne muscular dystrophy; hpf, hours post fertilization; MALDI, matrix-assisted laser-desorption; MO, morpholino oligonucleotide; MS, mass spectrometry; PCR, polymerase chain reaction; qPCR, quantitative PCR; siRNA, short interfering RNA.

Biotechnology and Genetic Engineering Reviews – Vol. 22, January 2006
0264-8725/06/22/77-99 $20.00 + $0.00 © Lavoisier/Intercept, 14 rue de Provigny, F-94236 Cachan cedex, France

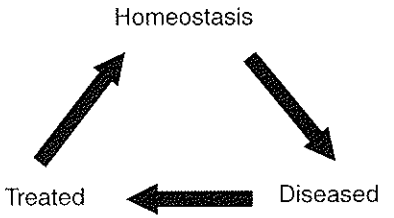

Figure 5.1. The three states of disease. Understanding the relationship between a diseased state and homeostasis is essential in developing effective treatments, while the relationship between the treated state and homeostasis is central to the effectiveness of the treatment.

effective treatment for genetic disease. Central to this process is the development of animal models to provide a platform for experimentation and analysis. The pivotal nature of the model organism needs to be considered carefully in order to allow a detailed understanding of human disease processes and biological pathways.

Genotype/phenotype relationships are not necessarily conserved among different organisms (Elsea and Lucas, 2002). There is essentially a reliance on some degree of species conservation in the biological processes that are being studied; the challenge is to identify a conserved response/process from one that is specific to a particular organism. One approach to deal with this challenge is to undertake comparative disease 'modelling' and analysis (Rissman, 2004).

COMPARATIVE MODELLING

When making comparisons between different organisms, we are dealing with diversity combined with the underlying conservation of fundamental biological processes. In using one organism to make inferences about another, it is necessary to determine how closely related they are in order to assess the potential for displaying appropriate or useful similarities. This determination has usually adopted the approach of constructing an 'evolutionary' tree, i.e. displaying relatedness on the basis of the predicted divergence of the different organisms. While this makes sense and generally reflects their similarity based on some easily observable characters, like morphological complexity, we are now beginning to realize that this certainly does not hold true for all comparisons. This realization adds uncertainty to the scientific rationale underlying the choice of a 'model' organism in order to make inferences of functional similarity with another.

Recently, it has been proposed that a more broad-based comparative modelling approach should be used to marginalize organism-specific influences, when they exist, and to reveal conserved biological processes. In this respect, the identification of gene co-expression networks offers a means of revealing this conservation. The complete genome sequences of humans and several model organisms have established a nearly complete list of genes required for cellular, developmental, and behavioural processes (Goffeau *et al.*, 1996; Myers *et al.*, 2000; Lander *et al.*, 2001; Venter *et al.*, 2001). These data have been used to undertake microarray-based analysis in order to highlight regulatory relationships among genes. The concept of gene co-expression networks has been used to uncover and identify conserved

processes and groups of genes across different model organisms (Stuart *et al.*, 2003). The idea behind this approach is that genes encoding for proteins that participate in the same pathways, or are part of the same protein complex, are often co-regulated. Furthermore, genes encoding for related functions often exhibit expression patterns that are correlated under a large number of diverse conditions (Stuart *et al.*, 2003). However, it is important to recognize that apparent co-regulation of genes does not necessarily imply that genes are functionally related, and that trying to dissect these relationships in a single species is difficult, as it is nearly impossible to distinguish coincidentally regulated genes from those that comprise an inter-related sub-network. The concept of evolutionary conservation is a powerful criterion to identify genes that are functionally important from a set of co-regulated genes. Co-regulation of a pair of genes over a large evolutionary distance possibly implies that the co-regulation confers a selective advantage, most likely because the genes are functionally related.

The gene co-expression network approach to identifying gene function has been suggested to be far more sensitive than scoring phenotypes resulting from loss-of-function mutants (Stuart *et al.*, 2003). The unpredictable nature of the response to loss-of-function experiments in different organisms illustrates the need for a comparative approach when trying to make inferences from studies in model organisms. A major limitation in realizing the advantages of a broad-based disease modelling approach is the limited number of existing model organisms that can be used in a comparative study. Whilst it is possible to make comparisons between organisms as distant as *Drosophila* and mice, modelling based on vertebrate model species is essential for developing a comparative analysis platform for human disease.

Zebrafish: a critical vertebrate model system

The zebrafish is the most popular vertebrate model species after the mouse and the rat. Although the zebrafish has been predominantly used to study vertebrate development, its potential as a platform for disease analysis and drug discovery is slowly being realized (Zon, 1999; Barut and Zon, 2000; Dodd *et al.*, 2000; Dooley and Zon, 2000; Pichler *et al.*, 2003). The zebrafish offers the potential to play a key role in a comparative disease analysis approach. However, significant technical developments need to be made to allow researchers to perform experiments in zebrafish that mimic those that can be performed in mice. This lack of technical parity with the mouse is a considerable hurdle to the zebrafish being fully accepted as a disease analysis platform.

The zebrafish is a small, freshwater, tropical fish species native to India (Eisen, 1996). The most significant feature of the zebrafish is that they have optically transparent embryos. An individual female can produce hundreds of eggs in each clutch, allowing large numbers of progeny to be generated. The eggs are large enough to be easily manipulated, and can be fertilized naturally or *in vitro*. The eggs develop rapidly *ex utero*, becoming free-swimming larvae within 4–7 days, and have a generation time of 3–4 months. The external development of the embryos lends itself to the easy observation of organogenesis and other developmental processes. Individual cells can be clearly resolved *in vivo* using simple microscopic techniques across a broad range of developmental stages. It is therefore possible to

track the fate of individual cells as they divide, migrate, and differentiate (Kimmel and Warga, 1988). Many techniques have been developed to achieve this, including labelling cells with lipophilic dyes or fluorescent tracers (Westerfield and Eisen, 1988; Eisen, 1991; Ho and Kimmel, 1993). More recently, transgenic lines have been constructed expressing fluorescent reporter genes, such as green fluorescent protein (GFP), under the control of tissue specific promoters, allowing the detailed observation of tissue development (Ju *et al.*, 1999). These techniques have expanded into different methods for assaying gene function. The microinjection of synthetic mRNA can be used during embryogenesis in gain-of-function assays (Hyatt and Ekker, 1999). Other methods, such as RNA caging, can be used at later stages to control the temporal and spatial expression of a gene (Ando *et al.*, 2001).

A key aspect in the development of zebrafish as a model organism has been the development of genetic and genomic resources. The zebrafish genome is 1.7 gigabases in size, which is slightly more than half that of other vertebrates, such as human and mouse, and is divided into 25 pairs of chromosomes. There are regions of synteny between the chromosomes of fish and mammals, which supports the assumption that there will be zebrafish orthologues for most human genes (Barbazuk *et al.*, 2000). There is believed to have been a genome duplication at the base of the teleost radiation 450 million years ago (mya), as well as a second duplication 100 mya, with as many as 20% of genes having been duplicated in fish compared to humans (Postlethwait *et al.*, 1998, 2000, 2004). It is thought that, over time, genes that have been duplicated have also functionally diverged. Many possible outcomes exist for duplicated genes. One outcome is where the function of ancestral genes is split between the two paralogous genes, which has been shown for Na(+), K(+)-ATPase, *Hox* and *Pax* genes (Force *et al.*, 1999; Postlethwait *et al.*, 2000, 2004; Serluca *et al.*, 2001). An alternative outcome is that one copy becomes inactivated, resulting in a pseudogene.

Significant effort has been spent in constructing physical and genetic maps of the zebrafish genome. Genetic maps have been constructed using both microsatellites (Knapik *et al.*, 1998; Postlethwait *et al.*, 1998; Shimoda *et al.*, 1999) and genes (Hukriede *et al.*, 1999; Kelly *et al.*, 2000; Woods *et al.*, 2000). The sequencing of the zebrafish genome was initiated in February 2001 by the Sanger Centre using a two-pronged approach: clone-based sequencing from BAC and PAC libraries, and whole genome shotgun sequencing from plasmids (http://www.sanger.ac.uk/Projects/D_rerio/). Further sequence information is being generated by the Washington University Zebrafish Genome Resource Project (http://zfish.wustl.edu/), which has developed cDNA libraries for EST sequencing.

DISEASE MODELLING USING THE ZEBRAFISH

Three main approaches exist for attempting to model human disease in the zebrafish. First, mutants can be isolated that have similarity to disease at either the genotype or phenotype level. These mutants have been reviewed extensively (Amatruda and Zon, 1999; Barut and Zon, 2000; Dooley and Zon, 2000; Shin and Fishman 2002; Berman *et al.*, 2003). Although a genetic screen is a powerful technique for identifying genes involved in specific pathways leading to specific morphological differences, the identification of the mutated gene can be a laborious process

involving positional cloning. Recently, screening methods have been developed for identifying specific mutations after random mutagenesis in a process termed target-selected gene inactivation (Wienholds *et al.*, 2002), although this approach is not widely accessible.

