UAlg
UNIVERSIDADE DO ALGARVE
DEPARTAMENTO DE CIÊNCIAS BIOMÉDICAS E MEDICINA

UNIVERSIDADE DO ALGARVE

DEPARTAMENTO DE CIÊNCIAS
BIOMÉDICAS E MEDICINA

Development of a Piggybac based direct reprogramming system for derivation of integration free induced pluripotent stem cells

Dino Emanuel Santos Matias

Tese

Mestrado em Ciências Biomédicas

Trabalho efectuado sob a orientação de:

Professor Doutor Gustavo Tiscornia

2013

UNIVERSIDADE DO ALGARVE

DEPARTAMENTO DE CIÊNCIAS BIOMÉDICAS E MEDICINA

Development of a Piggybac based direct reprogramming system for derivation of integration free induced pluripotent stem cells

Dissertação apresentada à Universidade do Algarve para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Ciências Biomédicas, realizada sob a orientação do Professor Doutor Gustavo Tiscornia (Departamento de Ciências Biomédicas e Medicina e Centro de Biomedicina Molecular e Estrutural - CBME).

Dino Emanuel Santos Matias

2013

Development of a Piggybac based direct reprogramming system for derivation of integration free induced pluripotent stem cells

Declaração de autoria de trabalho

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Abstract

Induced pluripotent stem cells (iPSc) have great potential for applications in regenerative medicine, disease modeling and basic research. Several methods have been developed for their derivation. The original method of Takahashi and Yamanaka involved the use of retroviral vectors which result in insertional mutagenesis, presence in the genome of potential oncogenes and effects of residual transgene expression on differentiation bias of each particular iPSc line. Other methods have been developed, using different viral vectors (adenovirus and Sendai virus), transient plasmid transfection, mRNA transduction, protein transduction and use of small molecules. However, these methods suffer from low efficiencies; can be extremely labor intensive, or both. An additional method makes use of the piggybac transposon, which has the advantage of inserting its payload into the host genome and being perfectly excised upon re-expression of the transposon transposase. Briefly, a policistronic cassette expressing Oct4, Sox2, Klf4 and C-Myc flanked by piggybac terminal repeats is delivered to the cells along with a plasmid transiently expressing piggybac transposase. Once reprogramming occurs, the cells are re-transfected with transposase and subclones free of tranposon integrations screened for. The procedure is therefore very labor intensive, requiring multiple manipulations and successive rounds of cloning and screening.

The original method for reprogramming with the the PiggyBac transposon was created by Woltjen et al in 2009 (schematized here) and describes a process with which it is possible to obtain insert-free iPSc. Insert-free iPSc enables the establishment of better cellular models of iPS and adds a new level of security to the use of these cells in regenerative medicine. Due to the fact that it was based on several low efficiency steps, the overall efficiency of the method is very low (<1%). Moreover, the stochastic transfection, integration, excision and the inexistence of an active way of selection leaves this method in need of extensive characterization and screening of the final clones.

In this work we aime to develop a non-integrative iPSc derivation system in which integration and excision of the transgenes can be controlled by simple media manipulations, avoiding labor intensive and potentially mutagenic procedures. To reach our goal we developed a two vector system which is simultaneously delivered to the

original population of fibroblasts. The first vector, Remo I, carries the reprogramming cassette and GFP under the regulation of a constitutive promoter (CAG). The second vector, Eneas, carries the piggybac transposase associated with an estrogen receptor fragment (ERT2), regulated in a TET-OFF fashion, and its equivalent reverse trans-activator associated with a positive-negative selection cassette under a constitutive promoter. We tested its functionality in HEK 293T cells.

The protocol is divided in two the following steps:

1) Obtaining acceptable transfection efficiency into human fibroblasts.
2) Testing the functionality of the construct
3) Determining the ideal concentration of DOX for repressing mPB-ERT2 expression
4) Determining the ideal concentration of TM for transposition into the genome
5) Determining the ideal Windows of no DOX/TM pulse for transposition into the genome
6) 3, 4 and 5) for transposition out of the genome
7) Determination of the ideal concentration of GCV for negative selection

We successfully demonstrated that ENEAS behaved as expected in terms of DOX regulation of the expression of mPB-ERT2. We also demonstrated that by delivering the plasmid into 293T HEK cells and manipulating the levels of DOX and TM in the medium, we could obtain puromycin resistant lines. The number of puromycin resistant colonies obtained was significantly higher when DOX as absent, suggesting that the colonies resulted from transposition events. Presence of TM added an extra layer of regulation, albeit weaker. Our PCR analysis, while not a clean as would be desired, suggested that transposition was indeed occurring, although a background level of random integration could not be ruled out. Finally, our attempt to determine whether we could use GVC to select clones that had successfully mobilized PB out of the genome was unsuccessful. Unexpectedly, 293T HEK cells that had been transfected with ENEAS and selected for puromycin resistance were insensitive to GCV.

Keywords: reprogramming, iPS, PiggyBac, transposon, transpositiono, cloning.

Resumo

Reprogramação celular é a tecnologia mais recente no campo da biologia celular e desenvolvimento. A possibilidade de reverter qualquer tipo celular a um estado de pluripotencia, a partir do qual possa ser diferenciado em qualquer outro do corpo humano, independentemente da célula de partida abriu toda uma panóplia de possibilidades e conceitos em Biomedicina. As propriedades intrínsecas das células iPS de divisão simétrica e pluripotencia, semelhantes às das células estaminais, tornam-se especialmente importantes para os ramos da medicina regenerativa e investigação biomédica, com realce para o desenvolvimento de modelos de doença *in vitro*, uma vez que contrariamente às células estaminais, só encontradas durante o desenvolvimento embrionário, as células iPS podem ser desenvolvidas a partir de um indivíduo adulto. Após obtenção de iPSs é possível obter virtualmente um número ilimitado de qualquer género celular do dador, inclusive células de difícil obtenção, por falta de casos, difícil isolamento, ou falta de casos clínicos.

O método original foi desenvolvido por Takahashi e Yamanaka e baseia-se na administração e expressão forçada de factores de transcrição, nomeadamente Oct4, Sox2, Klf4 e C-Myc (OSKC). Estes factores foram referenciados ao longo dos anos por estudos de transcriptómica comparativa entre vários clones de células estaminais *versus* células conhecidas como percursoras, ou progenitoras, conhecidas por manterem ainda algum grau de pluripotencia, e células somáticas adultas. Em 1962, John Gurdon mudou todo o campo conhecido como biologia do desenvolvimento ao gerar clones de rã através de transferência nuclear somática (SCNT), provando que toda a informação genética necessária para formar o indivíduo adulto permanece no núcleo das células somáticas, despertando assim uma revolução na área e ao desenvolvimento do conceito de reprogramação celular. De referir também a contribuição marcante de Davis e colegas que em 1987, usando um elegante método de extracção de ADN complementar detectaram um conjunto de três genes que se encontravam predominantemente expressos em mioblastos. A transcrição forçada de apenas um desses genes, Myod1, revelou-se capaz de converter fibroblastos em mioblastos capazes de expressar miosina, relatando assim, pela primeira vez um processo actualmente denominado por reprogramação directa ou transdiferenciação.

Em relação ao processo de reprogramação propriamente dito, ainda há muito para descobrir. No entanto alguns mecanismos já foram deslindados nestes sete anos. Apesar de já serem conhecidas novas combinações de factores que são capazes de reverter o estado de diferenciação celular estabelecido em diferentes tipos celulares, os quatro factores de Yamanaka, OSKM continuam a ser o padrão na área, principalmente pela sua robustez e capacidade de reprogramar a maioria dos tipos celulares. Os factores OSKM têm um efeito sinergético, funcionando em conjunto para ultrapassar os resilientes sistemas celulares intrínsecos de protecção de identidade. Assim, neste sistema, a expressão de Oct4 e Sox2 promove maioritariamente um efeito desestabilizador da ordem transcripcional estabelecida, recrutando NANOG, outro agente de pluripotencia e formando um núcleo autoregulatório de pluripotencia. Este núcleo, uma vez estabelecido activa vias de sinalização como a via da MAPK1 e WNT3, e sinalizando o grupo Polycomb. Por outro lado, os factores Klf4 e c-Myc, como reguladores da divisão celular, têm uma actividade mitogénica, obrigando a alterações constantes no estado epigenético da célula, aumentando fortemente a cinética e eficiência do processo.

Ainda que, ao momento, desenvolvimento de iPSs utilizando vectores virais integrativos de função semelhante aos utlizados na publicação original não seja já considerado um desafio a nível laboratorial, a transposição desses produtos para a investigação e aplicações clínicas tem sido problemático. Em primeiro lugar, existem evidências que na sua maioria, os clones de iPS mantêm algum nível de marcação epigenética reminiscente do tipo celular a partir do qual foram desenvolvidas; em segundo lugar, os métodos de entrega dos factores de reprogramação não integrativos são muito ineficientes, ao passo que os métodos integrativos induzem mutações insercionais, não sendo assim nem seguros nem desejáveis para aplicações clínicas; em terceiro e último lugar, desenvolvimento, expansão e caracterização de clones de iPS é um processo largamente moroso, lento e caro.

Assim, torna-se evidente a necessidade imperiosa de desenvolvimento de protocolos de entrega de factores e reprogramação que permitam uma transposição segura para a clinica e garantam a aquisição de células de melhor qualidade para a investigação, garantindo a qualidade e confiança nos resultados.

Ao aliar a eficiência e reprodutibilidade dos métodos de reprogramação integrativos com a possibilidade de posterior excisão, os transposões, nomeadamente o PiggyBac, devido à sua ínfima taxa de mutação após excisão, surge como vector muito promissor para a administração de factores com vista à reprogramação celular.

O presente trabalho surge neste contexto como uma contribuição para o desenvolvimento de um sistema de reprogramação baseado no transposão PiggyBac que seja simples e rápido, dispensando caracterização molecular após obtenção de iPSs.

