IN VIVO CELL REPROGRAMMING

current status and potential clinical applications

In Vivo Cell Reprogramming, or IVR, is a new technology that can be used to develop cures for incurable diseases. Using a combination of protein, small molecule or RNA drugs, one type of cell (*substrate cell*) can be reprogrammed to another type of cell (*product cell*). The *product cells* are usually the damaged or depleted cells associated with certain diseases and the *substrate cells* can be any type of cell that is abundant either in proximity to or inside the affected organs. IVR potentially combines the advantages of drug-based and cell-based therapies while avoiding the shortcomings of both.

By Dr Shili Wu, Jonathan Pang, Bobby Baravati, Jun Bao and Dr Yong Zhu

Embryonic stem cells (ESCs) are capable of differentiating into every type of cell in the human body. Previously, differentiation from ESCs to progenitor cells and eventually to somatic cells was believed to be irreversible; cell fate was thought to be committed. However, this paradigm has been overturned by the emergence and acceptance of the cell reprogramming phenomenon.

The concept of cell reprogramming is not new. It emerged in the 1960s: somatic cells were reprogrammed to stem cells using somatic cell nuclear transfer (SCNT)[1]. The groundbreaking discovery of induced Pluripotent stem cells (iPSC) by Dr Yamanaka and colleagues in 2006 once again indicated that cell fate can be reversed[2]. More importantly, Dr Yamanaka and his team showed that the change of cell fate can be accomplished using defined factors – in this case, by forced expression of four transcription factors: Oct4, Sox2, Klf4 and c-Myc. The field of cell reprogramming has exploded since then. In 2008, Dr Melton's group demonstrated that pancreatic exocrine cells can be transdifferentiated to β cells *in vivo* via viral transduction of several genes encoding pancreatic and β cell transcription factors, namely, Ngn3 (Neurog3), Pdx1, and MafA[3]. More recently, researchers have directly reprogrammed fibroblasts into both functional neurons[4] and into functional cardiomyocytes[5] *in vitro* using the same technique.

All these exciting discoveries have raised an intriguing question: can we directly reprogramme one type of cells to another type of cells *in vivo* (or *in situ*) for therapeutic purposes? In this article, we will discuss the concept of *in vivo* cell reprogramming, its potential clinical applications and possible challenges.

Target: difficult diseases

Although great advances have been made in biomedical research, scientists and clinicians are still looking for effective treatments for some of the most devastating diseases, such as diabetes, cardiovascular diseases, spinal cord injury, Parkinson's disease and Alzheimer's disease. Some

Table 1: Some incurable diseases and their associated deficient cell types

DISEASE	DEFICIENT CELL TYPE
Diabetes mellitus	β cells
Myocardial Infarction	Cardiomyocytes
Stroke	Neurons
TRAUMATIC BRAIN INJURY	**NEURONS**
Parkinson's disease	Dopaminergic neurons
Alzheimer's disease	Neurons

traditionally incurable diseases are listed in **Table 1**. A common feature of these diseases is that they are characterised by a deficiency or depletion of particular cells.

Potential treatments for these diseases include drug-based therapies and cell-based therapies. However, both have their limitations. Drug-based therapies usually treat the symptoms of disease only and make the patient chronically dependent upon them. Cell-based therapies are hampered by the scarcity of the source, immune rejection, substantial manufacturing and distribution costs, and high risks in development and commercialisation (**Table 2**).

Is there an approach that can combine the advantages of both drug-based and cell-based therapies while avoiding the shortcomings of both? Ideally, deficient cells are replenished *in situ* upon the treatment of therapeutics, eliminating the need for cell transplantation. There may be several approaches capable of accomplishing this. One way is to stimulate adult stem cells to produce the deficient cell type. Another way, which we call *in vivo* cell reprogramming or IVR, is to directly convert another type of cells nearby to the deficient cell type. The IVR approach potentially has a broader application, since the targeted somatic cells are more abundant than adult stem cells.

Model and mechanisms

The aim of *in vivo* cell reprogramming, also known as *in vivo* direct reprogramming or *in vivo* lineage reprogramming, is to achieve the 4Rs – cell regeneration, replacement, repair and rejuvenation: the holy grail of regenerative medicine. To further discuss IVR, it is first necessary to establish a working terminology for the field. Borrowing terms from enzymology, we named the starting cells we wish to reprogramme *substrate cells* and the reprogrammed cells *product cells* (**Figure 1**).