The second approach in generating human disease models is by transgenesis, or homologous recombination (HR)-based gene targeting; the latter technique typifying the standard approach taken in mice. Some successes have been reported in generating models of neurodegeneration and T-cell leukaemia in zebrafish through transgenesis (Tomasiewicz *et al.*, 2002; Langenau *et al.*, 2003). However, the generation of targeted gene knockout or knockdown in the zebrafish for disease analysis has yet to be realized due to the lack of stable gene targeting techniques. Although ES-like cells from zebrafish have been isolated (Ma *et al.*, 2001), the preference of non-homologous DNA-end joining over HR in zebrafish raises some doubt over the efficiency of HR-based gene targeting (Hagmann *et al.*, 1998; Dodd *et al.*, 2004b).

The third approach that has received significant attention is that of transient gene targeting. This approach has been applied to the study of development and signalling pathways, but has also been applied to disease modelling.

Transient gene targeting in the zebrafish

A significant development in transient gene targeting in the zebrafish has been the use of antisense morpholino-modified oligonucleotides to block the translation of specific genes (Nasevicius and Ekker, 2000). Morpholinos do not elicit RNAse H mediated mRNA degradation, but hinder ribosome binding to a target transcript, thus blocking translation (*Figure 5.2*). This technique has revolutionized gene-targeting in zebrafish, and has allowed researchers to target genes transiently during the first few days of development (Ekker, 2000; Heasman, 2002). Although this approach has proved of great benefit, several significant limitations exist. Most gene-targeting experiments utilizing morpholinos rely heavily on phenotype as a measure of targeting success. While this approach is reasonable, given the existence of known phenotypes associated with characterized mutants, it becomes more problematic if one is unsure about the expected phenotype; a circular problem if you are using gene-targeting to help uncover the function of genes.

It is important that inferences of genotype/phenotype relationships that are made based on the use of morpholinos involve an analysis of specificity between a target gene and a morpholino. Here lies one of the significant problems with targeting *via* translation inhibition. To measure the specificity and efficacy of targeting, protein-based studies must be undertaken. These studies would require either enzyme-specific assays, protein-specific antibodies, or the use of cross-reacting antibodies from another species. Unfortunately, a lack of these resources in the majority of cases limits the level of testing that is generally performed in the zebrafish. The problem in confirming gene-targeting specificity is further exacerbated by the use of scrambled or mismatched morpholinos as controls. Unreported data of non-specific phenotypic effects using scrambled/mismatched morpholino controls, and even using morpholinos designed against gene targets, give cause to consider how best to substantiate proposed genotype/phenotype relationships. Even with these

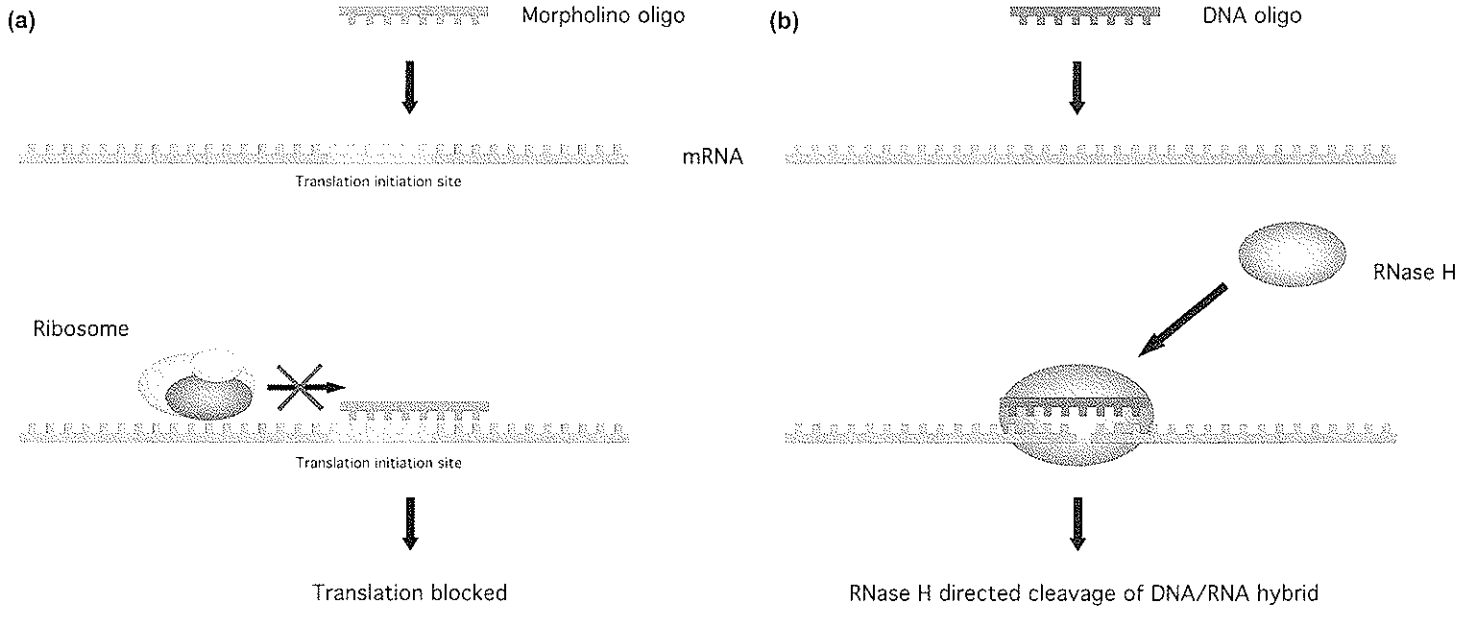

Figure 5.2. Antisense mechanisms. a) Morpholino-modified oligonucleotides inhibit gene expression by blocking translation initiation of mRNAs. b) DNA and RNA oligonucleotides inhibit gene expression by eliciting RNase H-mediated degradation of the target mRNA.

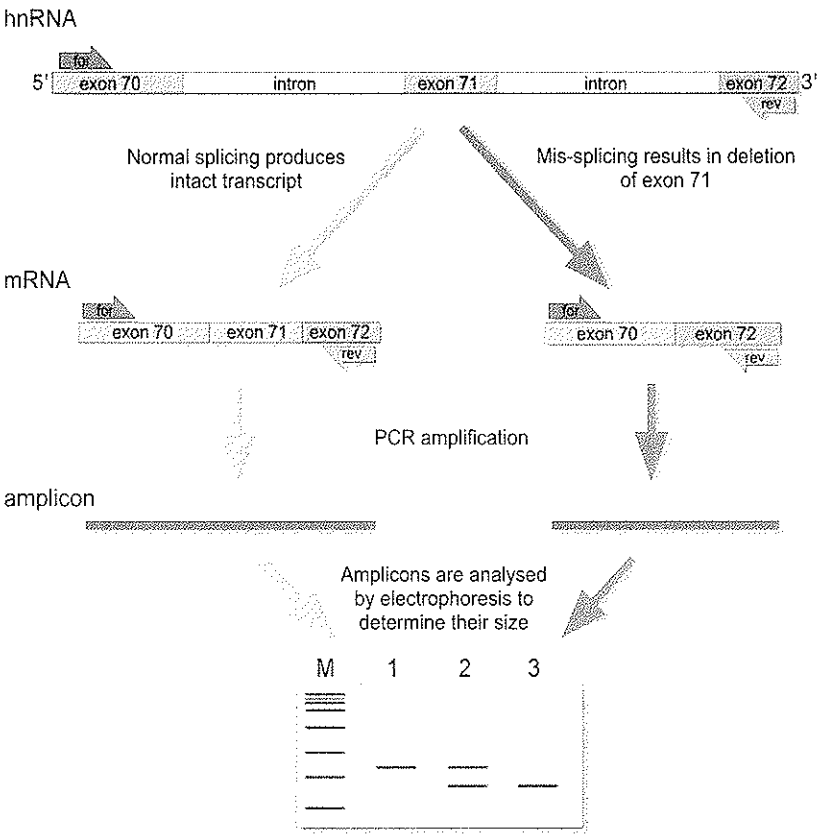

Figure 5.3. RT-PCR assay to analyse splicing events. This flow diagram illustrates the analytical process to detect morpholino-directed mis-splicing of a nuclear transcript (hnRNA). To determine if mis-splicing has occurred, mRNA is isolated and amplified using primers designed to bind exons flanking the targeted exon (arrows labelled 'for' and 'rev'). The PCR products are electrophoresed in an agarose gel and the size of the resulting amplicon(s) is estimated against a marker (M). Lane 1 of the conceptual gel illustrates normal splicing of the amplicon, the presence of a smaller fragment in lane 2 indicates that mis-splicing has been detected together with the normally spliced amplicon, whereas in lane 3 only a mis-spliced amplicon is detected.

limitations, morpholinos remain a popular tool for transient gene targeting in the zebrafish.