Por motivos de limitação de carga o sistema foi dividido em dois vectores a seres co-transfectados na população inicial de fibroblastos. O primeiro vector, *Remo*, é composto pela cassete de reprogramação contendo os factores OSKM e GFP sobre a expressão de um promotor constitutivo (CAG). O segundo vector, *Eneas*, é constituído pelo gene da transposase de codão optimizado para ratinho (mPB) associado a um receptor de estrogénio (ERT2) sobre a regulação de um sistema repressível por tetraciclina (TET-OFF). No extremo 3´ encontra-se ainda o transactivador reverso capaz de regular o promotor TET. Ambos os vectores contêm uma cassete de selecção positiva-negative e são flanqueados por sequências TR 5´ e 3´ específicas do transposão PiggyBac.

O protocolo consiste de duas partes: numa primeira instância, através de um pulso de expressão de mPB ambos os plasmídeos são transpostos para o genoma, permitindo expressão da cassete de reprogramação OSKM; posteriormente, após obtenção de colónias de iPSs, um segundo pulso seria imposto, de modo a permitir remobilização e obtenção de colónias iPS sem integração. Integração e mobilização serão controladas pela cassete de selecção positiva-negativa.

De modo a estabelecer o sistema os seguintes objectivos foram traçados:

a) Clonar ENEAS;
b) Escolha de um sistema de transfecção que permitisse altos níveis de transfecção em fibroblastos humanos.
c) Testar a funcionalidade do sistema em termos de:
    a. Concentração de DOX necessária para reprimir a expressão de mPB;
    b. Concentração de TM ideal para transposição;
    c. Janela temporal de expressão de mPB;

d. Definir a, b e c para o processo de remobilização;
e. Concentração ideal de GCV para selecção negativa.

A clonagem do vector ENEAS foi lograda e confirmada por digestão com quatro enzimas distintas e posterior sequenciação. A funcionalidade de ENEAS foi testada em células HEK 293T. Por RT-PCR provou-se a repressão da expressão de mPB à concentração de 2 µg/ml de DOX. A concentração de TM foi titulada e confirmada quer pela bibliografia quer pelos testes de transposição com diferentes janelas temporais. Três janelas temporais de expressão de mPB foram testadas: 24H, 48h e 72h, revelando-se a ultima a que melhores resultados gerava.

Após obtenção de clones resistentes a puromicina, uma fracção dos mesmos foram expandidos sobre meio selectivo e quatro prosseguiram para caracterização por PCR genómico, de modo a averiguar se se tratavam de clones gerados por transposição catalisada por mPB. Apesar do surgimento de algumas dificuldades devido a amplificação de produtos de PCR na população controlo de 293T, as evidências apontam para que, ainda que nenhum dos quatro clones escolhidos tenha sido obtido apenas por transposição, diferentes intensidades de produtos de amplificação sugerem existência de transposição no sistema.

Ainda que os resultados não tenham coincidido perfeitamente com o esperado, e uma nova bateria de testes tenha que ser desenhada de modo a aumentar o controlo sobre o sistema de transposição, provou-se a funcionalidade do sistema a nível molecular deixando esperança para que, em tempo, o conceito de concretize.

Palavras chave: reprogramação, iPS, PiggyBac, transposão, transposição, clonagem.

TABLE OF CONTENTS

FIGURES LIST

TABLES LIST

ABBREVIATIONS LIST

293T HEK	Human Embryonic Kidney 293 cells
ALS	Amyotrophic lateral sclerosis
BSA	Bovine serum albumin
CAG	Strong synthetic promotor
cDNA	Complementary DNA
CIP	Calf intestinal alkaline phosphatase
CMV	Cytomegalovirus promoter
CV	Crystal violet
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOX	Doxycycline
ERT2	Estrogen receptor
FBS	Fetal bovine serum
GCV	Ganciclovir
Ipsc	induced pluripotent stem cells
IRES	Internal ribosome entry site
LBamp	luria bertani medium with Ampicillin
O.N.	Over night
PB	PiggyBac transposon
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PGK	Phosphoglycerate kinase promoter
Puro	Puromycin
PuTK	Puromycin-Thymidine kinase fusion protein
RNA	Ribonucleic acid
RT	Retrotranscription
rtTA	reverse tetracycline transactivator
SMA	Spinal Muscular Atrophy
TET	Tetracycline
TK	Thymidine kinase
Tm	Melting temperature
TM	Tamoxifen
TR	Terminal repeat
TRE	Tetracycline response element
tTA	Tetracycline transactivator

CHAPTER 1: INTRODUCTION

1. Introduction

Cellular reprogramming is widely considered a very promising field with great potential for regenerative medicine. The possibility of generating a cell capable of making part of all the tissues of an adult body promises a new set of possibilities and strategies for biomedical research and regenerative medicine by enabling access to scarce cell types and theoretically autologous replacement. (Cherry and Daley 2012; Cahan and Daley 2013).

Up to the beginning of the second half of the XX century, it was unknown whether differentiated cells had achieved their state by regulating a complete genome or if they had lost parts of their genome during specialization. In 1962, John Gurdon revolutionized the field of developmental biology by generating cloned frogs by somatic cell nuclear transfer (SCNT), proving, thus, that all genetic information needed to form an entire individual was still present in the differentiated cell and that, somehow, factors present in the oocyte cytoplasm were able to reverse the differentiated state off the somatic cell. Regarding Waddington´s landscape, Gurdon´s findings would represent a movement uphill towards dedifferentiation (Gurdon 1962; Takahashi and Yamanaka 2013; Turksen 2013).

In 1987, Davis and colleagues, using an elegant method of complementary DNA extraction, detected a set of three genes that were predominantly expressed in myoblasts. Exogenous expression of one of those genes, Myod1, alone was sufficient to convert mouse fibroblasts to stable myosin-expressing myoblasts (Davis, Weintraub et al. 1987).

This was the first report of transdifferentiation to be found in the literature. Contrasting with the previously described dedifferentiation, transdifferentiation refers to the direct conversion of one cell type into another without ever reaching the state of pluripotency.

Based on the recent discoveries regarding cell reprogramming and cell fate imposition by a specific factor, Kazutoshi Takahashi and Shinya Yamanaka drew the experiment that would definitely change the paradigm of developmental biology and open a new set of possibilities for the fields of regenerative medicine and disease modeling. They selected a set of 24 genes that might have an effect on pluripotency, according with what was described for ES cell identity at that time, and screened it for

the minimal set of factors that enabled the reprogramming of adult fibroblasts into a pluripotent-like state when expressed ectopically. In this original study, exogenous expression of Oct4, Sox2, Klf4 and c-Myc (OSKM) were enough to reprogram mouse fibroblasts (Takahashi and Yamanaka 2006; Takahashi 2012; Warmflash, Arduini et al. 2012; Takahashi and Yamanaka 2013).

Although the iPS technology was only developed in 2006, the factors used in Yamanaka´s experiment had been known and associated with cellular pluripotency for some time. The study of gene expression associated with the formation of embryonic stem cells through microarray revealed a set of 230 genes with enhanced expression in stem cells relatively to somatic cells. Some of these genes that had already been shown to participate in embryogenesis and ES formation, as well as the factors that would be used to reprogram mouse and human somatic cells (Ramalho-Santos, Yoon et al. 2002; Warmflash, Arduini et al. 2012).

1.1 What happens during reprogramming?

Embryonic stem cells are characterized by indefinite symmetric self-renewal and ability to differentiate into the three germ layers *in vitro* and *in vivo*. Somatic cells are, in turn, defined by a functional state of differentiation in which its identity and function are maintained by an internal lineage and type specific transcriptional and epigenetic status. These two levels of regulation function as a robust embedded system of protection to cell identity, avoiding undesired transdifferention and imposing a division limit, protecting against accumulation of mutations and development of malignancy *in vivo*.

In order to reprogram somatic cells to a pluripotency state, several epigenetic and transcriptional roadblocks, specific of each cell type, must be overcome. Notwithstanding its robustness this mechanisms can now be surpassed by forced expression of key transcription factors, enabling the manipulation of cellular identity and cell fate. Ectopic expression of core pluripotency transcription factors like Oct4 and Sox2 has a dominant destabilizing effect over the existing transcriptional order, overruling the previous epigenetic and transcriptional state, and resulting in a stable phenotypic transformation to an embryonic-like state. Albeit is now current knowledge that reprogramming can still occur in absence of cell division, the co-expression of the

cell cycle regulator, Klf4 and c-Myc, a mitogenic factor, along with Oct4 and Sox2, greatly enhances the overall efficiency of the process. Combined, these four factors cooperatively confer binding stability or specificity, most probably after the "pioneer" function of either Oct4 or Sox2 that recognizes their target sites independently of the previous chromatin state. Once established a stable transcription, the factors act in two main ways: on the one hand, Oct4 and Sox2 activate the expression of NANOG and form an autoregulated nucleus of pluripotency that enhances the expression of other pluripotency factors and paradoxically activates pro-differentiation pathways like MAPK1 and WNT3, and represses differentiation by activating the chromatin regulator Polycomb group (Cahan and Daley 2013); on the other hand, Klf4 and c-Myc activity as mitogens greatly enhances the kinetics of the process most likely by imposing epigenetic modifications inherent to cell division which probably opens a window of opportunity for the establishment of pluripotent-like epigenetic state, facilitating the process of reprogramming (Fig. 1) (Egli, Birkhoff et al. 2008; El-Karim, Hagos et al. 2013).

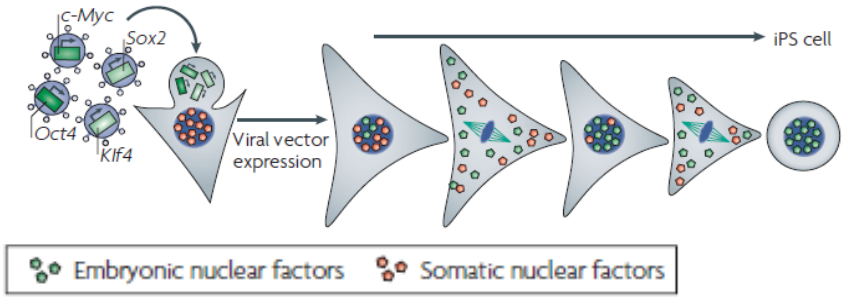

Fig. 1. Reprogramming with defined factors requires gradual replacement of ectopicly expressed factor with the endogenous circuitry of pluripotency. This process seems to be more efficient in rapidly dividing cells adapted from (Egli, Birkhoff et al. 2008).