Substrate cells are a cell type that is abundant and in proximity to, or within the affected organs. They should be phylogenetically very close to *product cells*' lineage. These three criteria – abundance, proximity and relevance – not only make the reprogramming easier, but also ensure that the newly created *product cells* function similarly to the original cells in the same microenvironment and in response to the same physiological signals. For example, hepatocytes or pancreatic acinar cells (*substrate cells*) can be converted to pancreatic β

Table 2: The limitations of drug-based therapies and cell-based therapies

	CONVENTIONAL DRUG-BASED THERAPIES	CELL-BASED THERAPIES
MODALITIES	Small molecules, proteins, RNAs...	Cells (including stem cells or their derivatives), tissues or organs
LIMITATIONS	Some diseases are untreatable	Scarcity of donors
	Treat or manage symptoms only	Immune rejection
	Chronic dependence on drugs	Unsustainable business models
	Expensive	– Manufacture and distribution costs
		– Regulatory risks
		– Development risks
		– Avoided by investors

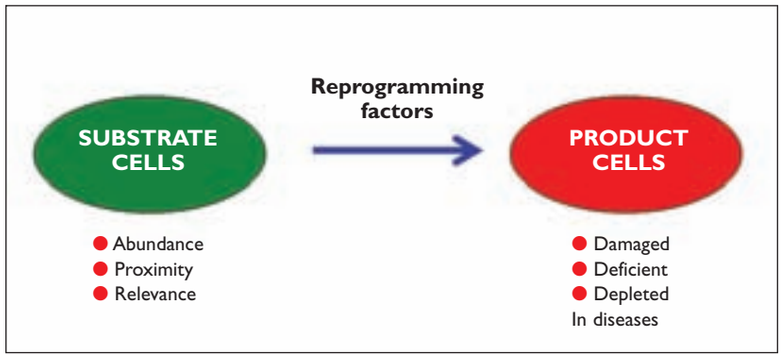

Figure 1: The basic model of IVR technology. Using a combination of protein drugs (and/or small molecule drugs), we can convert one type of cells to another type of cells. Borrowing terms from enzymology, we named the starting cells we wish to reprogramme *substrate cells* and the reprogrammed cells ready for use *product cells*. *Product cells* are usually the damaged or depleted cells associated with certain diseases, and *substrate cells* are a cell type that is abundant, in proximity to or inside the affected organs, and close to *product cells* phylogenetically

Table 3: Selection of *substrate* and *product cells* in targeting some diseases

DISEASE	SUBSTRATE CELL	PRODUCT CELL
Diabetes	Liver cells Pancreatic exocrine cells	β cell
Autoimmune diseases	T cells	Regulatory T cells (T_{reg}s)
Obesity	White fat cells	Brown fat cells
Stroke	Astrocytes	Neurons
Myocardial infarction	Cardiac fibroblasts	Cardiomyocytes

cells (*product cells*) to treat diabetes. More examples are given in **Table 3**.

Reprogramming can be through any of the five mechanisms: 1) Differentiation, 2) Transdifferentiation, 3) Retrodifferentiation, 4) Transdetermination and 5) Dedifferentiation. These mechanisms are illustrated in **Figure 2**.

Figure 2A is a view of cell differentiation using a simplified cell lineage tree. **Figure 2B** is adapted from the differentiation model proposed by Conrad Waddington 50 years ago[6]. The ball sitting atop a slope and poised to roll down into different branching ravines represents the totipotent zygote ready to differentiate into pluripotent and multipotent stem cells and eventually somatic cells and to take different fates.

Reprogramming factors are mainly transcription factors and chromosome remodelling proteins, or agents that can induce, affect or express these proteins. Their modalities include proteins, small molecules, RNAs (including siRNAs, miRNAs and mRNAs) or DNAs.

In applying the IVR technology, the key is selecting the right combination of reprogramming factors. How do we identify these factors after determining the *product cells* and the *substrate cells*? The answer may come from systems biology and developmental biology.