Recently, a different approach has been taken with antisense oligonucleotides, which involves splice site targeting (Draper *et al.*, 2001; Xu *et al.*, 2002; Yan *et al.*, 2002; Dodd *et al.*, 2004b). This approach uses antisense oligonucleotides to target either a splice acceptor or donor site, resulting in mis-splicing of the nuclear transcript, which can cause an in-frame deletion or frame-shift mutation. The efficacy of this approach can be measured at the transcript level by RT-PCR (*Figure 5.3*). Although this is an interesting technique that offers the ability to make specific changes to the transcript, more comparative studies are required to determine the

optimum backbone modification of the antisense oligonucleotide in order to achieve stability and target-specificity (Dodd *et al.*, 2004b).

The desire to regulate the expression of specific genes in a temporal and tissue-specific manner, and the need to target multiple genes at once, is driving researchers to consider alternative techniques to achieve gene targeting (Koopman, 2003). The first technique to offer the potential of regulated/tissue-specific gene targeting is that of gene silencing using double-stranded RNA (dsRNA). The use of gene silencing has revolutionized gene targeting in *C. elegans* and *Drosophila*, and is considered an exciting alternative to the use of homologous recombination in generating traditional gene knockouts in mice (Buckingham *et al.*, 2004).

DOUBLE-STRANDED RNA-BASED GENE TARGETING IN THE ZEBRAFISH

The use of dsRNA gene targeting has not gained wide acceptance in the zebrafish community, possibly because of early concerns about non-specific effects (Oates *et al.*, 2000). These experiments used long dsRNA, but recent developments in our understanding of gene silencing mechanisms in mammalian systems suggest that off-site effects using long dsRNA might be common to vertebrates. In this respect, short interfering RNAs (siRNAs) appear to exhibit efficacy in terms of mammalian gene targeting, while avoiding the off-site effects seen with long dsRNA (Hannon and Rossi, 2004; Meister and Tuschl, 2004).

In the absence of constructs that might enable the expression of siRNAs in zebrafish, this species lends itself to the direct microinjection of siRNAs. However, testing the efficacy of this approach, which is essentially uncontrolled and transient, requires the consideration of at least two key issues (*Figure 5.4*). The first issue concerns the time at which the gene silencing machinery becomes active during embryo development. There are several lines of evidence to suggest that this activation might happen very early in development. The first line of evidence comes from previous gene silencing experiments performed in zebrafish. Wargelius and colleagues targeted the *pax 2.1* gene and demonstrated a significant reduction in target RNA at 11 hpf (Wargelius *et al.*, 1999), and Li and colleagues targeted the zebrafish *ntl* gene and demonstrated a significant reduction of *ntl* mRNA at 12 hpf (Li *et al.*, 2000). Both these studies indicate that embryos have gene silencing potential at 11–12 hpf. The second line of evidence comes from the involvement of the gene silencing machinery in microRNA-based gene regulation mechanisms. Although it is uncertain what role RNA-based regulation plays in developmental pathways in zebrafish, it is likely to be a significant role (based on studies in other vertebrates). The second key issue in assessing the suitability of assaying siRNA efficacy in microinjected zebrafish embryos is the change in concentration of siRNAs as the embryo develops. The dilution and degradation of siRNAs will limit the developmental window where effective gene silencing can be achieved. The specific dynamics of these two processes might be expected to play a role in the assessment of siRNA-based gene silencing using an embryo-based transient assay.

TESTING siRNAS: PHENOCOPYING V. TRANSCRIPT ANALYSIS

A common approach to determine the effectiveness of siRNA-based gene targeting

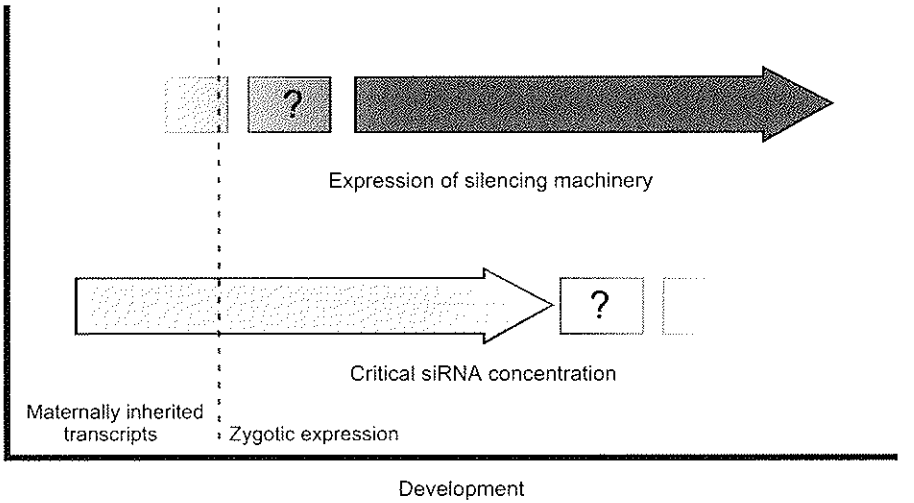

Figure 5.4. Critical gene silencing parameters. The expression of gene silencing machinery and the dilution of siRNAs over time are represented as critical factors in testing siRNA-based gene targeting.

in zebrafish embryos could involve phenocopying well-recognized developmental mutants. This approach was central to the development and demonstration of antisense techniques in zebrafish (Nasevicius and Ekker, 2000). Although popular, the suitability of this approach for measuring gene targeting should be carefully considered. The first issue is that phenocopying generally lacks any quantitative value. It is generally unclear what level of gene targeting is required to produce a specific phenotype, or what dynamic might exist in a particular genotype/phenotype relationship. This dilemma can be illustrated using two contrasting examples (*Figure 5.5*).

In the first example (*Figure 5.5a*), we consider targeting a gene that is acting in a system as a simple genetic switch. For the switch to turn on, transcript levels must accumulate to a level above a threshold at a particular time during development. In this instance, transcript levels must fall below a nominal threshold value within a narrow window of temporal expression in order to detect an altered phenotype compared to normal embryos. This situation requires limited gene targeting activity for an effect to be detected, with the phenotype possibly lasting long after the gene targeting activity has finished. In the second example (*Figure 5.5b*), we consider targeting a gene that has a dynamic genotype/phenotype relationship. Similar to the previous example, the targeted transcript level must fall below a threshold in order to interfere with the functional readout of the gene, but this reduction must be sustained for a phenotypic outcome to be detected. What is clear is that these two situations potentially require quite different levels of gene targeting to generate a mutant phenotype. Possibly, the first situation might make a gene targeting technique look more successful than the second. These examples illustrate that the ability to phenocopy a particular mutant is heavily influenced by its genotype/phenotype relationship.

