Journal of Genetics (2021)100:27
https://doi.org/10.1007/s12041-021-01278-2

© Indian Academy of Sciences

REVIEW ARTICLE

Transcription factor stoichiometry in cell fate determination

RAHUL KUMAR[1,2]* and AJAY KUMAR SHARMA[1]

[1]*University Department of Zoology, Vinoba Bhave University, Hazaribag 825 301, India*
[2]*Department of Zoology, Sheodeni Sao College, Kaler, Arwal 824 127, India*
*For correspondence. E-mail: rahuldayanand33@gmail.com.

Received 25 August 2020; revised 25 November 2020; accepted 9 December 2020

Abstract. Transcription factors play very important role in cell fate determination. There are many cell specific transcription factors which when expressed ectopically may lead to cell fate conversion or transdifferentiation. Many of these transcription factors function differently based on their levels and stoichiometry. Many different types of differentiated cells have been generated from other differentiated cell types by expressing different levels and stoichiometry of reprogramming factors. Many methodologies have been developed for efficient cell fate conversion by regulating the levels and stoichiometry of transcription factors in a particular cocktail that have therapeutic values. An approach called phenotypic activation which involves overexpression of putative transcription factors has been developed as a tool to discover new transcription factors and their targets. Transcription factor overexpression may also have toxic effects where non-specific electrostatic interactions and 'squelching' may lead to inhibition of many genes. Altered levels of transcription factors may have disastrous consequences like cancer. Recent developments like designing of artificial transcription factors, nanotechnology-based transcriptional tools and CRISPR-based transcription modules with capabilities of precise regulation of gene expression patterns hold huge potentials in the field of transcriptional therapeutics.

Keywords. artificial transcription factors; CRISPR; phenotypic activation; reprogramming factors; transdifferentiation.

> '*In the expression of the nucleic message, as well as in its reproduction, adaptation results from an elective rather than an instructive effect of the environment*'
>
> (Francois Jacob, Genetics of the Bacterial Cell, Nobel Lecture, 11 Dec 1965)

Introduction

Environmental changes may influence the genomic functions. Diversity among cells result both due to structural as well as functional changes on the level of genes. Functional changes may arise due to structural changes which involve both changes in the nucleotide sequence as well as changes like methylation of bases and histones, acetylation of histones and other modifications that alter genomic spatial organization. All these changes may be involved in the activation as well as the repression of genes for differential functioning of the cells under different conditions. On cellular level, the extracellular milieu and the cytoplasm have profound impact on the genomic functions. These factors may induce such changes in the genome which can even be inherited. All these changes are called epigenetic changes and these involve mainly covalent modifications of DNA bases (e.g., CpG methylation) and covalent modifications of histones. All these modifications constitute what we call the 'epigenetic signatures of the genome'. In the creation of these signatures, various proteins and noncoding RNAs are involved (Goldberg *et al.* 2007; Clancy *et al.* 2015). These signatures are the 'mirrors' of the active and repressed state of the chromatin which in turn reflect the functional and the silent state of a gene. A multicellular organism is made up of a large numbers of cells, which in turn form tissues. Cells within a particular type of tissue are specialized for specific functions which differ from those present in other tissue types. Such specializations result through the process of differentiation. Life of a multicellular organism starts with a multipotent single celled zygote that undergoes various

Published online: 30 April 2021

changes leading to formation of a multicellular embryo with large number of specialized cells of different form and function. Each type of differentiated cell has its own epigenetic signature. Most of these epigenetic signatures are maintained throughout the life of an organism.