The development process from a single cell to a complicated individual is hard-wired in the species' genome – the pathways to different cell fates are encoded by the Gene Regulatory Network (GRN)[7]. The GRN is an extensive network that

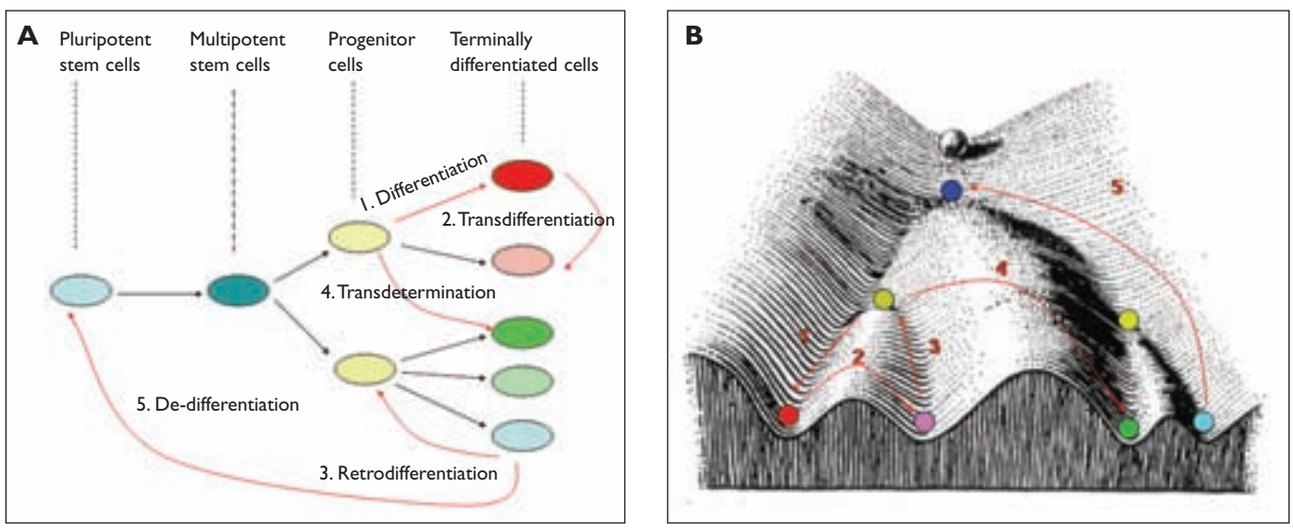

Figure 2: Mechanisms of cell reprogramming. **A** Linage tree view. The conversion from the *substrate cells* to the *product cells* can be through any of the five mechanisms: 1) Differentiation, 2) Transdifferentiation, 3) Retrodifferentiation, 4) Transdetermination and 5) Dedifferentiation, which are illustrated here **B** Stochastic model view. The same mechanisms are illustrated using the Waddington model

Table 4: Comparison of IVR modalities

MODALITY	STRENGTHS	WEAKNESSES
Small molecule	Easy to manufacture and administer	Development pathway less straightforward Less specific – off target effects and toxicities
RNA	Easy to manufacture Development straightforward	Problematic delivery: stability and specificity May need carriers to target specific cells
Protein	Development straightforward Very specific (cell type, distribution, and mechanisms)	Relatively high manufacture and distribution cost Inconvenient administration

consists of numerous genes and their products (nodes), and the regulatory relations among them (edges). Although all the cells in the human body share one GRN, the state of the GRN differs significantly among different cell types. For example, some parts of the GRN are active or lightened up in β cells, but are inactive or go dark in hepatocytes and *vice versa*. In fact, what distinguishes β cells from hepatocytes is their internal state of the GRN. The differences in GRN states are due to differences in cell environment, the legacy components from parent cells, and the history of cell lineages. The goal of the reprogramming process, then, is to perturb one GRN state into another. One way this can be done is to turn on parts of the GRN while simultaneously shutting down the other parts. Transcription factors are the convenient choices for perturbing GRN states because they occupy the top layers of the hierarchical GRN. They are the switches for different GRN states.

Comparing the GRN states of two different types of cells may uncover hundreds of transcription factors that differ significantly in expression. However, there is a shortcut through this maze. Among all of these nodes exists an active kernel that is highly conserved in the same type of cells among different species[7]. Since certain nodes of such conservative kernels in the *product cells* are inactive in the *substrate cells*, the protein products of such nodes are instrumental in the conversion process.

Using cues from systems biology and confirmation from experimental biology, it is not difficult to identify the reprogramming factors given the *substrate* and the *product cells*. Here we can only provide the general principles; a more detailed treatment is the subject of another article.