The whole process of reprogramming is still not fully known, which sustains many doubts regarding its future use in the biomedical field. Consequently, the scientific community designed a series of tests in order to access qualitatively which is the degree of a specific iPS clone. iPSC pluripotency was established in many ways. In order for a line to be classified as pluripotent it has to fulfill at least three of the following tests: a) ES-like gene expression and epigenetics; b) express extracellular markers of pluripotency, for example SSEA1; c) generate cells from the three germ layers *in vitro* or *in vivo*, d) participate in chimera formation and e) generate an embryo

by tetraploid complementation. Note that tests d) and e) cannot be performed in human iPS cells (Cherry and Daley 2012).

Given the potential of this technology, the field has become a scientific hotspot of interest. One technology that is needed is a reproducible method that enables the development of an integration free iPS of clinical grade. Although recent reports claim to reach a reprogramming efficiency of virtually 100% (Rais, Zviran et al. 2013), within the premises of research, and due to its simplicity and robustness, integrative methods, mostly viral vectors similar to the ones used by Yamanaka are still the most common (Egli, Birkhoff et al. 2008; Yamanaka 2012; El-Karim, Hagos et al. 2013; Takahashi and Yamanaka 2013; Turksen 2013).

1.2. Use of induced pluripotent stem cells in Regenerative medicine:

Induced pluripotent stem cells derivation in humans was first reported by Takahashi and Yamanaka in 2007 (Takahashi, Tanabe et al. 2007) and immediately recognized as a landmark breakthrough, extending Gurdon´s reversibility of development. Their basic result was that adult somatic cells were able to undergo a phenotypic makeover by retrovirus mediated over-expression of 4 transcription factors (Oct4, Sox2, Klf4 and c-Myc) resulting in embryonic stem cells like colonies (Yamanaka and Blau). Alike the already known mouse iPS cells, these cells were also found to self-renewal and differentiate in the three germ layers i.e., pluripotency. Since then, a number of methods for direct reprogramming have been developed and some basic characteristics of the reprogramming mechanism established (Hanna, Saha et al. ; Jaenisch and Young 2008; Marson, Levine et al. 2008; Gonzalez, Boue et al. 2011).

Although so far, all the attempts to generate functional specialized cells from embryonic stem (ES) cells have not been entirely successful, there is still the belief that a more comprehensive understanding of cellular identity and specific cellular niches would allow its *in vitro* reproduction with good results. Nevertheless, there are some experiments with mouse ES cells that have already been partly successful, as is the case of mouse blood cells, pancreatic cells and specialized neurons (Cherry and Daley 2012). Taking on these principles and, because environmental influences and epigenomic specificities are mostly erase through long passaging of iPS clones bringing these cells

even closer to ES cells, it is expectable that, in time, patient specific tissues could be derived, enabling , autologous transplantations of *in vitro* differentiated cells, thus avoiding the now necessary use of immune suppressors. Furthermore, in order to avoid resurgence of genetic diseases, homologous recombination of iPS clones could be an option, as proved with the correction of sickle-cell anemia in mouse (Tiscornia, Vivas et al. 2011; Robinton and Daley 2012).

From the point in which we stand, two general approaches can be followed in the future in order to restore a missing tissue or cell type: a) derivation of iPS cells from an easy accessible tissue (for example skin or blood), followed by *ex vivo* differentiation of the desired cell type which would be transplanted and integrated in the desired organ; b) direct conversion of somatic cells into the progenitor cell of the desired cell, thus creating a population that would affect cell regeneration *in situ* rather than directly replacing it. Although these two hypotheses may promise novel medical approaches to several diseases, the existing methods of reprogramming are still not able of creating clinical grade iPSC in a consistent and reproducible way on a clinically relevant amount of time (Cherry and Daley 2012).

1.3. Use of induced pluripotent stem cells in Disease modeling:

The study of human disease in the lab has been limited to studying human cells in vitro or developing animal models of the condition. Human tissue is usually relatively difficult to access, and often, the accessible cell types are not of interest to the disease, limiting the experimental approaches. While animal models have provided powerful experimental systems and allow studies at the organismal and systemic level, they suffer from the intrinsic disadvantage of not being human (Tiscornia, Vivas et al. 2011). The development of murine models tends to be slow and in most cases have been shown not to reproduce the human phenotype faithfully and there is growing realization that a high number of therapies developed in animal models fail in clinical trials (van der Worp, Howells et al. 2010). Perfect example of that is the case of amyotrophic lateral sclerosis and Gaucher´s Disease, in which the animal model may not always recapitulate exactly the human form of the disease, resulting in an active drug in mouse

but rather innocuous in the human case (Farfel-Becker, Vitner et al. 2011; Wu and Hochedlinger 2011).

Along with their potential for regenerative medicine applications, direct reprogramming offers an alternative paradigm for studying human disease: patient specific cells can be reprogrammed, characterized and differentiated to the disease relevant cell type. Thus, large amounts of relevant cell types carrying specific human disease causing mutations can be generated in culture for investigating basic pathogenic mechanism, hypothesis testing, toxicology studies, pharmacological compound screening and development of novel therapeutic approaches. In the few years, this approach has been demonstrated successfully for a number of diseases, including, including ALS, SMA , Rhett Syndrome, Long QT Syndrome, Pompe's Disease and others (Dimos, Rodolfa et al. 2008; Ebert, Yu et al. 2009; Marchetto, Carromeu et al. 2010; Huang, Chen et al. 2011; Itzhaki, Maizels et al. 2011; Tiscornia, Vivas et al. 2011). These encouraging new results demonstrated that the iPS technology makes a good target for modeling monogenic and multigenic diseases, creating the possibility of high throughput pharmaceutical analysis, settle the basis for a pharmacogenomics based medicine.

1.4. Methods of reprogramming:

Over the years, with the development of the field and the increased knowledge on reprogrammed cell behavior, scientists have been pushing forward towards the development and practical use of these cells in biomedicine. For iPSc to be used in regenerative medicine, issues of differentiation, genetic instability and ultimately the risk for tumorigenicity must be addressed. The classic method of reprogramming using retroviral vectors, though practical, is unsuited to clinical practice. Moreover, although the retroviral vectors stay mostly repressed in immature cells like, as is the case of iPSc, when differentiated, these cells might reactivate those inserts, acting as a counter stimulus to the differentiation process decreasing the final yield of the process, preventing in some cases a stable maintenance of the differentiated phenotype and producing an heterogeneous population in various states of differentiation. Likewise, cell culture and maturation protocols must be optimized. Until now, maintaining pluripotent cells and differentiating was a long, complicated and expensive process, which in most of the cases did not guarantee a high percentage of the desired cell type

or state of maturation, requiring several rounds of optimization for each tissue and organism, in order to maximize the yield, quality and thus, *in vivo* integration. Finally, in the case of disease modeling, some genetic diseases need crucial environment in order to develop is characteristic phenotype. Fulfilling these requisites will mark the next step forward in the pluripotency field (González, Boué et al. 2011; Vierbuchen and Wernig 2012).

Since the publication of the first method of reprogramming by retroviral delivery a broad number of new methods and approaches to induce pluripotency in somatic cells have been developed. Alongside with these new techniques a standard array of testes has been defined in order to properly address the contribution and improvement that each method brings onto the field.

As reference before, there are two major research applications when considering the iPS field: the first one is more directed to the basic research fields, focusing on deciphering the mechanisms of reprogramming, differentiation and cell identity and a second one, is more direct towards relevant clinical applications and iPS based therapies. For the present being, a reprogramming strategy that fits all the purposes and ends of application for the field has still not been designed. So, instead, we have an array of strategies that can be used, depending on the final application to obtain a useful iPS population.

In the first case, when clinical grade cells are not required, but rather a fast, efficient and reproducible method is desired, and the presence of foreign DNA or expression of possible tumorigenic factor is tolerable, as is the case of the study of mechanisms of disease and the development of models of disease for drug testing, the use of integrative inducible lentiviruses provides efficient and robust results. In the second case, on the other hand, when clinical applications are considered and avoidance of genomic modifications is strictly necessary for safety reasons. Thus, non-integrative methods represent, until now, the only reasonable option. Non-integrative methods can be divided into four fundamental categories: episomal delivery, RNA delivery, protein delivery, integration-defective viral delivery. Despite the fact that these approaches are theoretically safe, they are difficult to apply, due to poor transfection efficiencies, cell survival, long reprogramming kinetics and possibly other unknown limitations which result in high inefficiency and poor reproducibility (González, Boué et al. 2011).

In this context of need for a high efficiency method that preserves the original genomic sequence, transposable systems, may be good candidates as factor delivery vectors, due to their merged characteristics of integrative vector, with stable expression and relatively high reprogramming efficiency, along with the possibility of being later excised from the genome. Within the several transposons known to transpose in mammal cells, most belong to either the TC1 family, like Sleeping Beauty; the hAT-like Tol2; Frog Prince, or to the PiggyBac superfamily. The PiggyBac transposon was first isolated from cabbage looper moth *Trichoplusia ni* in 1989. Nevertheless, only recently has emerged as a reliable tool in molecular biology for cargo delivery and insertional mutagenesis, quickly becoming the option of choice due to its cargo capacity, absence of transposase overexpression repression and, most of all, its perfect cut and paste mobilization, restituting the original sequence in over 98% after remobilization (Fig. 2) (Li, Pettitt et al. 2013). The first iPSCs reprogrammed with a PiggyBac transposon were obtained by Woltjen et al in 2009. Derivation insert-free iPSc enables the establishment of better cellular models of iPS and adds a new level of security to the use of these cells in biomedical applications. Nevertheless, due to the fact that it was based on several low efficiency steps consisting of consecutive transfections, with plasmid dilution overtime resulted in an overall low efficiency. Moreover, the constant possibility of random integration forces a molecular characterization of each clone, which ended up being laborious, time demanding and expensive (Kim and Pyykko 2011).

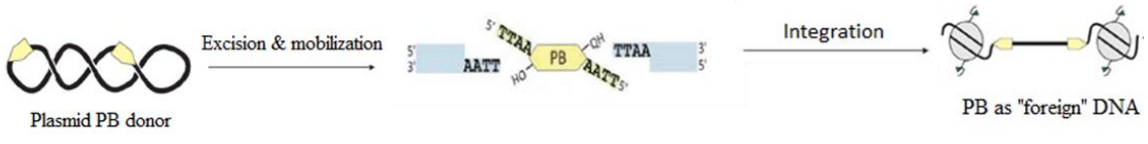

Fig. 2. Model of PB cut and paste mobilization from a plasmid. PB transposase recognizes and integrates into TTAA sequences with preference for euchromatin regions. Adapted from (Li, Pettitt et al. 2013).