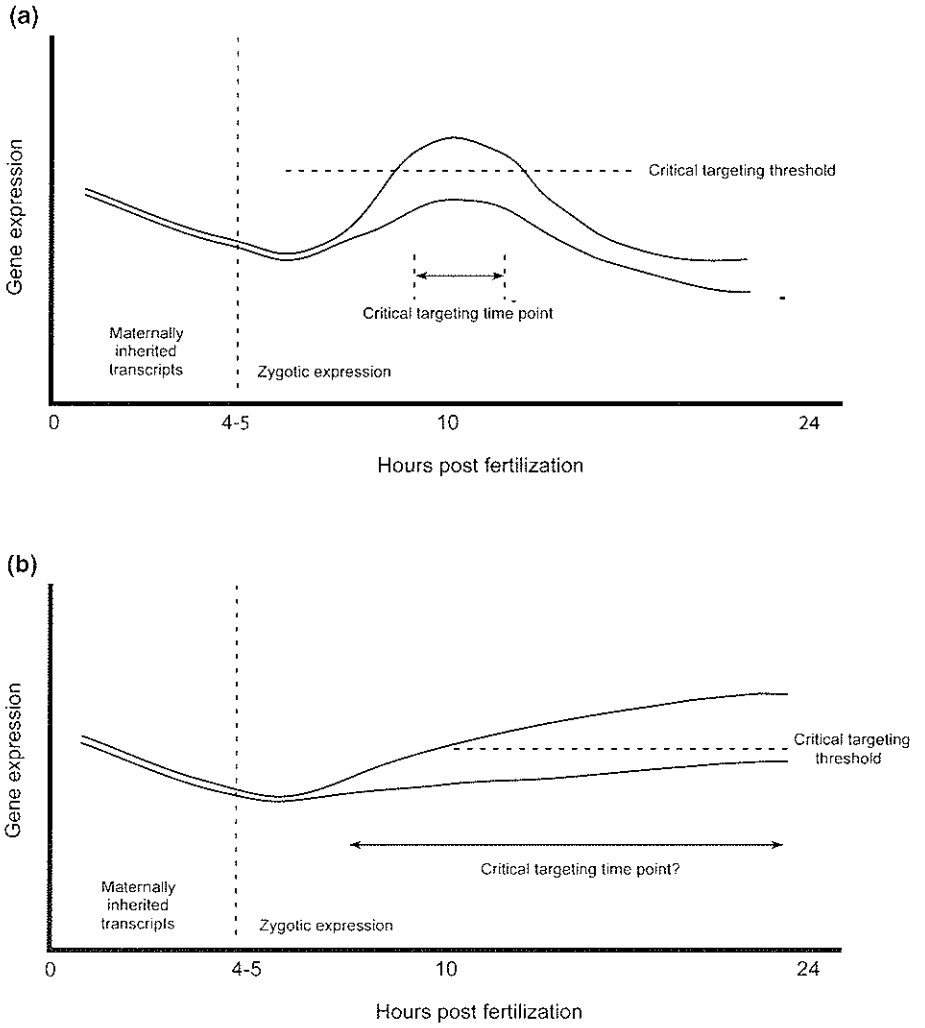

Figure 5.5. Gene targeting dynamics. The potential of transient gene targeting during early development depends on the expression profile of the target gene. A gene (a) with a restricted temporal expression profile requires less targeting activity compared to a gene (b) that is continuously expressed. Plots represent potential mRNA profiles of normal and treated embryos.

Unfortunately, usually little is known about the processes that influence the expression of a particular genotype, and thus what events lead to a particular phenotype. In this situation, it is difficult to use a phenotypic measure, which is the result of a highly complicated and largely unknown series of processes in order to make direct quantitative inferences about a genotypic or transcriptional state.

An alternative to the phenocopying approach is to directly measure the gene targeting effect. In this respect, the fact that siRNAs affect the integrity of a gene's transcript enables quantitative PCR (qPCR) assays to be undertaken using primers

that flank the site of siRNA cleavage. Importantly, there is also a need to assess the relationship between a deficit in transcript levels and the yield of the encoded protein, and if relevant, its spatial localization. The aim here is to distinguish between a gene silencing effect that leads to no observable phenotype, and a gene that is resistant to targeting. With these issues in mind, it is proposed that siRNA-based gene targeting in the zebrafish should be quantified and related to what is known about the dynamics of the system being targeted before significant conclusions about a gene's function can be inferred.

The importance of measuring gene targeting at both the transcript and protein level directly influences the choice of a target gene. Specific antibodies are available for few zebrafish genes, and this presents a significant limitation in using most genes associated with well-characterized mutants. As a means of overcoming these limitations, a resource-intensive focus should be adopted. In this respect, modelling the human neuromuscular disorder, Duchenne muscular dystrophy (DMD), offers the resources required to assess the efficacy of siRNA-based gene targeting in the zebrafish, while also offering outcomes in terms of functional genomics related to this disorder. In this respect, although several animal models exist for DMD, there are many outstanding unanswered questions concerning the dystrophin gene. Among these is the need to explore the temporal expression profile of the large variety of transcripts expressed by this gene, and to understand the role of each transcript's protein product.

Zebrafish dystrophin gene as a reporter for siRNA action

Zebrafish are well suited for the study of muscle development since somitic muscle comprises a large portion of the body. Fish somitic muscle forms initially from segmented paraxial mesoderm, and muscle fibres differentiate to span an entire somite in the anterior–posterior axis. By 24 h, the somite adopts a distinctive chevron shape, with the dorsal and ventral halves being separated by a sheet of matrix called the vertical myoseptum. These myosepta, along with the notochord, provide the attachment sites for somitic muscle fibres. It has been suggested that these structures are essentially laminar tendons that transmit force to the notochord and the vertebral column (Bassett and Currie, 2003).

Much of our knowledge of dystrophin function has been derived from studying mice that are null mutants of the dystrophin gene, such as the *mdx* and *mdx3Cv* mutants, which mimic some of the features that characterize human DMD (Bulfield *et al.*, 1984; Sicinski *et al.*, 1989; Cox *et al.*, 1993). Importantly, the pathology manifested by these mice does not mimic entirely that found in human patients, while the developmental role of many of the full-length, as well as shorter, dystrophin isoforms expressed in tissues other than muscle remains unclear (Blake *et al.*, 1999; Dodd *et al.*, 2004a).

In terms of the zebrafish, this species contains gene orthologues of dystrophin, and the sarcoglycans and dystroglycans that comprise the dystrophin glycoprotein complex (DGC; see *Figure 5.6*; Bolanos-Jimenez *et al.*, 2001a,b; Chambers *et al.*, 2001, 2003). The zebrafish *duchenne muscular dystrophy* (*dmd*) gene also expresses a number of isoforms that are similar in size to those found in human tissues (Bolanos-Jimenez *et al.*, 2001b). Zebrafish dystrophin and the DGC are found in the

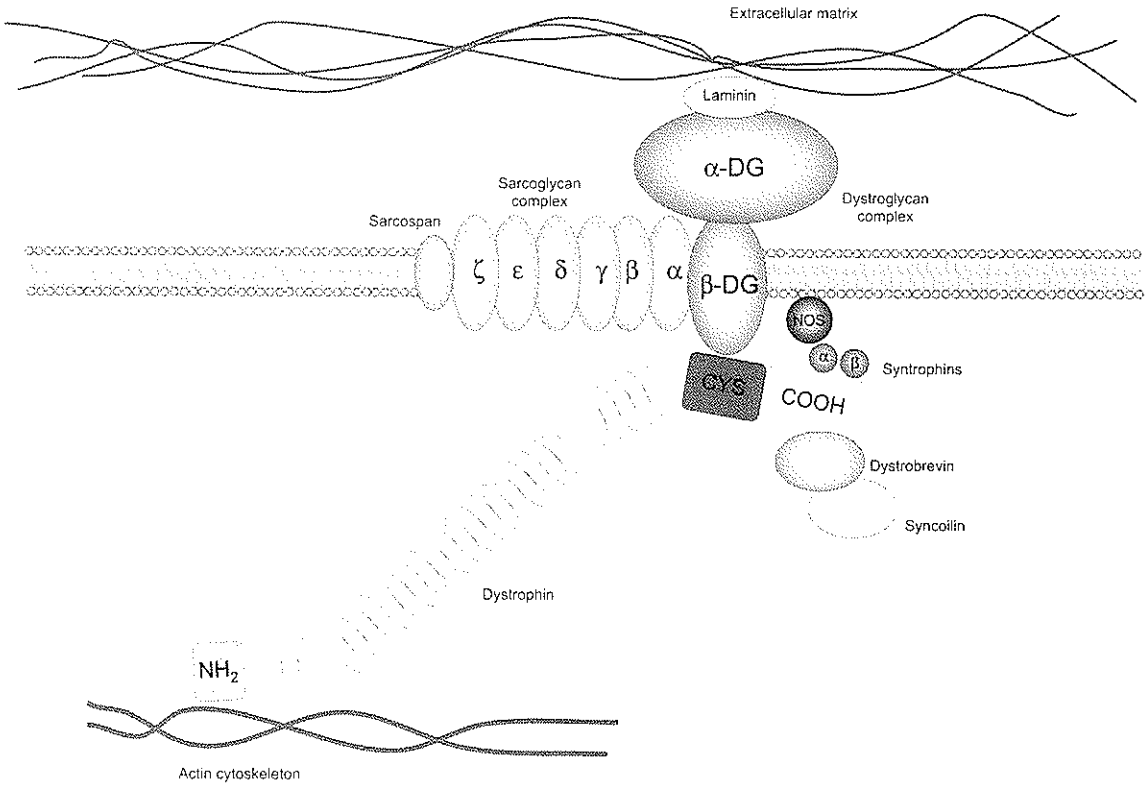

Figure 5.6. The dystrophin glycoprotein complex (DGC). This schematic representation of the DGC outlines the various proteins that are involved in the complex. Dystrophin interacts with cytoplasmic, transmembrane, and extracellular proteins having a critical role in the maintenance of stability by creating a link between the contractile machinery in the cell and the extracellular matrix (diagram does not reflect the exact membrane organization of the DGC).

sarcolemma, with strong localization to transverse myosepta (Chambers *et al.*, 2003). Interestingly, zebrafish treated with a morpholino (MO) targeted to the translational start site of the full-length 427 kDa isoform expressed in muscle (Dp427m; Guyon *et al.*, 2003), and the *sapje* zebrafish mutant that carries a nonsense mutation in exon 4 of the *dmd* gene (Bassett *et al.*, 2003), exhibit a lack of localization of dystrophin in the myoseptae. Recent studies on the disruption of the DGC have only involved MO-based translational knockdown of the dystroglycan gene (Parsons *et al.*, 2002). This research showed that dystroglycan is essential for normal muscle fibre differentiation, and that the lack of dystroglycan results in irregular arrays of sarcomeric protein. In contrast, the recent publication of Bassett and colleagues in examining the *sapje* zebrafish mutant showed that the deficiency of Dp427m in zebrafish embryos caused a muscle attachment failure and progressive loss of muscle integrity (Basset *et al.*, 2003). Taken together, these data support the view that the zebrafish offers an amenable system for examining muscle disease at the gene expression, protein, and cellular levels, and that this species offers advantages in modelling the early events of neuromuscular disorders compared to the more conventional use of the mouse (Bassett and Currie, 2003).