One key feature of the active kernel module for a specific cell is that it enters a 'lock on' state once the cells' fate is decided. The 'lock on' state is maintained through several positive feedback loops initially and then through epigenetic events in the long term. Therefore, after introducing several key transcription factor proteins and initiating a new GRN state, the external reprogramming factors are no longer necessary. Even when these external reprogramming factors are withdrawn from the environment, the cells will retain the new fate because the new GRN state is maintained internally and becomes very stable, ie 'locked on'.

Strategy

Any effective cure towards the difficult diseases should fulfill two requirements. First, we need to identify the mechanism behind cell deficiency and halt this cell-destruction mechanism. Second, we need to find a way to replenish the cells.

In certain diseases, cell reprogramming can be used to meet the latter requirement and integrated into a holistic therapeutic strategy. For example, in stroke patients, blood clots are cleared and the ruptured/blocked blood vessels are repaired at the injury sites through surgery. During the procedure, some agents can be administered to initiate cell reprogramming and to rewire the neural circuitry. In some other diseases, cell reprogramming can be used for both requirements. For example, in type I diabetes, the immune system destroyed the patient's own β cells, through a process that is still not fully understood. In this case, T cells can be first reprogrammed to T_{reg} cells to stop the self-destruction,

References

1 Gurdon, JB. Adult frogs derived from the nuclei of single somatic cells. Developmental Biology 4, 256-273 (1962).

2 Takahashi, K and Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676 (2006).

3 Zhou, Q, Brown, J, Kanarek, A. Rajagopal, J and Melton, DA. In vivo reprogramming of adult pancreatic exocrine cells to β-cells. Nature 455, 627-632 (2008).

4 Vierbuchen, T, Ostermerier, A, Pang, ZP, Kokubu, Y, Südhof, T and Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. Doi:10.1038/nature08797 (2010),

5 Ieda, M, Fu, JD, Delgado-Olguin, P, Vedantham, V, Hayashi, Y, Bruneau, BG and Srivastava, D. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell 142, 375-386 (2010).

6 Waddington, CH. The strategy of the genes. A discussion of some aspects of theoretical biology. Alen & Unwin, 1957.

7 Davidson, EH. The regulatory genome: gene regulatory networks in development and evolution. Academic Press, USA, 2006.

8 Warren, L, Manos, PD, Ahfeldt, T, Loh, Y-H, Li, H, Lau, F, Ebina, W, Mandal, PK, Smith, ZD, Meissner, A, Daley, GQ, Brack, AS, Collins, JJ, Cowan, C, Schlaeger, TM and Rossi, DJ. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 7, 618-630 (2010).

9 Wadia, JS and Dowdy, SF. Protein transduction technology. Current Opinion in Biotechnology 13, 52-56 (2002).

Continued on page 87

followed by reprogramming of acinar cells to insulin producing β cells.

Modality

The reprogramming factors can be proteins, small molecules, RNAs (including siRNAs, miRNAs and mRNAs) or DNAs (using virus or plasmid as vectors). From a practical standpoint, the DNA approach or gene therapy may have the most difficulties in reaching the clinics, as shown by the tortuous history in the gene therapy field. However, the DNA approach is very useful in screening and selecting the right combination of reprogramming factors in the research and discovery stage, as demonstrated in several studies[2-5]. The strengths and limitations of the other modalities – small molecules, RNAs and proteins – are outlined in **Table 4**.

Certain small molecule drugs can exert the reprogramming effects by changing the epigenetic landscape, affecting signal transduction, or activating/suppressing transcription factors. But finding the right small molecule drugs is not straightforward – it requires extensive screening and optimisation. An advantage of small molecule drugs is that they are easy to synthesise, transport and administer, thus the costs of manufacture, distribution and administration are low. However, like all other small molecule drugs, there are specificity issues: they can act on multiple types of cells, have off-target effects, or distribute all over the body.

The development of RNA reprogramming agents is straightforward. Once the reprogramming factors are identified, the respective RNAs can be generated and evaluated[8]. But the biggest problem for the RNA approach is delivery: barriers to effective RNA delivery include low stability, low specificity and high toxicity upon systemic administration. Several years ago, siRNA was the hottest therapeutic option in the pipeline of the biopharmaceutical industry. But major pharmas' excitement over RNA interference has faded due to the delivery problem. A recent example came last November when Roche announced it would end its in-house RNAi research.