1.5. Objectives

The objective of the present work is to provide an improved and simplified PiggyBac system for reprogramming. We developed a two vector system which is simultaneously delivered to the original population of fibroblasts. The first vector, ***Remo,*** carries the reprogramming cassette and GFP under the regulation of a constitutive promoter (CAG). The second vector, *Eneas*, carries the piggyback mouse transposase (mPB) associated with an estrogen receptor fragment (ERT2), regulated in a Tetracyclin repressible (TET-OFF) fashion, and its equivalent reverse trans-activator associated with a positive-negative selection cassette under a constitutive promoter. Both constructs carry a positive-negative selection cassette and are flanked by the PiggyBac 5´ and 3´terminal repeats. This method is designed to be time and work efficient, hopefully resulting in insert-free iPSc within six weeks. The protocol is divided in two parts: first, through expression pulse of mPB, both Remo and Eneas get transposed into the genome, allowing expression of the OSKM in a stable manner; secondly, after obtaining iPSCs a second pulse of mPB expression will remobilize the inserts, resulting insert-free iPSCs. Integration and remobilization will be controlled recurring to the positive negative selection of PuroTk positive-negative marker, being that, during the first pulse of mPB expression, cells carrying the insert will become puromycin resistant and after the second pulse, only cells that have lost all the insert will be resistant to ganciclovir. Ultimately, iPS derivation will be achieved simply by controlling the media without the need for clonal expansion and further characterization. This transposable system would, in theory, work the following way:

1) A plasmid containing a piggyback transposon would be delivered to the cells to be reprogrammed.
2) The transposon would contain the following genetic elements:
    a) A policystronic reprograming cassette composed of human Oct4, Sox2, Klf4, c-Myc and GFP linked by 2A self-cleaving peptides driven by a constitutive CAG promoter.
    b) A constitutive PGK promoter driving a tTA transactivator (TET-OFF system), followed by an IRES element and a PuTK gene fusion for positive-negative selection.

c) A TET responsive CMV minimal promoter expressing mPB transposase linked to a tamoxifen inducible nuclear translocation ERT2 element.
d) Elements a, b and c would be flanked by transposon TR.

After transfection, the cells would be subjected to a limited pulse of no DOX and TM. PGK would constitutively express the tTA protein. During the windows of no DOX/TM treatment, the absence of DOX would cause the tTA protein to activate expression of mPB-ERT2. TM would induce translocation of the mPB-ERT2 protein to the nucleus, where it would catalyze transposition of the transposon into the host genome. The transposition pulse would be terminated by adding DOX and withholding TM from the medium. Theoretically, the remainder of the plasmid backbone would be lost by dilution. Constitutive expression of the reprogramming cassette would result in iPSc colony formation. These colonies would constitutively express both puromycin resistance (positive selection marker) and TK (negative selection marker). As iPSc reprogramming is in itself a selective event, puromycin selection would be unnecessary. Once the iPSc line was obtained, it would be subject to a second pulse of DOX/TM, which would result in re-expression of mPB-ERT2, translocation to the nucleus and excisional transposition of PB out of the genome. As this event can result in either permanent excision or reinsertion into the genome, the cells would be selected with ganciclovir for clones having lost the transposon and therefore no longer expressing the PuTK fusion. Therefore, an integration free iPSc line would result. Fine tuning of the system would require:

8) Obtaining acceptable transfection efficiency into human fibroblasts.
9) Testing the functionality of the construct
10) Determining the ideal concentration of DOX for repressing mPB-ERT2 expression
11) Determining the ideal concentration of TM for transposition into the genome
12) Determining the ideal Windows of no DOX/TM pulse for transposition into the genome
13) 3, 4 and 5) for transposition out of the genome
14) Determination of the ideal concentration of GCV for negative selection

An initial problem with the system was that the cargo of genetic elements required for the strategy (approximately 16 kb) was well over the carrying capacity of the PB transposon (approximately 8 kb). If cargo exceeds the PB carrying capacity, transposition efficiency drops markedly. Therefore, a two PB system was designed: One transposon would deliver the reprogramming cassette while the second would deliver the TET-OFF system and the mPB-ERT2 elements. Both transposons would express the PuTK positive/negative selection elements. Hence, the strategy would involve delivery of two transposons, induction of transposition, reprogramming, induction of excision and finally selection for ganciclovir resistant clones.

CHAPTER 2: MATERIALS AND METHODS

2. Materials and Methods

2.1. Cloning of EneasI

This system was designed to work with 5´PTK3´TETO-CMV-mPB-L3ERT2-hPGK-tTA hereby known as ENEAS in association with a 5´CAG-OSKM-GFP 3´, hereby REMO I. Partial construction of ENEAS had advanced to an intermediate stage where the only element missing was the TRE-CMV-mPB-ERT2 cassette. This cassette had been previously constructed (plasmid pCR BT TETO CMV-mPB-ERT2).

To conclude the cloning of ENEAS the following steps were taken, as schematized on Fig.3, and described next. The TRE-CMV-mPB-ERT2 cassette was excised from the plasmid pCR BT TETO CMV-mPB-ERT2 by digestion with restriction enzymes NHeI and ClaI. The recipient vector was digested with NheI and ClaI and then treated with CIP to prevent concatemerization. Both fragments were ligated using T4 DNA ligase and transformed into elctrocompetent bacteria which were then selected in LB agar with ampicillin (1µg/mL). Plasmid DNA was extracted from 18 ampicillin resistant colonies and analyzed by restriction digest with BamHI, HindIII and NcoI. All the reagents were commercially acquired from *New England Biolabs*.

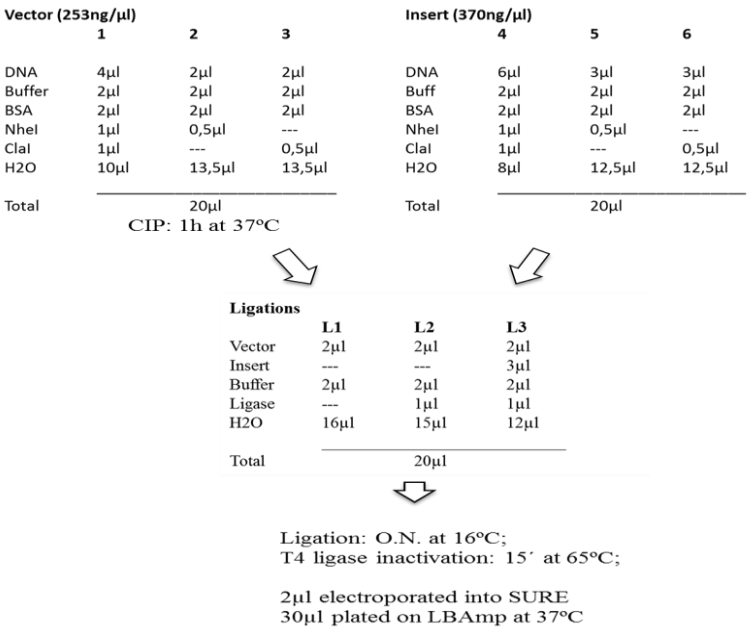

Fig. 3. Experimental setup of ENEAS cloning procedure. First both plasmids were digested with NheI and ClaI: 1 and 4- double digestion; 2-3 and 4-6 -single digestion controls (On the top). Secondly, both isolated insert and destination vector were ligated using T4 ligase (midle). Finally, the ligation reaction was electroporated into SURE cells which were selected for ampicillin resistance.

2.2 Characterization of ENEAS

In order to test the functionality of the system and determine if the construct functions as predicted and fine tune it accordingly with the objectives, i.e., obtain an integration efficiency that maximizes the number of insert bearing cells with minimal number of individual integrations which would enable an optimal loss of integration thoμgh mobilization, thus maximizing the efficiency of the system. As a matter of practicality the tests were performed in 293T HEK cells with a passage number inferior to 20. The first part of the characterization (1) is designed to guarantee the responsiveness of the system to DOX, TM, evaluate the ability to generate Puromycin resistant colonies in different conditions. In a second phase (2), the obtained clones were physically isolated and expanded in order to determine the conditions of remobilization, and selection with GcV. Fine tuning was performed using the following tests:

A. Obtaining acceptable transfection efficiency and test the functionality of the construct
    1) Determining the ideal concentration of DOX for repressing mPB-ERT2 expression
    2) Determining the ideal concentration of TM for transposition into the genome
    3) Determining the ideal Windows of no DOX/TM pulse for transposition into the genome

B. Fine tuning transposon excision:
    1) Clonal expansion of puro resistant clones and diagnostic genomic PCR
    2) Repeat tests A1, A2 and A3 for excision
    3) Determining the ideal concentration of GcV for negative selection.

A. Fine tuning of integration and working conditions

Initially, we determined the minimal amount of puromycin concentration required to kill at least 50% of non-transfected 293T cells. To do so, 500,000 293T cells were seeded in a 10 cm tissue culture dish and subjected increasing concentrations of puromycin (0, 250, 500, 1000 and 2000 μg/ml). Cells were observed for 10 days. Cells subjected to no puromycin or 250μg/ml puromycin continued to grow and achieved confluency after 5 days. A concentration of 250 μg/ml resulted in aproximately 80%

cell death, while higher concentrations of puromycin obliterated the whole cell population. Therefore, we chose 500µg/ml as our puromycin selection concentration.

1) Determining the ideal concentration of DOX for repressing mPB-ERT2 expression

We set up a series of experiments with the goal of determining the correct procedure to obtain transposition of PB into the genome of our target cells. Proceeding by trial and error, after a few attempts we determined that the following procedure should be followed:

1) Seed 200,000 293T cells in a 35 mm dish, incubate overnight
2) Next morning add the appropriate amount of DOX (0, 2 and 10 µg/ml) and TM (1 µM) to each well
3) After six hours of pre-incubation in DOX/TM conditions, transfect 100ng of ENEAS into each dish, using Fugene6 (*promega*), following given protocol.
4) Maintain the conditions for 72 hs
5) Add DOX and eliminate TM (+DOX/-TM) from all dishes in order to repress further transposition
6) Trypsinize and passage the cells into 10 cm dishes, maintain +DOX/-TM conditions and select for puromycin resistance with 500 ng/ml.
7) Harvest the cells and Proceed to RT-PCR.