Transcription factors are involved in regulating genetic functions at a larger scale because these factors have direct access to the genome. Transcription factors are key components of environmental responses (Gonzalez 2016). Transcription factors play an important role in the process of differentiation. In fact, many transcription factors are specific to a specific cell type. Transcription factors play an important role in determining cell fate during development. They are involved in generation and maintenance of cell type specific features including their epigenetic signatures. Even, the level of a transcription factor is very important for normal development and appropriate cell fate determination. In this review, it has been explained how different levels of the same transcription factor may lead to different cell fates with acquisition of different epigenetic signatures followed by an overview of different methods developed to alter transcription factor stoichiometry artificially. In chemistry, stoichiometry refers to the quantitative relationship between the amount of all the reactants and the products in a reaction. Likewise, transcription factors do not work in isolation. The level of a particular transcription factor in a cell at a particular time is always relative to other transcription factors and other molecules. Also many transcription factors work in cascade, i.e. activation of one transcription factor may trigger direct activation or repression of other transcription factors. One transcription factor may also influence expression of other transcription factors. Therefore, the term 'stoichiometry' simply refers to the amount of a transcription factor (which is always relative to other transcription factors) in a cell at a particular time (Darnell 2020). The fluctuation in levels of one transcription factor may bring out changes in cell fate by affecting levels of other transcription factors apart from affecting synthesis of other effector molecules and nontranscription factor proteins. Not all transcription factors exhibit similar effects. Some transcription factors may act as master regulators and may have more profound effects on the cellular functions than others. For example, microphthalmia associated transcription factor (MITF) is considered the master regulator of melanocyte development (Levy, Khaled and Fisher 2006; Lin and Fisher 2007) and (sex determining region y)-box9 (SOX9) transcription factor is considered the master regulator of chondrogenesis (Lefebvre and Dvir-Ginzberg 2017). Transcription factor stoichiometry of a particular cell type plays an important role in keeping the cell in its differentiated state. Disturbances causing changes in this stoichiometry may result in cell fate switching and even cancer. Present review article deals with the importance of the relative levels or stoichiometry of transcription factors in determination and maintenance of cell fate, how these levels can be altered artificially to cause specific set of cells to acquire different cell fate more efficiently, how scientists have created specific type of stem cells just by altering the stoichiometry of a set of transcription factors in a cocktail, how overexpression of specific transcription factors have led to the discovery of many other transcription factors in yeast, how overexpression of certain transcription factors may even lead to toxicity in some biological systems and other recent advances in the field of transcription factor biology.

Role of transcription factors in cell fate determination: an experimental overview

Recently, scientists have shown that even terminally differentiated somatic cells can be reprogrammed back again to acquire its pluripotent properties. The first evidence for such reprogramming came from the works of Briggs and King (1952) and Gurdon *et al.* (1958) which proves that all somatic cells of the body of an organism contains same nucleus with same number of chromosomes with same genes. Hence, all the somatic cells of the body of a multicellular organism, though different in structure and function, are genomically equivalent (i.e., they show genomic equivalence). The scientific world was surprised when Ian Wilmut (1997) declared the generation of a cloned sheep, Dolly by transplanting a somatic nucleus in an enucleated egg. These experiments were based on direct transfer of nucleus where factors present in the egg cytoplasm reprogramme the differentiated nucleus to acquire new epigenetic signatures. In a quest to discover factors present in egg cytoplasm causing such effect on differentiated nucleus, it was found that most of these reprogramming factors are transcription factors. Ectopic expression or overexpression of tissue specific transcription factors are also known to induce cell fate switching or transdifferentiation where one type of differentiated cell can become differentiated cell of another type. The earliest evidence of cell fate conversion has come from the seminal work of Weintraub *et al.* (1989), where the ectopic expression of muscle specific basic helix-loop-helix (bHLH) transcription factor named myoD in mouse embryonic fibroblasts converted these cells to myoblasts. Many similar experiments were conducted subsequently by different groups across the globe (Cozar-Castellano and Stewart 2005; Graf 2011; Morris and Daley 2013) (figure 1).

Recently the complete conversion of mouse and human fibroblasts to functional melanocytes by overexpressing only three transcription factors, namely MITF, (sex determining region y)-box 10 (SOX10) and paired box gene 3 (PAX3) has been reported (Yang *et al.* 2014). One more advancement in this area came from the landmark experiments of Takahashi and Yamanaka (2006) for which Yamnaka was awarded Noble prize of Physiology or Medicine in 2012 with Gurdon. In their work, they generated induced pluripotent stem cells (iPSCs) by forced expression of only four transcription factors, namely octamer-binding transcription factor 4 (Oct4), also known as POU domain, class