A recent study of siRNA-based targeting of the zebrafish *dmd* gene transcripts showed a reduction in transcript levels, and delayed localization of Dp427m (Dodd *et al.*, 2004a). This level of analysis involved the use of qPCR, protein localization by immunohistochemistry, as well as electron microscopic analysis of zebrafish muscle. However, it should be stressed that other measures of compromised muscle function could also be brought to bear in examining perturbed *dmd* gene expression in the zebrafish.

IMAGING CALCIUM FLUX IN THE ZEBRAFISH

Calcium homeostasis is critical to many aspects of muscle development and function (Ferrari and Spitzer, 1999; Blake *et al.*, 2002). Early reports have shown that resting levels of intracellular calcium ($[Ca^{2+}]_i$) might be abnormally high in dystrophin-deficient muscles, and it has been suggested that this is an important cause of the pathophysiological process leading to dysfunction, and ultimately to cell death (Blake *et al.*, 2002). In this context, the development of fluorescence-based Ca^{2+} reporters (Tsien, 1980) represents a major technological advancement, and they have been applied extensively to isolated muscle fibres, myotubes, and myocytes (Culligan and Ohlendieck, 2002).

It is now generally accepted that the Ca^{2+} levels in human DMD and *mdx* mouse muscles can be abnormally and persistently elevated. Indeed, it has been suggested that contraction-induced sarcolemmal disruption in mature dystrophic myotubes begins a Ca^{2+}-mediated cascade of events that leads to abnormally active Ca^{2+}-specific leak channels, accelerated Ca^{2+} entry through these channels, and increasingly higher levels of Ca^{2+}-dependent proteolysis (Alderton and Steinhardt, 2000). This positive feedback loop in contracting dystrophic myotubes would result in sustained higher Ca^{2+}-dependent proteolytic rates, further abnormalities, and ultimately in cell death. It should be noted, however, that current Ca^{2+} measuring techniques have not been successfully applied to intact animals. Thus, a major gap in our knowledge still exists: what is the nature of the $[Ca^{2+}]_i$ changes in the

developing and functional muscle cells of an intact dystrophic animal compared to a control? This question could be addressed through the use of Ca^{2+} imaging techniques to explore signalling events during zebrafish embryogenesis. These techniques could comprise low-resolution, luminescence-based imaging, followed by high-resolution, fluorescence-based confocal analysis involving the use of ultra-sensitive aequorins (Shimomura *et al.*, 1990) to image free cytosolic Ca^{2+}.

Aequorins are non-toxic; do not buffer Ca^{2+} to any great extent; involve ultra-low background signal (there is no autoluminescence analogous to autofluorescence); show inherent contrast enhancement (i.e. light increases in a relationship that approximates the second power of the Ca^{2+} concentration); have a very wide dynamic range (from 0.01–10 μM *in vivo*); can be transgenically expressed and targeted; and they can be used to image Ca^{2+} in intact zebrafish embryos (Webb and Miller, 2003a,b). The advantage of using aequorin-based imaging is that it allows continuous, non-disturbing imaging of intact zebrafish embryos over the entire period of time required for development of the musculature (i.e. from ~10 to ~40 hpf). By injecting embryos with aequorin at the 64–128-cell stage, cells can be loaded that subsequently form the muscles. Despite the low resolution, aequorin-based imaging can be used to establish where and when Ca^{2+} transients arise in the developing musculature for subsequent examination *via* high-resolution, 2-photon confocal microscopy. It is tempting to speculate that whole organism studies of Ca^{2+} flux could be used also to support a mathematical model of this biological process (Sneyd *et al.*, 2004).

Thus far, we have argued for transcript and protein studies to form a critical part of functional genomics in the zebrafish, coupled with other measures that provide biological data regarding specific disease outcomes. In terms of transcript studies, we shall not discuss the use of microarray technology that has been presented by us elsewhere (Love *et al.*, 2004; Pichler *et al.*, 2004a,b). However, the use of advanced proteomic analysis for the zebrafish has received little attention in the literature. The following discussion seeks to address this deficiency.

Zebrafish proteomics

As a model of disease states, the zebrafish provides an analytical platform to investigate pathologies that are directly reflected by changes in the proteome (the protein complement of the genome). Importantly, the proteome is a more complex and dynamic entity than the genome as it constitutes a mix of alternatively spliced variants and post-translationally modified protein species that vary in form and amount depending on the physiological state at the time of sampling. Since they are the functional output of genes, proteins provide a range of potential molecular markers, and therefore a direct means for profiling a disease state (Mirkes *et al.*, 2003; Shrader *et al.*, 2003). Germane to this discussion has been the emergence of functional proteomics, which reflects a paradigm shift from mere protein identification to defining protein function. In this respect, the vast array of individual proteins isolated from cells, a tissue, or an entire organism, such as the developing zebrafish embryo, can be studied in order to rationalize the physiological changes manifested in a disease condition (Hondermarck, 2004).

In analysing the proteome, proteins have to be extracted and solubilized. A range of

new detergents has been developed to assist in the sequential extraction of proteins since no one set of conditions will be ideal for all proteins; some will always remain recalcitrant to any one treatment. Obtaining subsets, such as membrane proteins, is a useful way to reduce the complexity of the sample, as well as narrowing the isoelectric point or molecular weight range of the protein fraction of interest *via* pre-fractionation and microdialysis. New precipitant and immunoaffinity resins help enrich proteins of interest and remove contaminating substances and high abundance proteins so as to accommodate the 10^4-fold range in concentration observed.

The analysis of proteins has recently advanced to such an extent that performing detailed studies on small numbers of embryos, or even single embryos, is now a realistic and potentially rewarding endeavour. The two most common approaches for the analysis of complex protein mixtures are based on either separation of intact proteins by 2-dimensional (2-D) gel electrophoresis before preparing tryptic peptides, or the separation of peptides themselves derived from digests of the total protein mixture using a combination of ion exchange and reversed-phase HPLC. This step of separation is followed in both instances by MS-based identification. These principles are outlined in *Figure 5.7*. In terms of identification, several major technological improvements are worth noting. These improvements, discussed below, include differential in-gel electrophoresis (DIGE), the increased sensitivity and range of operational modes of modern mass spectrometers, and the range of bioinformatics tools now available for interrogating databases for the identification of proteins of interest.

DIGE (www.gehealthcare.com) is a form of multiplexed dye analysis that highlights proteins exhibiting a change in molecular weight or charge as a result of a post-translational modification, or by a change in expression level, and is observable as a shift in position or intensity, respectively, when analysed by 2-D gel electrophoresis (Tong *et al.*, 2001). Diseased and normal samples are reacted with different reactive fluorescent dyes, then mixed together and subjected to 2-D electrophoresis in the same gel before scanning at optimal wavelengths to visualize each fluorophor independently. Image analysis can then differentiate changes in protein spot intensity to target a reduced set of proteins that specifically change in the disease state. Even a 20% change in intensity over replicated gel sets can be statistically significant. Moreover, fewer replicated gels are required since the gel samples are multiplexed.