RT-PCR

The total RNA of each of the conditions 0, 2 and 10 µg/ml of DOX was extracted using the Trizol protocol for RNA extraction and retro-transcribed into cDNA using random hexamers as primers. Once obtained, the cDNA served as a PCR template for two sets of primers in order to amplify two fragments: an ERT2 region of 300bp and a fragment of the transactivator (tTA). Both PCRs were catalyzed by a Go Taq polymerase (*promega*), using the brand suggested cycling conditions with a melting temperature of 57ºC and 35 cycles.

2) Determining the ideal concentration of TM for transposition into the genome

To evaluate the effect of TM in transcriptionally active clones we designed the following experiment:

1) Seed 200,000 293T cells in a 35 mm dish, incubate overnight
2) Next morning add the appropriate amount of TM (0,25µM, 0,5µM, 1mM and 1,5mM) to each well
3) After six hours of pre-incubation in DOX/TM conditions, transfect 100ng of ENEAS into each dish, using Fugene6 (*promega*), following given protocol.
4) Maintain the conditions for 72 hs
5) Add DOX and eliminate TM (+DOX/-TM) from all dishes in order to repress further transposition
6) Trypsinize and passage the cells into 10 cm dishes, maintain +DOX/-TM conditions and select for puromycin resistance with 500 ng/ml.
7) Proceed to crystal violet (CV) coloration.
    a. Wash carefully the colonies with PBS;
    b. Fixate cells for 1 min with 4% PFA;
    c. Remove PFA and add 0.1% CV solution for 30 minutes ;
    d. Remove CV and wash with tap water.

3) Determining the ideal Windows of no DOX/TM pulse for transposition into the genome

Having both the DOX and TM titrated to working concentrations, we tried to determine the mPB expression pulse that would result in consecutive appearance of colonies consistent with what was expected. Three expression windows were tested: 24h, 48h and 72h. Although colonies were evident in all time points, the 72h window rendered more consistent and evident differences, being repeated five times with similar outcomes. Following the outline of the previous experiments, a set of conditions, was created base on table 1.

Table 1. Set of conditions used to optimize the window of expression for ENEAS transposition.

Well	#1	#2	#3	#4	#5
ENEAS	---	100ng	100ng	100ng	100ng
DOX	---	---	2µg/mL	---	2µg/mL
TM	---	---	---	1µM TM	1µM TM

The procedure was the following:

1) Seed 200,000 293T cells in a 35 mm dish, incubate overnight
2) Next morning add the appropriate amount of DOX and TM (table 1) to each well
3) After six hours of pre-incubation in DOX/TM conditions, transfect 100ng of ENEAS into each dish, using Fugene6 (*promega*), following given protocol.
4) Maintain the conditions for 72 hs
5) Add DOX and eliminate TM (+DOX/-TM) from all dishes in order to repress further transposition
6) Trypsinize and passage the cells into 10 cm dishes, maintain +DOX/-TM conditions and select for puromycin resistance with 500 ng/ml.
7) Proceed to crystal violet (CV) coloration.
    a. Wash carefully the colonies with PBS;
    b. Fixate cells for 1 min with 4% PFA;
    c. Remove PFA and add 0.1% CV solution for 30 minutes ;

Remove CV and wash with tap water

B. Fine tuning transposon excision

After establishing all the previous parameters and be able to consistently reproduce the same result, consisting with the theoretic prevision, the protocol A-3 was once again repeated. Instead of proceeding to CV coloration, 12 individual clones were physically isolated and expanded.

1) Clonal expansion of puro resistant clones and diagnostic genomic PCR

The isolated clones were kept under puro selection (500ng/ml) and DOX repression (2µg/ml) for two passages and then frozen in complete DMEM+ 10% DMSO. Genomic DNA was extracted from four of the puro resistant clones, using a *Quiagen* QIAamp DNA Mini Kit. Samples were tested for integration by PCR, using a KAPATaq hotstart

polymerase (from KAPABIOSYSTEMS). The reaction was setup as described in table 2, using the commercial cycling parameters with a melting temperature of 52ºC and 35 cycles. The resulting PCR products were ran in a 2% agarose gel and images were acquired using a Chemidoc imaging system (BioRad).

Table 2. KAPATaq PCR setup

KAPATaq PCR reaction		
	Final concentration	Volume
Buffer	1x	5µl
Forward primer	1µM	0,25µl
Reverse primer	1µM	0,25µl
$MgCl_2$(25mM)	5µM	0,5µl
dNTPs	0,2mM	0,5µl
H_2O		11,05µl
DMSO	5%	1,25µl
Template DNA	50ng	1µl
KAPATaq Polymerase	1 unit	0,2µl
Total		25µl

.

CHAPTER 3: RESULTS

3. Results

This sytem had been already partially designed and constructed before the comencement of this project. The PB transposon carrying the reprogramming cassette and the rtTA-IRES-PuTK elements (called Romulus) had already been completed (see fig 4). The second PB transposon, carrying the mPB-ERT2 and the rtTA-IRES-PuTK elements (called ENEAS) was partially constructed. The first task of this project was to complete the cloning of ENEAS.

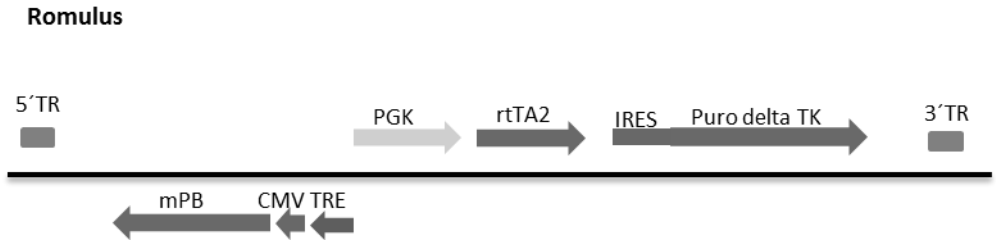

Fig. 4. Schematic representation of Romulus.

3.1. Cloning of ENEAS:

The cloning strategy is illustrated in Fig 3. Partial construction of ENEAS had advanced to an intermediate stage where the only element missing was the TRE-CMV-mPB-ERT2 cassette. This cassette had been previously constructed (plasmid pCR BT TETO CMV-mPB-ERT2), Fig.5. Therefore, the TRE-CMV-mPB-ERT2 cassette was excised from the plasmid pCR BT TETO CMV-mPB-ERT2 by digestion with restriction enzymes NHeI and ClaI and gel purified. The recipient vector was similarly digested with NheI and ClaI and gel purified. Both DNA fragments were ligate using T4 DNA ligase and transformed into elctrocompetent bacteria. Plasmid DNA was extracted from a number of ampicillin resistant colonies and analyzed by restriction digest with HindIII (fig.6A). Several clones with the expected restriction fragment pattern were obtained. One particular cloned was amplified and further analyzed by additional restriction digest analysis (fig. 6B) and confirmed by sequencing to have the correct sequence in all genetic elements (data not shown).

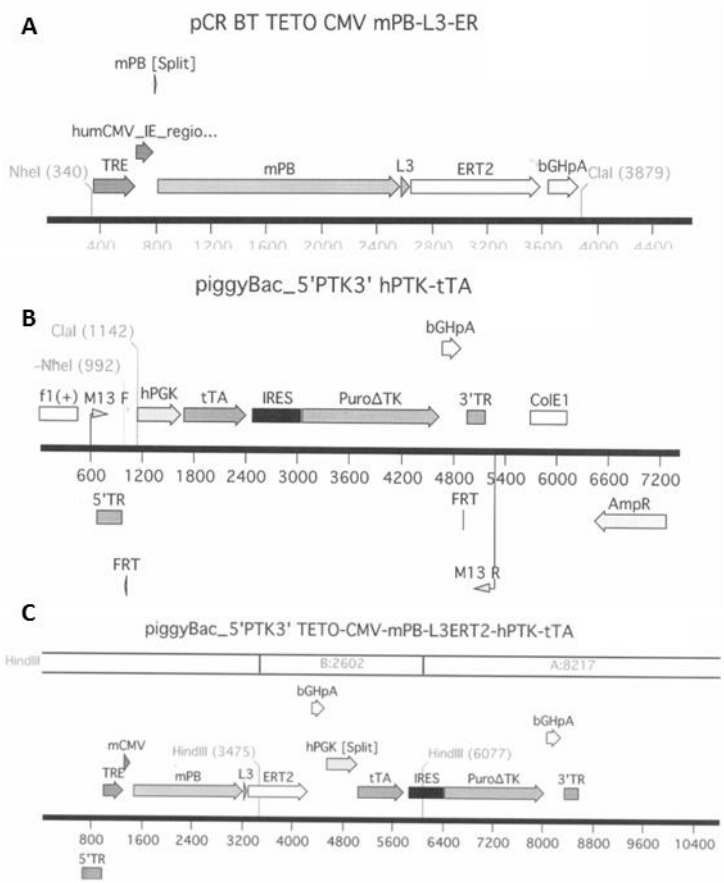

Fig. 5. Schematic representation of the genetic constructs used to clone ENEAS (A and B) and ENEAS (C). In C we can also observe the expected restriction fragments of a restriction with HindIII (8,2Kb and 2,6Kb).