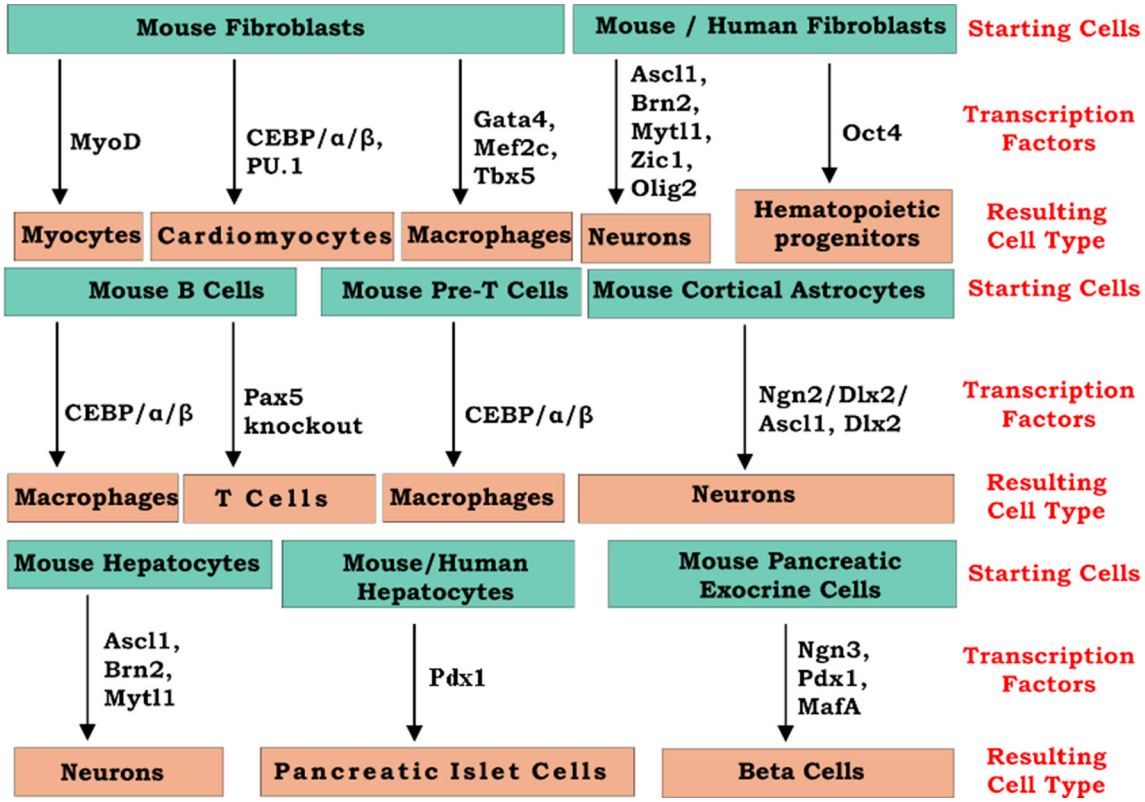

Figure 1. Some examples of transcription factor-induced transdifferentiation.

5, transcription factor 1 (POU5F1), (sex determining region y)-box 2 (Sox2), kruppel-like factor 4 (Klf4) and gene name derived from avian virus, myelocytomatosis (Myc). These transcription factors can induce such molecular changes in a somatic cell that leads to change in the epigenetic signature of the genome from that of a fibroblast to that of a stem cell.

These factors not only regulate transcription via acting on the target genes but also interact with a large number of proteins involved in activation and repression of genes, chromatin remodelling, and generation, maintenance and removal of epigenetic signatures like SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4 (SMARCA4), histone deacetylase 1 (HDAC1), histone deacetylase 2 (HDAC2), histone deacetylase 5 (HDAC5), histone deacetylase 7 (HDAC7), etc. These proteins, e.g., lead to gain of DNA trimethylation gradually on histone H3 Lysine residue 4, 27 and 36 (H3K4me3, H3K27me3 and H3K36me3). The cells produced via these manipulations are very much similar to embryonic stem cells (Schmidt and Plath 2012) but the efficiency of generation of iPSCs using this cocktail of transcription factors is very low (1%). Therefore, many modified strategies have been developed where addition of additional factors with above mentioned core factors like Glis1, Sall4 and Nanog has increased the efficiency (Schmidt and Plath 2012). Even for transdifferentiation experiments, 'shortcut' methodologies have been developed using these core factors where transient expression of these factors prior to expression of the differentiation inducing transcription factors of our choice enhances the efficiency of generation of specific differentiated cell like cardiomyocytes from other differentiated cells like mouse tail tip fibroblasts (Morris and Daley 2013). It should be noted that all differentiated cells are not equally plastic to fate-switching rather some cells like fibroblasts show more differentiation plasticity than cells like neurons (Sanchez Alvarado and Yamanaka 2014). Thus, same methodologies are not equally efficient for all types of cells. These experiments do not tell us about the effects of different levels of transcription factors on cell fate determination, transdifferentiation and induction of pluripotency, and the ratio or stoichiometry of these factors in the cocktail used in the experiments. All these issues will be dealt in subsequent sections.