Protein-based treatment may be more practical than RNA-based therapeutics. But to make it work, two hurdles in drug delivery have to be overcome. First, the reprogramming factors are usually transcription factors and chromosome remodelling proteins – they have to enter the nuclei to be functional. But in their native forms these proteins cannot be delivered into the cells. To overcome this obstacle, these proteins can be fused with protein transduction domains (PTDs). As a result, the engineered proteins are transducible, ie, capable of penetrating cells and trafficking into nuclei when placed in cell culture medium or around the affected organs[9]. Secondly, most PTDs are not selective among different cell types. Ideally, the reprogramming proteins will only enter the *substrate cells* upon administration. To achieve this end, proteins can be engineered to target specific cell types. One strategy is to make the PTD with high affinity for certain cell types. A cell-specific PTD can be selected by screening a peptide library based on the phage display technology[10]. Another way is to construct a fusion protein in which the transducible component is an antibody against a cell-specific surface receptor[11]. The fusion protein can enter the cells that specifically express the receptor through endocytosis.

Protein-based IVR can be very specific in terms of mechanism, location and cell type. Its drawbacks are the same as other protein drugs: the manufacturing and distribution cost is relatively high, compared to small molecule and RNA agents; and unlike small molecule drugs, they can not be taken orally.

The modality of IVR therapeutics is not exclusive – for example, proteins and small molecules may be used in conjunction to achieve a higher reprogramming efficiency than that achieved by proteins alone[12].

Protein-based IVR: example

In April 2009, a collaboration between our team and The Scripps Research Institute resulted in a major breakthrough in stem cell research[12]. For the first time mouse fibroblast cells were converted to iPSCs without the use of any exogenous genetic material. We engineered and produced four transcription factor proteins – oct4, sox2, klf4 and c-myc – that are capable of penetrating into and reprogramming cells. By using recombinant proteins we eliminated any risk of modifying the target cell genome, making it possible to use iPSC and its derivative cells in a clinical setting. More importantly, the study established the feasibility of using protein-based IVR in clinical settings.

One application of the IVR technology is in the treatment of diabetes. Our strategy is to reprogramme liver and/or pancreatic exocrine cells to insulin-producing β-cells using the combination of transducible proteins Pdx1-PTD, Ngn3-PTD and MafA-PTD. Some preliminary data are presented below.

Six CD-1 mice were divided into two groups: the treatment group and the control group. Transducible protein Pdx1-PTD (1mg/kg), Ngn3-

Continued from page 85

10 Brown, KC. New approaches for cell-specific targeting: identification of cell-selective peptides from combinatorial libraries. Current Opinion in Chemical Biology 4, 16-21 (2000).
11 Song, E, Zhu, P, Lee, S-K, Chowdhury, D, Kussman, S, Dykxhoorn, DM, Feng, Y, Palliser, D, Weiner, DB, Shankar, P, Marasco, WA, Lieberman, J. Antibody mediated *in vivo* delivery of small interfering RNAs via cell-surface receptors. Nature Biotechnology 23, 709-717 (2005).
12 Zhou, H, Wu, S, Joo, JY, Zhu, S, Han, DW, Lin, T, Trauger, S, Bien, G, Yao, S, Zhu, Y, Siuzdak, G, Schöler, HR, Duan, L and Ding, S. Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins. Cell Stem Cell 4, 581 (2009).
13 Ber, I, Shternhall, K, Perl, S, Ohanuna, Z, Goldberg, I, Barshack, I, Benvenisti-Zarum, L, Meivar-Levy, I and Ferber, S. Functional, Persistent and extended liver to pancreas transdifferentiation. The Journal of Biological Chemistry 278, 31950-31957 (2003).