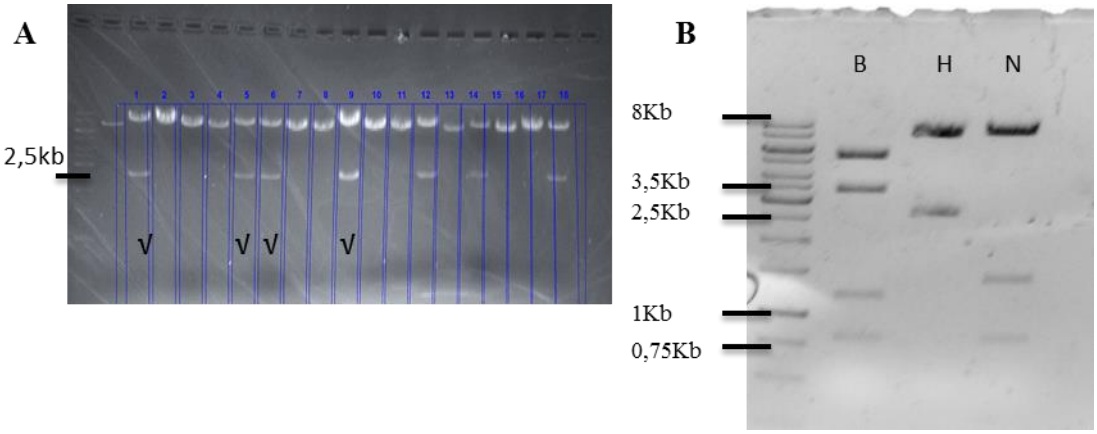

Fig. 6. A: Restriction digest of the 18 ENEAS clones with HindIII. Seven of the 18 clones presented the expected band of 2,5Kb. The marked clones were expanded and saved for further analysis.. B:Diagnostic restriction, cuts of Eneas with BamHI (B), HindIII (H) and NcoI (N). All the bands correspond to the expected sizes: B- 5,47Kb, 3,36Kb 1,19Kb and 0,796Kb; H- 8,22Kb and 2,Kb; N- 8,6Kb, 1,39Kb and 0,798Kb.

3.2. Characterization of ENEAS

Is the mPB-ERT2 cassette expression responsive to DOX?

We next sought to determine whether ENEAs would work as predicted. In a first step, we asked whether mPB-ERT2 expression could be controlled by manipulating the levels of DOX in the medium. To do so, we transiently transfected ENEAS plasmid into 293T cells using the Fugene6 transfection reagent (Roche) and cultured the cells for 72 hs in absence of DOX and in two concentrations of DOX (2 µg/ml and 10 µg/ml). Cells were washed, harvested and extracted for total RNA, which was then analysed by RT-PCR.

The results of the analysis (see Fig 7), indicated that strong expression of mPB-ERT2 was obsevable in absence of DOX, while strong repression of expression was obtained with both DOX concentrations used. In contrast, and as expected, constitutive expression of tTA was unaffected by presence of absence of DOX in the media. Furthemore, we confirmed that the mPB-ERT2 band observed in absence of DOX (300 bp)indeed represented the correct sequence, as digestion with HindIII restriction enzyme yeilded 2 bands of the expected length (176 bp and 124 bp).

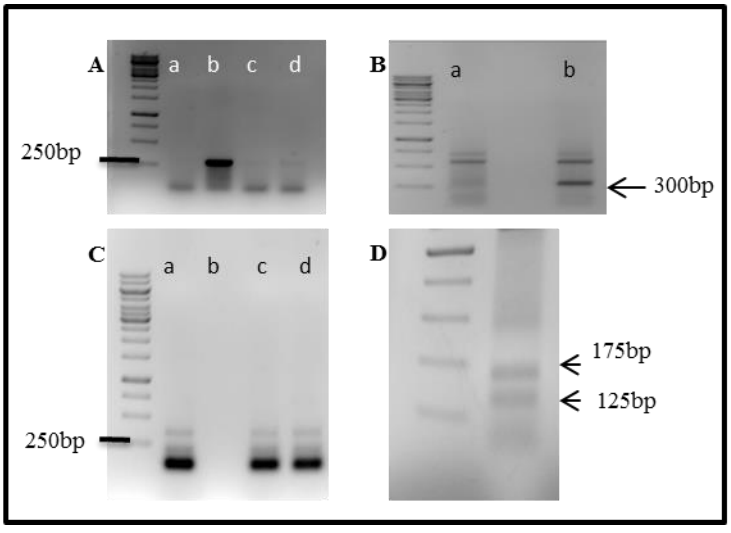

Fig. 7.Panel showing the transcriptional analysis of ENEAS in 293T HEK cells through amplification of na ERT sequence (Aand B) in order to check repression by DOX, and amplification of a tTA to confirm constitutive expression of the transactivatior. The amplification o ERT was repetad as in A. The sample (b) was run against a control (a) of picture B, and simultaneously cut with HindIII to confirm identity of the amplicon (picture D). A- the control (a) of 293T cells without transfection of ENEAS was run against ENEAS clones without any DOX in the media (b), and respectively, with 2 and 10 µg/ml of DOX (c and d). C- expression of the tTa sequence was compared between a non-transfected 293T HEK population (b) with transfected cells without DOX (a), with 2 and 10 µg/ml of DOX (c and d). In the gels A, B and C a 1Kb ladder was used, whereas in D, due to to the small size of the digestion products, a 100bp ladder was used.

3.2.1 Does ENEAS function as expected in an *in vitro* cell transfection experiment?

We sought to establish the parameters for correct functioning of ENEAS for obtaining transposon containing colonies of cells. For reasons of practicality, we chose to work with 293T HEK cells. If ENEAS worked as designed, when transfected into cells its behavior should depend on the presence or absence of DOX in the medium. In absence of DOX (but presence of TM), PB should transpose into the genome of recipient cells and create puromycin resistant clones. In contrast, in the presence of DOX (and presence of TM), the mPB-ERT2 cassette should not be expressed and therefore transposition should not occur and puromycin resistant colonies should not be obtained.

Initially, we determined the minimal amount of puromycin concentration required to kill at least 50% of non-transfected 293T cells. To do so, 500,000 293T cells were seeded in a 10 cm tissue culture dish and subjected increasing concentrations of puromycin (0, 250, 500, 1000 and 2000 µg/ml). Cell were observed for 10 days. Cells subjected to no puromycin or 250 µg/ml puromycin continued to grow and achieved confluency after 5 days. A concentration of 250 µg/ml resulted in aproximately 80% cell death, while higher concentrations of puromycin obliterated the whole cell population. Therefore, we chose 500 µg/ml as our puromycin selection concentration.

We set up a series of experiments with the goal of determining the correct procedure to obtain transposition of PB into the genome of our target cells. Proceeding by trial and error, after a few attempts we determined that the following procedure should be followed:

1) Seed 200,000 293T cells in a 35 mm dish, incubate overnight

2) Next morning add the appropriate amount of DOX (0, 2 and 10 µg/ml) and TM (1 mM) to each well

3) After six hours of pre-incubation in DOX/TM conditions, transfect 100 ng of ENEAS into each dish

4) Maintain the conditions for 72 hs

5) Add DOX and eliminate TM (+DOX/-TM) from all dishes in order to repress further transposition

6) Trypsinize and passage the cells into 10 cm dishes, maintain +DOX/-TM conditions and select for puromycin resistance with 500 µg/ml.

After 2 weeks of puromycin selection, puromycin resistant colonies were aparent and styained with crystal violet. Results of a typical experiment are shown (Fig 8):

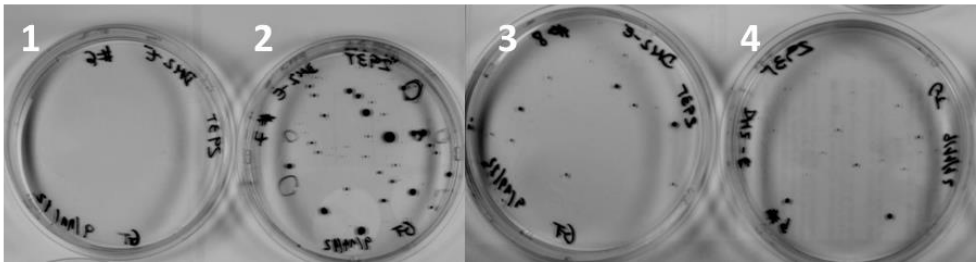

Fig. 8. Crystal violet coloration of Puro resistant colonies of 293T cells transfected with Eneas (100ng). Dox titration to shutdown mPB expression. 1- No Dox no Eneas; 2- Eneas no Dox; 3- Eneas+ Dox 2µg/mL; 4- Eneas + Dox 10 µg/mL; 5- Eneas + Dox 20 µg/mL.

As shown in Fig.8, the system behaved reasonably as expected. When ENEAS was transfected into 293T HEK cells and the cells subjected to a 72 hr window in absence of DOX and presence of TM, a number of puromycin resistant clones appeared in the dish. However, if 2 or 10 µg/ml of DOX were added to the medium during the 72 hr window, the number of puromycin resistant colonies droped significantly, although not to zero. These results showed that overall, the system seemed to be working as expected, and that the number of colonies correlated with the presence or absence of DOX in the culture medium. This result led us to bilieve that the colonies that were appearing in absence of DOX were due to transposition events.

3.2.2. What is the effect of TM on the system?

In our strategy, the levels of transposition are designed to be controlled by two parameters: the level of DOX in the culture medium (which would determine the level of expression of mPB-ERT2) and the presence or absence of TM (which would enhance the cytoplasmic-nuclear transport of mPB through interaction with the ERT2 domain of the fusion protein. The experiment described in the previous section was conducted in the presence of TM, ie, in conditions that maximized the translocation into the nucelus of whatever amount of mPB-ERT2 was being expressed. We then asked what were the relative contributions of each of the two levels of control. To determine this, we set up the following experiment:

Well #1: ENEAS not transfected

Well #2: ENEAS transfected, 0 µg/ml DOX, no TM

Well #3: ENEAS transfected, 2 µg/ml DOX, no TM

Well #4: ENEAS transfected, 0 µg/ml DOX, 1mM TM

Well #5: ENEAS transfected, 2 µg/ml DOX, 1mM TM

As before, the number of puromycin resistant colonies surviving after two weeks of puromycin selection was scored by crystal violet staining. The results are shown in Fig 9.

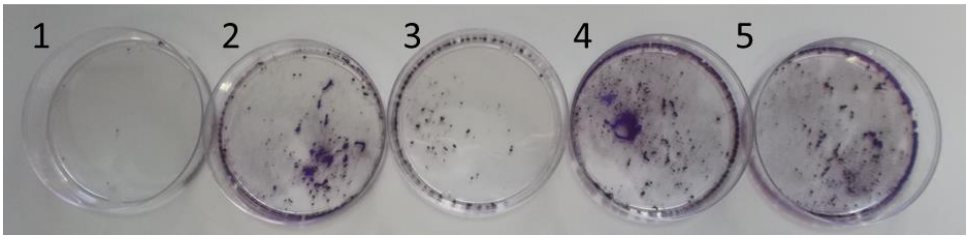

Fig. 9. Crystal violet coloration of colonies obtained after the selection with puromycin (puro) of the 72h expression window. Number 1 represents the negative control, consisting of a 293T HEK population treated with puro. Numbers 2 through 5 represent populations transfected with ENEAS (100ng) and treated with the following conditions: 2- no DOX and no TM; 3- Dox without TM; 4- TM without DOX; and 5- DOX and TM. The concentrations were for DOX 2µg/ml and 1mM for TM.