Transcription factor stoichiometry in cell fate switching

Above discussions were focussed on the role of transcription factors in cell fate determination. In this section, we will explore the role of different levels of transcription factors in cell fate determination. MITF, the master regulator of melanocytes, and also a major melanoma oncogene is found to be amplified in 30–40% of melanomas and is necessary for the survival and proliferation of melanomas. Thus, it is considered as a lineage survival or lineage addiction

oncogene (Steingrímsson *et al.* 2004; Garraway *et al.* 2005; Garraway and Lander 2013). Different levels of MITF shows different effects in melanomas. Low levels of MITF arrest cells in G1 stage of cell cycle and cause stem cell like properties and invasiveness (Carriera *et al.* 2006), intermediate levels increases proliferation, and higher levels again leads to G1 arrest but also differentiation. Therefore, MITF acts as a molecular rheostat (Strub *et al.* 2011; Carriera *et al.* 2006; Hoek and Golding 2010; Cheli *et al.* 2011; Ploper *et al.* 2015). Also, overexpression of MITF with two other transcription factors, Sox10 and Pax3, lead to the conversion of mouse and human fibroblasts to functional melanocytes (Yang *et al.* 2014). The fact that how the same transcription factor like MITF, just because of its different levels at different time points, has such diverse effects is intriguing and experimentally altering the expression levels of such transcription factors may lead to fascinating outcomes.

Two different types of stem cells are known that are derived from early mouse embryos of two different developmental stages: one called as embryonic stem cells (ESCs) from embryos at 3.5 days (E3.5) and another called as epiblast stem cells (EpiSCs) from embryos at 5.5 days (E5.5). These cells can also be produced by low level expression of reprogramming factors Oct4, Sox2, Klf4 and Myc in differentiated cells like fibroblasts. Recently a distinct stem cell was discovered by overexpressing the same reprogramming factors and maintaining their high expression levels in the differentiated cells. These new type of cells have been named as F-class cells, named after the fuzzy appearance of cell colonies in culture (Tonge *et al.* 2014) (figure 2). This finding resulted from an international collaborative work as a part of the project, Project Grandiose (Benevento *et al.* 2014; Clancy *et al.* 2014; Hussein *et al.* 2014; Lee *et al.* 2014; Wu and Izpisua Belmonte 2014; Tonge *et al.* 2014). This finding shows how different levels of the same transcription factor combination can have different effects on the cell fate (figure 2).

Although these cells express many genes that are also expressed by other pluripotent cells like genes involved in mesenchymal-epithelial transition (MET), e.g., Nanog and Sall4. It has been found that high levels of reprogramming factors in these cells lead to activation of many genes that are known to be inactive in ESCs or iPSCs of Yamanaka and many genes that are known to be active in ESCs or iPSCs of Yamanaka, has been found to be inactivated or repressed in this new stem cell state. For example, a number of pluripotency related cell adhesion genes, e.g. cadherin-1 (Cdh1) and epithelial cell adhesion molecule (Epcam); ESC-like markers, e.g. intercellular adhesion molecule 1 (Icam1), nuclear receptor 5A (Nr5a), v-myc avian myelocytomatosis viral oncogene homolog 1, lung carcinoma derived (Mycl1) and other genes, e.g. cluster of differentiation 44 (CD44), that are known to be active in ESCs and iPSCs, have been found to be repressed in F-class cells. These cells also showed high expression of many developmental genes like glioma-associated oncogene family zinc finger 1 (Gli1), islet-1 (Isl1) and *v-kit* hardy-zuckerman 4 feline sarcoma viral oncogene homolog (Kit) (Hussein *et al.* 2014).

There are also significant changes in the epigenetic signatures. Here, high expression levels of reprogramming factors lead to widespread loss of trimethylation of histone H3 lysine residue 27 (H3K27me3). This represents opening of the chromatin, i.e. active gene state and this is known to occur in ESCs but maintenance of high levels of these factors leads to reacquisition of these epigenetic signatures that is stable and is different from ESC-like state (Hussein *et al.* 2014; Lee *et al.* 2014). T-box transcription factor 5 (Tbx5), GATA binding protein 4 (Gata4) and myocyte-specific enhancer factor 2c (Mef2c), the combination of these three transcription factors have been reported to constitute the minimum requirement for directly