Advances in mass spectrometry have been instrumental in increasing the sensitivity so that sub-femtomole levels of proteins can be analysed by a range of mass spectrometry systems. The two common ionization modes are matrix-assisted laser-desorption (MALDI) and nanoelectrospray. These techniques can be coupled to different mass analysers, depending on the type of information that is required. Time-of-flight mass spectrometers generate primarily peptide mass fingerprint profile information, whereas ion-trap or multistage instruments produce tandem (MS–MS) mass spectral data *via* a collision-induced dissociation step, resulting in the fragmentation of individual peptides to produce actual amino acid sequence information. Search algorithms are now capable of quickly and automatically matching acquired spectra with spectra derived from virtual digests of protein sequence information. Nanoelectrospray at flow rates of sub microlitre per minute using 75 μm reversed phase columns ensures that small quantities of hydrolysed proteins

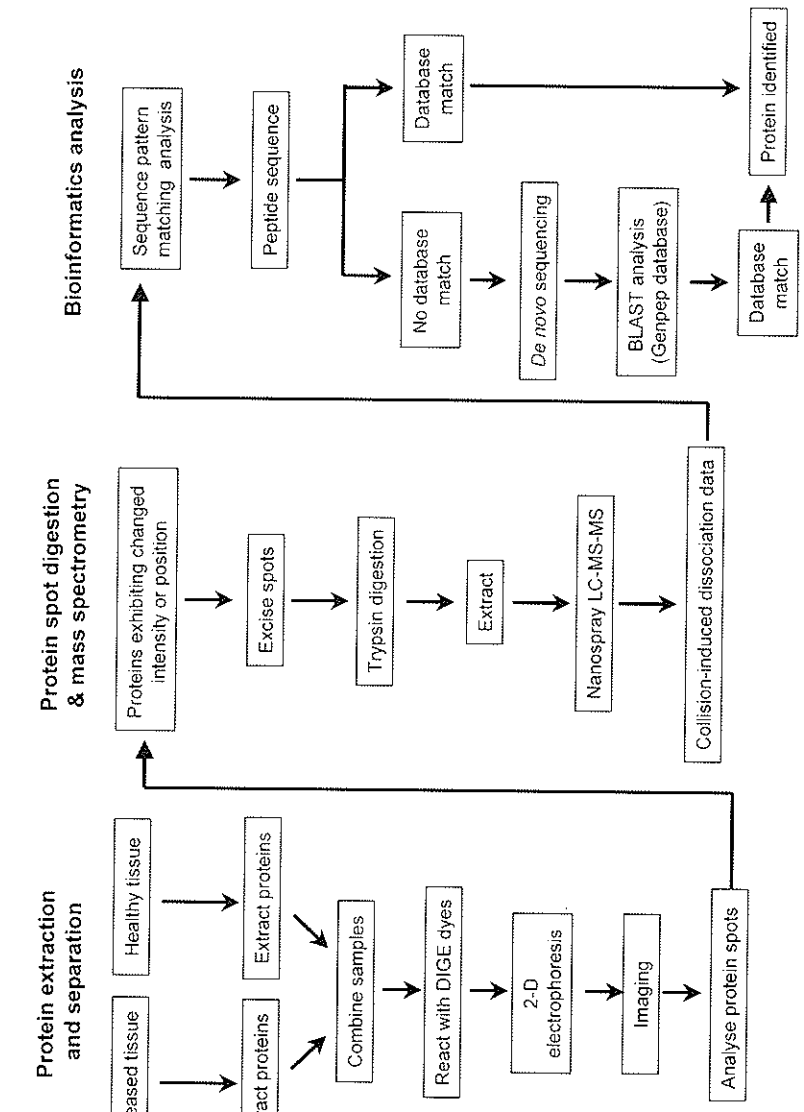

Figure 5.7. Flow diagram of a proteomic-based comparative study of normal and diseased tissue that is applicable to human disease modelling in the zebrafish.

can be separated efficiently in low volume with fast data-dependent scanning to capture the rapid concentration transients and coincident peaks.

The linear ion-trap mass spectrometer with an ion cyclotron resonance cell producing data that is interpretable by Fourier transform analysis has become a new benchmark in the analysis of proteins. An example is the ThermoFinnigan LTQ-FT MS (www.thermo.com) with its high mass resolution, accurate mass assignments, and MS^n capability, such that database search results can be more conclusively assigned. Moreover, using a top-down approach, proteins of mass <30 kDa can be introduced by infusion and fragmented *in situ*, circumventing the need for protease digestion (Wu *et al.*, 2004). The full applicability of this recent development awaits further experimentation.

Bioinformatics plays a pivotal role in the analysis regime dictating which method is likely to be most successful, conditional on how much genome sequence is available. In the case of the zebrafish, genome sequence data are updated daily (www.ensembl.org) with a completion date of the entire genome estimated some time in 2005. The use of peptide mass fingerprinting is of limited benefit where genome coverage is low and only a few representative protein sequences exist. Since the zebrafish does not have close relatives that have already been sequenced, then the direct comparison of MS–MS data with orthologues is not particularly viable, unless the peptide sequences are identical (Kwon *et al.*, 2003). Nevertheless, it is still worth passing the data across what sequences are available. Most information will likely be gleaned from *de novo* sequencing and BLAST analysis against a wider phylogenetic comparison. *De novo* sequencing uses peak differences of the MS–MS data to fit amino acid residues iteratively, and has the advantage of detecting post-translational modifications. Orthologue comparisons are then performed using BLAST analysis.

With the maturation of current proteomic technologies, increasing attention is being directed towards the analysis of complexes that proteins form and interact with. Protein–protein and protein–ligand interactions are fundamental to all cellular processes, not least of these being signalling pathways. Any alteration to these associations is central to the development of many diseases, including neurodegenerative disorders, infectious diseases, and cancers. It is now typical to see the use of bi-functional coupling reagents to covalently link complexed components together for analysis, assorted affinity methods for pull-down of intact complexes, and blue gels for the separation of native protein complexes (Schagger and von Jagow, 1991; Camacho-Carvajal *et al.*, 2004). These, and other developments, ensure that proteomics continues to develop as an important approach for deciphering disease conditions, which are ideally performed in model organisms such as the zebrafish.

Zebrafish as an integrated disease analysis and drug discovery platform

The final part of the disease analysis process is the development of therapies. Therapeutic approaches for genetic disorders at present fall into two general categories: 'gene therapy' *via* gene replacement, or the traditional use of drugs to target a pathogenic process or confer some compensatory effect.

The zebrafish has recently received significant attention as a model system for

drug discovery due to the ability to screen potential drug compounds in a high throughput manner. This approach has been reviewed by the author's laboratory, which can fall into the area of chemical genomics (Pichler *et al.*, 2003; Love *et al.*, 2004). This interest has also initiated the development of several zebrafish-based drug development companies (Pichler *et al.*, 2004b). Although the process is conceptually simple, its successful application requires developments in two key areas. The first of these is concerned with producing models of pathology for subsequent screening of compounds, and the second is the development of targeted high throughput assays to measure the potential impact of compounds on biological processes.

The second area is directly related to the development of therapeutic drugs for disease. In this process, zebrafish might be manipulated to recapitulate a specific disease-related pathology or process, which would allow the analysis of compounds that could interfere with or modulate the disease process. This approach is discussed widely in the literature, but at present is one which we are far from realizing. The central problem directly relates to two critical factors discussed here, which are the need to develop robust and simple gene targeting methods, and the ability to assess disease processes in zebrafish.

Another key area in the development of zebrafish-based drug screening platforms is how best to design assays for testing a drug's action. This problem goes hand-in-hand with model development. An assay must be applicable (eventually) to a high throughput process, but offer the ability to answer potentially complex questions related to the disease process. It should not go unnoticed that a high throughput approach lends itself to screening thousands of small molecules, which has received recent attention with respect to the use of zebrafish embryos (Pichler *et al.*, 2003; Goldsmith, 2004; Love *et al.*, 2004; Peterson, 2004; Peterson *et al.*, 2004). The principal outcome to these studies should be to provide high biological value to the disease being studied, while at the same time achieving a throughput that serves as a drug discovery platform.

Conclusions

The discussion presented here attempts to bring together different disciplines to converge on what we see as the unifying platform of zebrafish. Although this species has received little attention in terms of developing genetic manipulation techniques, we are of the view that siRNA-based gene targeting in the zebrafish offers a facile means of establishing human disease models. Although the work thus far has only involved the direct injection of siRNAs, the need is to develop zebrafish expression constructs in order to achieve temporal and spatial regulation of siRNAs (Dodd *et al.*, 2004a). It is important to realize, however, that modelling human disease requires interrogation at multiple biological levels, and that only a few of these levels are presented here.

The added value of working with the zebrafish is that drug discovery and therapeutics can be studied in an accessible manner. Mention has been made of imaging calcium flux in the developing embryos, but imaging technologies are currently being developed that offer the means of capturing other biological events within the living animal (Huisken *et al.*, 2004; Prescher *et al.*, 2004). All these methods should provide

the critical means by which to develop and assess zebrafish models of human disease, and provide an essential platform for high throughput drug screening approaches.

Acknowledgements

We acknowledge funding support from the Maurice and Phyllis Paykel Trust, the University of Auckland Research Committee, and the Lottery Grants Board of New Zealand.

References

ALDERTON, J.M. AND STEINHARDT, R.A. (2000). Calcium influx through calcium leak channels is responsible for the elevated levels of calcium-dependent proteolysis in dystrophic myotubes. *Journal of Biological Chemistry* **275**, 9452–9460.

AMATRUDA, J.F. AND ZON, L.I. (1999). Dissecting haematopoiesis and disease using the zebrafish. *Developmental Biology* **216**, 1–15.