The overall number of violet colonies (overall 'violetness' of the dishes) provides a semiquantitative measure of transposition efficiency. The results suggest that in absence of TM in the medium, most of the regulation is afforded by the presence or absence of DOX in the culture medium, while adding TM enhances the level of transposition generally, but not by a large amount.

Are the puromycin resistant clones obtained due to transposition events or random integration?

While the behavior of the system suggested that the puromycin resistant colonies we were obtaining were due to PB transposition events, ie, higher number of colonies when mPB-ERT2 was expressed (absence of DOX) and translocated to the nucleus (presence of TM), it was important to verify that this was indeed the case, as the whole strategy of achieving non-integrative iPSc colonies would depend crucially on whether the reprograming cassette had been integrated into the genome by a transposition event or not.

In order to do this, 12 individual puromycin resistant colonies were individually isolated and expanded. Genomic DNA was extracted from 4 lines and we attempted to determine what type of integration had occured by scoring for presence or absence of different regions of the plasmid originally transfected into the cells by PCR, as shown in Fig 10.

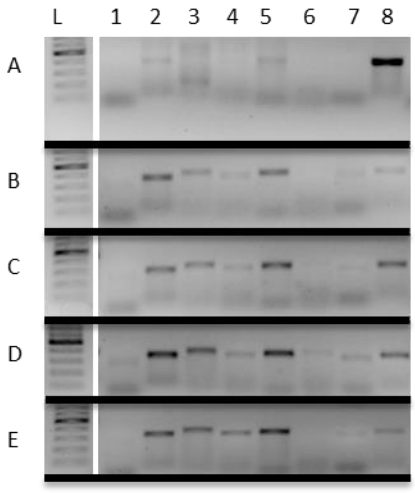

Fig. 10. Composed figure of the genomic PCRs performed on the four individual puro-resistant clones (B, C, D and E) and the control for each of the PCR primer pairs (A), using 293T HEK genomic DNA. The PCR primer pairs are represented with the numbers 1 to 8 and represent PCR1 to PCR8

In order to score for presence or absence of different regions of the original plasmid, primers for 8 PCR products were designed to scan the full length of the plasmid. All PCR products were between 300 and 400 bp long. PCR product 1 was upstream of the PB 5´TR. PCR products 2, 3, 4, 5 and 6 spanned different regions internal to the PB TR's, ie, mapped within the transposon and PCR products 7 and 8 mapped downstream of the PB 3´TR.

Results are shown in Fig.11.

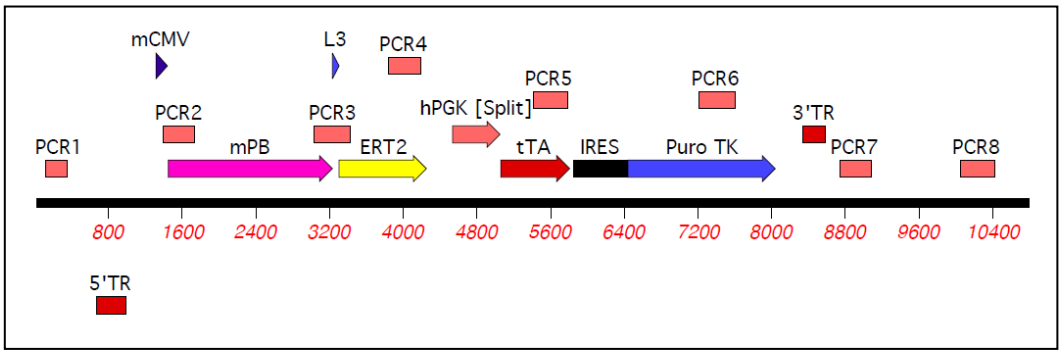

Fig. 11. Schematic representation of Eneas with the amplification region of each of the diagnostic primer pairs (PCR1 to PCR8)

3.2.3. Several points are worth mentioning:

Despite repeated attempts, we were unable to get a clean negative result using genomic DNA from untransfected 293T HEK cells. As can be seen in panel A, PCR products 2 and 5 seemed slightly positive, suggesting that early on a contamination of ENEAS into non transfected cells occured. Unexpectedly, a strong band corresponding to PCR product 8 was evident, suggesting the presence of this plasmid region in the original cell line, presumably due to the original manipulations performed on 293T HEK cells in order to immortalize them. Of note, PCR product 1 was absent from all 4 lines tested, while PCR products 2, 3, 4 and 5 (representing sequences internal to the pB TRs) were uniformly positive in all 4 lines. PCR product 6, mapping to a CG rich sequence in the PuTK region of the transposon did not amplify its product, except for a faint band in clone 3. We surmise that the primers performed poorly to the CG rich nature of the target. PCR product number 8, showing a strong band in the negative (panel A), was also postive in all four clones and therefore uninformative. PCR product 6, also mapping to plasmid sequences downtream of the 3´TR, was absent from the negative (panel A), but also present in the four clones analyzed. The PCR results analyzing the 5´TR end of the construct seem clear: plasmid sequences upstream of the 5´TR are absent, while PCR products 2, 3, 4 and 5 downstream of the 5´TR (within the transposon), are uniformly present in all 4 clones, as would be expected from a transposition event. The situation is less clear at the 3´TR end of the construct due to the amplification of PCR products 7 and 8, which could be due to a) presence of plasmid sequences in the original 293T HEK cells, or b) presence of a background level of random integration of plasmid sequences in addition to a transposition event. In sum, and though the experiment was not air tight, the results are consistent with integration by transposition in a genetic background already containing plasmid sequences.

3.3. What are the conditions of DOX, TM and duration of pulse required to mobilize the transposon integration out of the host genome?

We then proceeded to attempt to provide our puromycin resistant lines with a second pulse of expression of mPB-ERT2 in order to mobilize the transposon out of the host genome rendering it resistant to GCV.

To do so, a puromycin resistant line was chosen. Cells were seeded and given a similar pulse of absence of DOX combined with 1mM tamoxifen for 72 hs. Then DOX 2 µg/ml was added and TM eliminated from the culture medium to stop further transposition. Control cells recieved no expression pulse. Both control and tester cells were cultured in presence of GCV. Our expectation was that cells that were puromycin resistant would also express TK and therefore be sensitive to GCV, while cells that had lost the transposon would be resistant to GCV. Unexpectedly, cells that had not been subject to the transposon-mobilization protocol were unaffected by presence of GCV, and therefore our negative selection strategy failed to work in this particular experiment.

CHAPTER 4: DISCUSSION AND CONCLUSION

4. Discussion

One of the first preoccupations in the field of direct reprogramming after its discovery in 2006 was how to develop reprogramming technologies that would avoid the derivation of iPSc lines harboring exogenous genes or having undergone insertional mutagenesis. The initial Yamanaka approach used individual retroviral vectors expressing OCT4, SOX2, KLF4 and c-Myc. The resulting iPSc had multiple insertions of retroviruses which despite having been silenced in the pluripotent state remained a potential problem in terms of potential cell therapy applications. A number of approaches were tested by many groups worldwide to reduce the risks associated with exogenous gene sequences and insertional mutagenesis. Some of the main approaches have been reviewed (see Gonzalez et al— Reprogramming a la carte) are summarized below:

1) In order to reduce the number of insertional mutagenesis events resulting from the use of multiple individual retroviral factors, reprograming lentiviral vectors were designed carrying polycistronic reprogramming cassettes consisting of several reprogramming genes linked by IRES or 2A self-cleaving peptide sequences. This approach minimized the risk of insertional mutagenesis to a minimium of one lentiviral insertion (ref).

2) LoxP recombination sites were inserted into the retroviral vectors; once iPSc lines were derived, transient expression of CRE recombinase was used to excise the reprogramming cassette. This approach eliminated the problem of presence of exogenous sequences, but not the problem of insertional mutagenesis.

3) Non integrative viral vectors were used, such as Sendai Virus or Adenovirus. While these approaches were successful, the reprogramming efficiency was exceedingly low.

4) Reprogramming was attempted by expression of transgenes from transiently expressed plasmids, under the rationale that perhaps a strong initial reprogramming pulse obtained without insertion of reprogramming sequences into the genome would suffice. However, this approach had low efficiency in mice and did not work in human cells.

5) Use of episomal vectors based on the Epstein Bar virus. Plasmids bearing the EBNA1 origin of replication and a reprograming cassette are introduced into the cells by electroporation and are maintained episomaly. While they are present, they maintain a reprogramming pulse, but eventually they are lost, resulting in integration free iPSc.

6) Use of synthetic mRNA for the reprogramming factors. Messenger RNAs specifically synthesized to avoid the PKR response are transfected into cells to achive reprogramming.

7) Transduction of recombinant protein has been reported, but has very low efficiency.

8) Use of the PB transposon: this system involved transfection of two plasmids into cells to be reprogrammed. The first plasmid contained a PB transposon carrying a reprograming cassette. The second plasmid expressed mPB transposase. This resulted in transposition of the reprogramming PB into the genome and derivation of iPSc. In order to eliminate the transposon, the PB transposase coding plasmid was re-delivered to the iPSc and lines subcloned and tested for loss of the PB transposon by southern. In addition, insertion sites were mapped and analyzed to determine if any mutation event persisted.

All of these approaches suffer from being inefficient, laborious and/or expensive. Therefore, we set out to design a system that would be simple, efficient and inexpensive, and that would result in iPSc lines free of reprogramming factors and insertional mutagenesis events.

We aimed to design a PB based reprograming strategy in which all the elements would be in place to a) express the reprogramming genes and b) express the genetic elements that control PB transposition by simple manipulations in the culture medium.

According to the design, the PB transposon would contain 3 elements:

1) A reprograming cassette

2) A constitutive promoter driving the tTA transactivator and a PuTK cassette and

3) A DOX inducible promoter driving the piggybac transposase fused to an ERT2 nuclear localization domain.