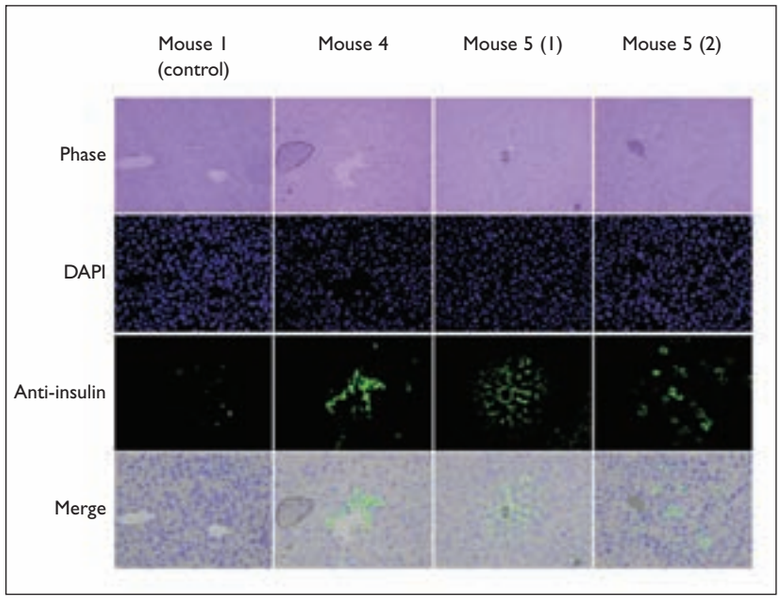

Figure 3: Immunofluorescent analysis of liver tissues of mice treated with control protein or with Pdx1-PTD, Ngn3-PTD and MafA-PTD proteins. Mice 1, 2 and 3 were from the control group, and mice 4, 5 and 6 were treated with the three proteins. The first row of panels are bright phase images, the second are fluorescent images with DAPI staining, and the third are fluorescent images with insulin antibody staining. Each image on the fourth row of panels is the merge image of the three images above. Each column of panels are for the same field

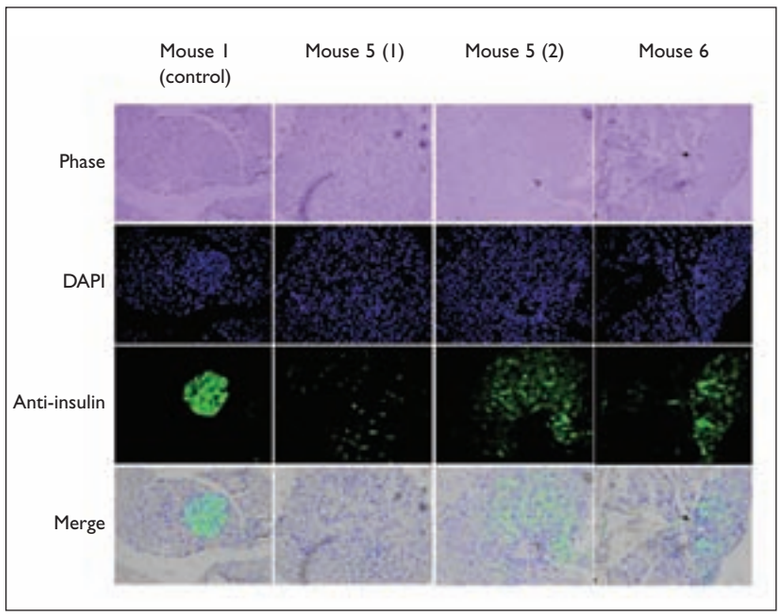

Figure 4: Immunofluorescent analysis of pancreatic tissues of mice treated with control protein or with Pdx1-PTD, Ngn3-PTD and MafA-PTD proteins. Mice 1, 2 and 3 were from the control group, and mice 4, 5, and 6 were treated with the three proteins. The first row of panels are bright phase images, the second are fluorescent images with DAPI staining, and the third are fluorescent images with insulin antibody staining. Each image on the fourth row of panels is the merge image of the three images above. Each column of panels are for the same field

PTD (1mg/kg), and MafA-PTD (1mg/kg) were injected into each mouse by intraperitoneal (IP) in treatment group (Mouse-4, Mouse-5 and Mouse-6) and BSA (1mg/kg) was injected into each mouse in the control group (Mouse-1, Mouse-2 and Mouse-3). Injections were repeated every day for seven days. Mice of both treatment and control group were sacrificed on day 10, three days after completion of injections. The mouse liver and pancreas were washed with 1X PBS and fixed by 4% paraformaldehyde for overnight. Then the liver and pancreatic tissues were processed by standard Paraffin Embedding protocol. The Tissue sections, 5-micro in thickness, were prepared routinely with histology microtomes and mounted on standard histology glass slides. The wax in tissues was dissolved by xylene during processing of tissue sections. Tissue sectioning and histologic and immunohistochemical staining were performed using routine methods. For indirect fluorescent-antibody (IFA) assay, the slides were blocked with 0.05% Tween 20 (TBST) and 3% BSA for one hour at RT and were incubated with mouse anti-insulin antibody (Invitrogen) at 4°C overnight. The slides were washed three times with PBS for 15 minutes at RT and incubated with fluorescein isothiocyanate (FITC) conjugated swine anti-mouse antibody (KPL) for two hours at RT. The same concentration of Mouse IgG was used as isotype control. Anti DAPI antibody was added to slides as a nuclear marker (**Figures 3 and 4**).