ANDO, H., FURUTA, T., TSIEN, R.Y. AND OKAMOTO, H. (2001). Photo-mediated gene activation using caged RNA/DNA in zebrafish embryos. *Nature Genetics* **28**, 317–325.

BARBAZUK, W.B., KORF, I., KADAVI, C. *ET AL.* (2000). The syntenic relationship of the zebrafish and human genomes. *Genome Research* **10**, 1351–1358.

BARUT, B.A. AND ZON, L.I. (2000). Realizing the potential of zebrafish as a model for human disease. *Physiological Genomics* **2**, 49–51.

BASSETT, D.I. AND CURRIE, P.D. (2003). The zebrafish as a model for muscular dystrophy and congenital myopathy. *Human Molecular Genetics* **12**, 265–270.

BASSETT, D.I., BRYSON-RICHARDSON, R.J., DAGGETT, D.F., GAUTIER ,P., KEENAN, D.G. AND CURRIE, P.D. (2003). Dystrophin is required for the formation of stable muscle attachments in the zebrafish embryo. *Development* **130**, 5851–5860.

BERMAN, J., HSU, K. AND LOOK, A. (2003). Zebrafish as a model organism for blood diseases. *British Journal of Haematology* **123**, 568–576.

BLAKE, D.J., HAWKES, R., BENSON, R. AND BEESLEY, P.W. (1999). Different dystrophin-like complexes are expressed in neurons and glia. *Journal of Cell Biology* **147**, 645–657.

BLAKE, D.J., WEIR, A., NEWEY, S.E. AND DAVIES, K.E. (2002). Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiological Reviews* **82**, 291–329.

BOLANOS-JIMENEZ, F., BORDAIS, A., BEHRA, M. *ET AL.* (2001a). Molecular cloning and characterization of dystrophin and Dp71, two products of the Duchenne Muscular Dystrophy gene, in zebrafish. *Gene* **274**, 217–226.

BOLANOS-JIMENEZ, F., BORDAIS, A., BEHRA, M., STRAHLE, U., SAHEL, J. AND RENDON, A. (2001b). Dystrophin and Dp71, two products of the DMD gene, show a different pattern of expression during embryonic development in zebrafish. *Mechanisms of Development* **102**, 239–241.

BUCKINGHAM, S.D., ESMAELI, B., WOOD, M. AND SATTELLE, D.B. (2004). RNA interference: from model organisms towards therapy for neural and neuromuscular disorders. *Human Molecular Genetics* **13**, R275–R288.

BULFIELD, G., SILLER, W.G., WIGHT, P.A. AND MOORE, K.J. (1984). X chromosome-linked muscular dystrophy (*mdx*) in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 1189–1192.

CAMACHO-CARVAJAL, M.M., WOLLSCHEID, B., AEBERSOLD, R., STEIMLE, V. AND SCHAMEL, W.W. (2004). Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. *Molecular and Cellular Proteomics* **3**, 176–182.

CHAMBERS, S.P., DODD, A., OVERALL, R. *ET AL.* (2001). Dystrophin in adult zebrafish muscle. *Biochemical and Biophysical Research Communications* **286**, 478–483.

CHAMBERS, S.P., ANDERSON, L.V., MAGUIRE, G.M., DODD, A. AND LOVE, D.R. (2003). Sarcoglycans of the zebrafish: orthology and localization to the sarcolemma and myosepta of muscle. *Biochemical and Biophysical Research Communications* **303**, 488–495.

COX, G.A., PHELPS, S.F., CHAPMAN, V.M. AND CHAMBERLAIN, J.S. (1993). New *mdx* mutation disrupts expression of muscle and non-muscle isoforms of dystrophin. *Nature Genetics* **4**, 87–93.

CULLIGAN, K.G. AND OHLENDIECK, K. (2002). Abnormal calcium handling in muscular dystrophy. *Basic and Applied Myology* **12**, 147–157.

DODD, A., CURTIS, P.M., WILLIAMS, L.C. AND LOVE, D.R. (2000). Zebrafish: bridging the gap between development and disease. *Human Molecular Genetics* **9**, 2443–2449.

DODD, A., CHAMBERS, S.P. AND LOVE, D.R. (2004a). Short interfering RNA-mediated gene targeting in the zebrafish. *FEBS Letters* **561**, 89–93.

DODD, A., CHAMBERS, S.P., NIELSEN, P.E. AND LOVE, D.R. (2004b). Modelling human disease by gene targeting. The zebrafish: cellular and developmental biology (second edition). In: *Methods in cell biology, volume 76*. Eds. H.W. Detrich, M. Westerfield and L.I. Zon, pp 593–612. USA: Elsevier Inc.

DOOLEY, K. AND ZON, L.I. (2000). Zebrafish: a model system for the study of human disease. *Current Opinion in Genetics and Development* **10**, 252–256.

DRAPER, B.W., MORCOS, P.A. AND KIMMEL, C.B. (2001). Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: a quantifiable method for gene knockdown. *Genesis* **30**, 154–156.

EISEN, J.S. (1991). Determination of primary motoneuron identity in developing zebrafish embryos. *Science* **252**, 569–572.

EISEN, J.S. (1996). Zebrafish make a big splash. *Cell* **87**, 969–977.

EKKER, S.C. (2000). Morphants: a new systematic vertebrate functional genomics approach. *Yeast* **17**, 302–306.

ELSEA, S. AND LUCAS, R. (2002). The mousetrap: what we can learn when the mouse model does not mimic the human disease. *Institute for Laboratory Animal Research Journal* **43**, 66–79.

FERRARI, M.B. AND SPITZER, N.C. (1999). Calcium signalling in the developing *Xenopus* myotome. *Developmental Biology* **213**, 269–282.

FORCE, A., LYNCH, M., PICKETT, F.B., AMORES, A., YAN, Y.L. AND POSTLETHWAIT, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545.

GOFFEAU, A., BARRELL, B.G., BUSSEY, H. *ET AL.* (1996). Life with 6000 genes. *Science* **274**, 546–567.

GOLDSMITH, P. (2004). Zebrafish as a pharmacological tool: the how, why and when. *Current Opinion in Pharmacology* **4**, 504–512.

GUYON, J.R., MOSLEY, A.N., ZHOU, Y. *ET AL.* (2003). The dystrophin associated complex in zebrafish. *Human Molecular Genetics* **12**, 601–615.

HAGMANN, M., BRUGGMANN, R., XUE, L. *ET AL.* (1998). Homologous recombination and DNA-end joining reactions in zygotes and early embryos of zebrafish (*Danio rerio*) and *Drosophila melanogaster*. *Biological Chemistry* **379**, 673–681.

HANNON, G.J. AND ROSSI, J.J. (2004). Unlocking the potential of the human genome with RNA interference. *Nature* **431**, 371–378.

HEASMAN, J. (2002). Morpholino oligos: making sense of antisense? *Developmental Biology* **243**, 209–214.

HO, R.K. AND KIMMEL, C.B. (1993). Commitment of cell fate in the early zebrafish embryo. *Science* **261**, 109–111.

HONDERMARCK, H. (Ed.) (2004). *Proteomics: biomedical and pharmaceutical applications*. The Netherlands: Kluwer Academic Publishers.

HUISKEN, J., SWOGER, J., DEL BENE, F., WITTBRODT, J. AND STELZER, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* **305**, 1007–1009.

HUKRIEDE, N.A., JOLY, L., TSANG, M. *ET AL.* (1999). Radiation hybrid mapping of the zebrafish genome. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 9745–9750.

HYATT, T.M. AND EKKER, S.C. (1999). Vectors and techniques for ectopic gene expression in zebrafish. *Methods in Cell Biology* **59**, 117–126.

JU, B., XU, Y., HE, J. *ET AL.* (1999). Faithful expression of green fluorescent protein (GFP) in

transgenic zebrafish embryos under control of zebrafish gene promoters. *Developmental Genetics* **25**, 158–167.

KELLY, P.D., CHU, F., WOODS, I.G. *ET AL.* (2000). Genetic linkage mapping of zebrafish genes and ESTs. *Genome Research* **10**, 558–567.

KIMMEL, C.B. AND WARGA, R.M. (1988). Cell lineage and developmental potential of cells in the zebrafish embryo. *Trends in Genetics* **4**, 68–74.

KNAPIK, E.W., GOODMAN, A., EKKER, M. *ET AL.* (1998). A microsatellite genetic linkage map for zebrafish (*Danio rerio*). *Nature Genetics* **18**, 338–343.

KOOPMAN, P. (2003). Gender benders. *Nature* **426**, 241.

KWON, K.H., KIM, M., KIM, J.Y. *ET AL.* (2003). Efficiency improvement of peptide identification for an organism without complete genome sequence, using expressed sequence tag database and tandem mass spectral data. *Proteomics* **3**, 2305–2309.