Once this plasmid was transiently delivered to the cells, a window of culture conditions favoring mPB-ERT2 expression and translocation to the nucleus would allow the reprograming transposon to insert into the host genome. Stable expression of the reprograming cassette from its genomic locus would result in reprogramming to the iPSc state. So far, our system is identical to the method of Yusa et al (2009) in that the reprograming cassette is delivered to the genome by use of a piggybac vector. However, in the system of Yusa et al, transposase was delivered transiently twice. While the first delivery is trivial, the second delivery implies a round of subcloning and characterization which is extremely laborious. In our system, two new elements are introduced into the design. The first is a PB transposase inducible expression cassette, which would allow expression of transposase by simple manipulation of the culture medium by addition or elimination of DOX and TM. Manipulation for expression of transposase would make induction of transposition into or out of the genome trivial. The second element is a negative selection cassette expressing TK. A cell expressing TK will be sensitive to GCV. Once iPSc were derived, a second window of transposon mobilization would be applied to the cells and clones having lost the transposon would be selected for resistance to GCV.

A first problem is that all the elements required for the strategy do not comfortably fit in the PB transposon. The full cargo is around 16 kb, well above th 8 KB maximum cargo capacity of PB. A modified version of PB has been reported in which mutation of the PB TRs and optimization of the transposase codon usage has resulted in a system capable of mobilizing up to 18 kb of cargo, but this system is proprietary and unavailable to us. Therefore we opted for designing the binary system described in this project. We are aware that having two PB transposons introduces additional conditions to our system. Ideally we would need to determine the transfection parameters that would result in integration of a minimal amount of both transposons, ideally only one copy of each. This would require careful titration of the plasmid amounts initially transfected into the cells to be reprogrammed. Integration of high numbers of transposons would be undesirable, as when PB transposes out of the genome it re-integrates into the genome with a frequency of 50%. Therefore, high copy numbers would result in no clones surviving the GCV selection. Furthermore, tightness of regulation of the transposase by DOX and TM is also a concern, as leaky expression

of transposase might result in ongoing transposition events during reprogramming that would be mutagenic.

This thesis describes progress towards developing this system. Initially, the construction of the two PB system was one step away from completion. The first plasmid (called Remo) had been constructed and consisted of a reprograming cassette consisting of OCT4, SOX2, KLF4, c-Myc and GFP driven by a CAG promoter and a PuTK cassette driven by a CMV promoter. The second plasmid was partially built and required a last cloning step consisting in the insertion of a TRE-CMV-mPB-ERT2 cassette into a PB backbone already containing a polycistronic construct consisting of a tTA transactivator linked to a PuTK cassette driven by a PGK promoter. This cloning step was the first task of this thesis.

We proceeded to test the plasmid. In order to do so, we successfully demonstrated that it behaved as expected in terms of DOX regulation of the expression of mPB-ERT2. We also demonstrated that by delivering the plasmid into 293T HEK cells and manipulating the levels of DOX and TM in the medium, we could obtain puromycin resistant lines. The number of puromycin resistant colonies obtained was significantly higher when DOX as absent, suggesting that the colonies resulted from transposition events. Presence of TM added an extra layer of regulation, albeit weaker. Our PCR analysis, while not a clean as would be desired, suggested that transposition was indeed occurring, although a background level of random integration could not be ruled out. Finally, our attempt to determine whether we could use GVC to select clones that had successfully mobilized PB out of the genome was unsuccessful. Unexpectedly, 293T HEK cells that had been transfected with ENEAS and selected for puromycin resistance were insensitive to GCV. We see two possible explanations for this. One possibility is that the level of expression of the fusion protein PuTK (driven from a PKG promoter) is high enough to result in resistance to puromycin, but too low to result in sensitivity to GCV. If so, PGK would have to be replaced by a stronger promoter. The second possibility is that the high metabolism of 293T HEK cells renders them insensitive to GCV, or that they are insensitive to GCV for other unknown reasons, although we have found no report of this in the literature. In this sense, 293T HEK cells were a poor choice of cell to work with. The hypothesis of high metabolism could be tested by culturing the 293T HEK cells in a less rich medium to slow down their metabolism, but is probably not worthwhile, as cell to be reprogrammed would be

primary cells such as fibroblasts or keratinocytes. In sum, this thesis has contributed a stage in the development of this system.

Future directions:

The next step would be to transfect both Remo and Eneas into mouse fibroblasts. The total amount of Remo and Eneas would have to be titrated and tested. Too much plasmid might result in high levels of transposition and a high number of integrated transposon copies. This in turn would result in low efficiency of mobilization of the transposon out of the genome. On the other hand, low levels of plasmids would result in low reprogramming efficiencies. Another parameter which needs to be investigated would be the relative ratio of Remo to Eneas. The objective would be to find the ratio of Remo and Eneas that provides the minimum amount of Remo required for reprogramming and the minimum amount of Eneas that provides regulable expression of mPB-ERT2. iPSc lines would need to be derived, expanded, subjected to a second transposon mobilization pulse and selected with GCV. Finally, the resulting iPSc lines would need to be characterized in terms of pluripotency, differentiation, karyotipic integrity and most importantly, absence of plasmid sequences and mutagenic events.

.

.

CHAPTER 5: REFERENCES

5. REFERENCES

Cahan, P. and G. Daley (2013). "Origins and implications of pluripotent stem cell variability and heterogeneity." Nature reviews. Molecular cell biology **14**(6): 357-368.

Cherry, A. and G. Daley (2012). "Reprogramming cellular identity for regenerative medicine." Cell **148**(6): 1110-1122.

Davis, R., H. Weintraub, et al. (1987). "Expression of a single transfected cDNA converts fibroblasts to myoblasts." Cell **51**(6): 987-1000.

Dimos, J. T., K. T. Rodolfa, et al. (2008). "Induced Pluripotent Stem Cells Generated from Patients with ALS Can Be Differentiated into Motor Neurons." Science **321**.

Ebert, A., J. Yu, et al. (2009). "Induced pluripotent stem cells from a spinal muscular atrophy patient." Nature **457**(7227): 277-280.

Egli, D., G. Birkhoff, et al. (2008). "Mediators of reprogramming: transcription factors and transitions through mitosis." Nature reviews. Molecular cell biology **9**(7): 505-516.

El-Karim, E., E. Hagos, et al. (2013). "Krüppel-like factor 4 regulates genetic stability in mouse embryonic fibroblasts." Molecular cancer **12**: 89.

Farfel-Becker, T., E. Vitner, et al. (2011). "Animal models for Gaucher disease research." Disease models & mechanisms **4**(6): 746-752.

González, F., S. Boué, et al. (2011). "Methods for making induced pluripotent stem cells: reprogramming à la carte." Nature reviews. Genetics **12**(4): 231-242.

Gonzalez, F., S. Boue, et al. (2011). "Methods for making induced pluripotent stem cells: reprogramming a la carte." Nat Rev Genet **12**(4): 231-242.

Gurdon, J. (1962). "The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles." Journal of embryology and experimental morphology **10**: 622-640.

Hanna, J. H., K. Saha, et al. "Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues." Cell **143**(4): 508-525.

Huang, H. P., P. H. Chen, et al. (2011). "Human Pompe disease-induced pluripotent stem cells for pathogenesis modeling, drug testing and disease marker identification." Human Molecular Genetics **20**.

Itzhaki, I., L. Maizels, et al. (2011). "Modelling the long QT syndrome with induced pluripotent stem cells." Nature **471**(7337): 225-229.

Jaenisch, R. and R. Young (2008). "Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming." Cell **132**(4): 567-582.

Kim, A. and I. Pyykko (2011). "Size matters: versatile use of PiggyBac transposons as a genetic manipulation tool." Molecular and cellular biochemistry **354**(1-2): 301-309.

Marchetto, M., C. Carromeu, et al. (2010). "A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells." Cell **143**(4): 527-539.

Marson, A., S. S. Levine, et al. (2008). "Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells." Cell **134**(3): 521-533.

Rais, Y., A. Zviran, et al. (2013). "Deterministic direct reprogramming of somatic cells to pluripotency." Nature.

Ramalho-Santos, M., S. Yoon, et al. (2002). ""Stemness": transcriptional profiling of embryonic and adult stem cells." Science (New York, N.Y.) **298**(5593): 597-600.

Robinton, D. and G. Daley (2012). "The promise of induced pluripotent stem cells in research and therapy." Nature **481**(7381): 295-305.

Takahashi, K. (2012). "Cellular reprogramming--lowering gravity on Waddington's epigenetic landscape." Journal of cell science **125**(Pt 11): 2553-2560.

Takahashi, K., K. Tanabe, et al. (2007). "Induction of pluripotent stem cells from adult human fibroblasts by defined factors." Cell **131**(5): 861-872.

Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." Cell **126**(4): 663-676.

Takahashi, K. and S. Yamanaka (2013). "Induced pluripotent stem cells in medicine and biology." Development (Cambridge, England) **140**(12): 2457-2461.

Tiscornia, G., E. Vivas, et al. (2011). "Diseases in a dish: modeling human genetic disorders using induced pluripotent cells." Nature medicine **17**(12): 1570-1576.

Turksen, K. (2013). "End of inevitability: programming and reprogramming." Stem cell reviews **9**(4): 385-387.

van der Worp, H., D. Howells, et al. (2010). "Can animal models of disease reliably inform human studies?" PLoS medicine **7**(3).

Vierbuchen, T. and M. Wernig (2012). "Molecular roadblocks for cellular reprogramming." Molecular cell **47**(6): 827-838.

Warmflash, A., B. Arduini, et al. (2012). "The molecular circuitry underlying pluripotency in embryonic stem cells." Wiley interdisciplinary reviews. Systems biology and medicine **4**(5): 443-456.

Wu, S. and K. Hochedlinger (2011). "Harnessing the potential of induced pluripotent stem cells for regenerative medicine." Nature cell biology **13**(5): 497-505.

Yamanaka, S. (2012). "Induced pluripotent stem cells: past, present, and future." Cell stem cell **10**(6): 678-684.

Yamanaka, S. and H. M. Blau "Nuclear reprogramming to a pluripotent state by three approaches." Nature **465**(7299): 704-712..