The results showed that the treatment group had more insulin-producing cells in livers compared with the control group (**Figure 3**). There are some insulin positive cells in the liver of control animals (mouse 1, **Figure 3**). This observation is not surprising since the same kind of leakiness in mice has been reported in the literature before[13]. In the livers of protein-treated mice, clusters of insulin-positive cells emerged, many of which were concentrated in the periportal areas.

In the pancreas of mice treated with the three transducible proteins, clusters of insulin-producing cells were observed beside the normal insulin-intensive β islets. Different from the native β islets, these clusters of insulin-producing cells are more mosaic and diffusive (**Figure 4**).

Conclusions

IVR potentially combines the advantages of drug-based and cell-based therapies while avoiding the shortcomings of both. Since the modality of IVR therapeutics are proteins, small molecules or RNAs, their risk profiles and cost structures of development and commercialisation are similar to

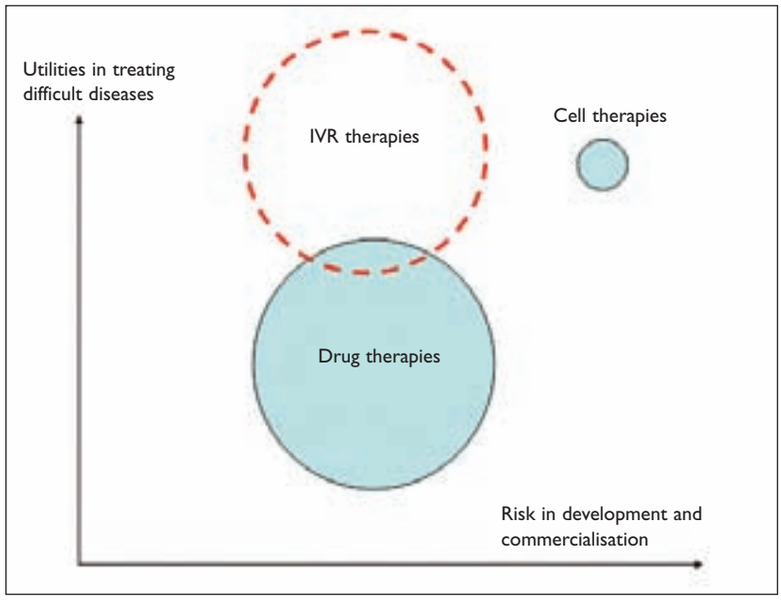

Figure 5: New strategy: *in vivo* cell reprogramming therapies. The size of the area enclosed by each circle indicates how big the market for that type of treatment is. The dotted circle is the projected market for *in vivo* cell reprogramming therapies. The risk profiles and cost structures of development and commercialisation of IVR therapeutics are similar to those of drug-based therapies. However, by using a cell-based mechanism, IVR therapeutics can achieve therapeutic benefits that cannot be achieved by conventional drugs

those of drug-based therapies (**Figure 5**). However, by using a cell-based mechanism, IVR therapeutics can achieve therapeutic benefits that cannot be achieved by conventional drugs. In addition, the newly created *product cells* are sustainable. Therefore a short course of treatment can provide long-term clinical benefits. Moreover, contrary to most situations in cell or tissue transplantation, new *product cells* will not cause immune rejection.

IVR can be applied to many disease areas, including diabetes, autoimmune diseases and stroke. As shown in **Figure 5**, we expect the future market for IVR therapeutics to be as big as that for drug-based therapeutics. To overcome potential challenges from the concept to the clinics, such as issues of efficiency, specificity and immunogenicity, the advancement of this field requires a collective effort from researchers, clinicians, institutes and companies. Given enough investment and commitment, the research in this promising area will ripen in the coming decade, forever revolutionising medicine. DDW

Dr Shili Wu *is Director of Technology at VivoScript, a US-based biotech company established to commercialise a disruptive innovation – IVR technology – in regenerative medicine. Messrs* ***Jonathan Pang*** *and* ***Bobby Baravati*** *are research associates at VivoScript. Mr* ***Jun Bao*** *is CEO of the company.* ***Dr Yong Zhu*** *is Vice President, Research and Business Development, at VivoScript.*