LANDER, E.S., LINTON, L.M., BIRREN, B. *ET AL.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

LANGENAU, D.M., TRAVER, D., FERRANDO, A.A. *ET AL.* (2003). Myc-induced T cell leukaemia in transgenic zebrafish. *Science* **299**, 887–890.

LI, Y.X., FARRELL, M.J., LIU, R., MOHANTY, N. AND KIRBY, M.L. (2000). Double-stranded RNA injection produces null phenotypes in zebrafish. *Developmental Biology* **217**, 394–405.

LOVE, D.R., PICHLER, F.B., DODD, A., COPP, B.R. AND GREENWOOD, D.R. (2004). Technology for high-throughput screens: the present and the future using zebrafish. *Current Opinion in Biotechnology* **15**, 1–8.

MA, C., FAN, L., GANASSIN, R., BOLS, N. AND COLLODI, P. (2001). Production of zebrafish germline chimeras from embryo cell cultures. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2461–2466.

MEISTER, G. AND TUSCHL, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349.

MIRKES, P., MCCLURE, M., HEINDEL, J. AND SANDER, M. (2003). Developmental toxicology in the 21st century: multidisciplinary approaches using model organisms and genomics. *Birth Defects Research Part A: Clinical and Molecular Teratology* **67**, 21–34.

MURPHY, E. (1985). The dynamics of quantifiable homeostasis. VI: Processes of fractional order. *American Journal of Medical Genetics* **21**, 385–394.

MYERS, E.W., SUTTON, G.G., DELCHER, A.L. *ET AL.* (2000). A whole-genome assembly of *Drosophila*. *Science* **287**, 2196–2204.

NASEVICIUS, A. AND EKKER, S.C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nature Genetics* **26**, 216–220.

OATES, A.C., BRUCE, A.E. AND HO, R.K. (2000). Too much interference: injection of double-stranded RNA has nonspecific effects in the zebrafish embryo. *Developmental Biology* **224**, 20–28.

PARSONS, M.J., CAMPOS, I., HIRST, E.M. AND STEMPLE, D.L. (2002). Removal of dystroglycan causes severe muscular dystrophy in zebrafish embryos. *Development* **129**, 3505–3512.

PETERSON, R.T. (2004). Discovery of therapeutic targets by phenotype-based zebrafish screens. *Drug Discovery Today: Technologies* **1** (1), 49–54.

PETERSON, R.T., SHAW, S.Y., PETERSON, T.A. *ET AL.* (2004). Chemical suppression of a genetic mutation in a zebrafish model for aortic coarctation. *Nature Biotechnology* **22**, 595–599.

PICHLER, F.B., LAURENSON, S., WILLIAMS, L.C., DODD, A., COPP, B.R. AND LOVE, D.R. (2003). Chemical discovery and global gene expression analysis in zebrafish. *Nature Biotechnology* **21**, 879–883.

PICHLER, F., BLACK, M., WILLIAMS, L.C. AND LOVE, D.R. (2004a). Design, normalization, and analysis of spotted microarray data. The zebrafish: genetics, genomics and informatics (second edition). In: *Methods in cell biology, volume 77*. Eds. H.W. Detrich, M. Westerfield and L.I. Zon, pp 521–543. USA: Elsevier Inc.

PICHLER, F.B., DODD, A. AND LOVE, D.R. (2004b). Global gene expression analysis in the zebrafish: the challenge and the promise. *Drug Discovery Today: Technologies* **1** (2), 79–84.

POSTLETHWAIT, J.H., YAN, Y.L., GATES, M.A. *ET AL.* (1998). Vertebrate genome evolution and the zebrafish gene map. *Nature Genetics* **18**, 345–349.

POSTLETHWAIT, J.H., WOODS, I.G., NGO-HAZELETT, P. *ET AL.* (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Research* **10**, 1890–1902.

POSTLETHWAIT, J., AMORES, A., CRESKO, W., SINGER, A. AND YAN, Y.-L. (2004). Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends in Genetics* **20**, 481–490.

PRESCHER, J.A., DUBE, D.H. AND BERTOZZI, C.R. (2004). Chemical remodelling of cell surfaces in living animals. *Nature* **430**, 873–877.

RISSMAN, E. (2004). Thinking outside the mouse box: the importance of comparatory laboratory animal models in research. *Institute for Laboratory Animal Research Journal* **45**, 1–3.

SCHAGGER, H. AND VON JAGOW, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Analytical Biochemistry* **199**, 223–231.

SERLUCA, F.C., SIDOW, A., MABLY, J.D. AND FISHMAN, M.C. (2001). Partitioning of tissue expression accompanies multiple duplications of the Na+/K+ ATPase alpha subunit gene. *Genome Research* **11**, 1625–1631.

SHIMODA, N., KNAPIK, E.W., ZINITI, J. *ET AL.* (1999). Zebrafish genetic map with 2000 microsatellite markers. *Genomics* **58**, 219–232.

SHIMOMURA, O., INOUYE, S., MUSICKI, B. AND KISHI, Y. (1990). Recombinant aequorins and semisynthetic recombinant aequorins. *Biochemical Journal* **270**, 309–312.

SHIN, J.T. AND FISHMAN, M.C. (2002). From zebrafish to human: modular medical models. *Annual Review of Genomics and Human* Genetics **3**, 311–340.

SHRADER, E., HENRY, T., GREELEY, M. AND BRADLEY, B. (2003). Proteomics in zebrafish exposed to endocrine disrupting chemicals. *Ecotoxicology* **12**, 485–488.

SICINSKI, P., GENG, Y., RYDER-COOK, A.S., BARNARD, E.A., DARLISON, M.G. AND BARNARD, P.J. (1989). The molecular basis of muscular dystrophy in the *mdx* mouse: a point mutation. *Science* **244**, 1578–1580.

SNEYD, J., TSANEVA-ATANOSOVA, K., YULE, D.I., THOMPSON, J.L. AND SHUTTLEWORTH, T.J. (2004). Control of calcium oscillations by membrane fluxes. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1392–1396.

STUART, J., SEGALL, E., KOLLER, D. AND KIM, S. (2003). A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**, 249–255.

TOMASIEWICZ, H.G., FLAHERTY, D.B., SORIA, J.P. AND WOOD, J.G. (2002). Transgenic zebrafish model of neurodegeneration. *Journal of Neuroscience Research* **70**, 734–745.

TONG, R., SHAW, J., MIDDLETON, B. *ET AL.* (2001). Validation and development of fluorescence two-dimensional difference gel electrophoresis proteomics technology. *Proteomics* **1**, 377–396.

TSIEN, R.Y. (1980). New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures. *Biochemistry* **19**, 2396–2404.

VENTER, J.C., ADAMS, M.D., MYERS, E.W. *ET AL.* (2001). The sequence of the human genome. *Science* **291**, 1304–1351.

WARGELIUS, A., ELLINGSEN, S. AND FJOSE, A. (1999). Double-stranded RNA induces specific developmental defects in zebrafish embryos. *Biochemical and Biophysical Research Communications* **263**, 156–161.

WEBB, S.E. AND MILLER, A.L. (2003a). Imaging intercellular calcium waves during late epiboly in intact zebrafish embryos. *Zygote* **11**, 175–182.

WEBB, S.E. AND MILLER, A.L. (2003b). Calcium signalling during embryonic development. *Nature Reviews Molecular Cell Biology* **4**, 539–551.

WESTERFIELD, M. AND EISEN, J.S. (1988). Neuromuscular specificity: pathfinding by identified motor growth cones in a vertebrate embryo. *Trends in Neuroscience* **11**, 18–22.

WIENHOLDS, E.T., SCHULTE-MERKER, S., WALDERICH, B. AND PLASTERK, R.H. (2002). Target-selected inactivation of the zebrafish rag1 gene. *Science* **297**, 99–102.

WOODS, I.G., KELLY, P.D., CHU, F. *ET AL.* (2000). A comparative map of the zebrafish genome. *Genome Research* **10**, 1903–1914.

WU, S.L., JARDINE, I., HANCOCK, W.S. AND KARGER, B.L. (2004). A new and sensitive on-line liquid chromatography/mass spectrometric approach for top-down protein analysis: the comprehensive analysis of human growth hormone in an *E. coli* lysate using a hybrid linear

ion trap/Fourier transform ion cyclotron resonance mass spectrometer. *Rapid Communications in Mass Spectrometry* **18**, 2201–2207.

XU, X., MEILER, S.E., ZHONG, T.P. *ET AL.* (2002). Cardiomyopathy in zebrafish due to mutation in an alternatively spliced exon of titin. *Nature Genetics* **30**, 205–209.

YAN, Y.L., MILLER, C.T., NISSEN, R.M. *ET AL.* (2002). A zebrafish sox9 gene required for cartilage morphogenesis. *Development* **129**, 5065–5079.

ZON, L.I. (1999). Zebrafish: a new model for human disease. *Genome Research* **9**, 99–100.