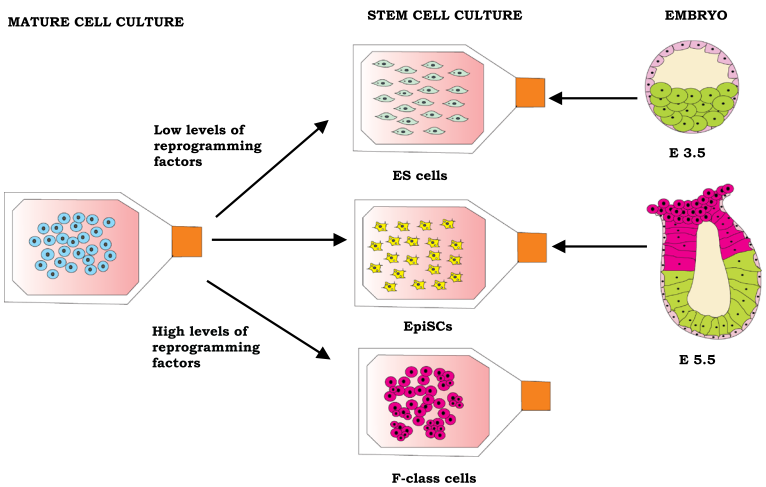

Figure 2. Distinct types of stem cells. ES cells, embryonic stem cells; EpiS cells, epiblast stem cells; F-class cells generated by reprogramming mature differentiated cells using different levels of the same reprogramming factors (explanation is given elsewhere in the review).

reprogramming mouse cardiac fibroblasts to cardiomyocytes (Ieda *et al.* 2010). For this reprogramming experiment, various methods were developed for increasing the reprogramming efficiency but most methods proved inefficient in generating stable induced cardiomyocytes. Later, it was found that optimal balance of expression of these three factors greatly improved the reprogramming efficiency. Thus the problem of inefficient reprogramming arising with most methods was due to use of heterogenous mixtures of the transcription factors without any concern of the ratio of each factor in the mixture. In most of these cases, cells were transfected with mixture of viruses containing insert of individual transcription factors. This led to variable and less efficient reprogramming. To counter this problem, a polycistronic construct was created for the expression of reprogramming factors in a homogenous ratio for improving the reprogramming efficiency. In this study, six polycistronic constructs were generated with identical self-cleaving 2A peptide (P2A and T2A) (Radcliffe and Mitrophanous 2004; Szymazak *et al.* 2004, Kim *et al.* 2011) to incorporate all possible orders of Tbx5, Gata4 and Mef2c in a single transgene, i.e. Mef2c-Gata4-Tbx5 (MGT), Mef2c-Tbx5-Gata4 (MTG), Gata4-Mef2c-Tbx5 (GMT), Gata4-Tbx5-Mef2c (GTM), Tbx5-Mef2c-Gata4 (TMG), and Tbx5-Gata4-Mef2c (TGM) (figure 3) (Wang *et al.* 2015).

When polycistronic vectors were used for transduction, a significant positive difference was observed in the reprogramming efficiency. Particularly, the combination of two vectors, Mef2c-Gata4-Tbx5 and Mef2c-Tbx5-Gata4, with high expression levels of Mef2c and low expression levels of Gata4 and Tbx5 led to enhancement in reprogramming. Therefore, levels of different transcription factors with respect to each other in a cocktail, i.e., their stoichiometry has profound influence on the reprogramming efficiency (Muraoka and Ieda 2015), thus, on the cell fate determination as well.

Phenotypic activation: a tool to discover new transcription factors

As already discussed, different levels of the same transcription factor can activate different set of genes. Expression of many gene are based on factors like temperature, nutrient availability, cell density, growth factors, toxins, etc. There are many unknown genes of putative transcription factors that are functional only at some specific time point under specific known or unknown conditions, that too for a very small time period. Thus, the study of targets and functions of such transcription factors become difficult. To overcome this issue, a new approach has been developed where the overexpression of the transcription factor under study lead to activation of its target genes (both known and unknown), that often (not always) gets reflected by the change in phenotype of the cells. This approach, called the phenotypic activation, led to discovery and characterization of many transcription factors and their targets in yeast (Chua *et al.* 2006), e.g. Toe1, Toe2 and Toe3 transcription factors were discovered and characterized in fission yeast *Schizosaccharomyces pombe* following the same approach of phenotypic activation (Vachon *et al.* 2013). Here, Toe1 overexpression produces elongated cells in fission yeast *Schizosaccharomyces pombe* (figure 4). Thus, experimental

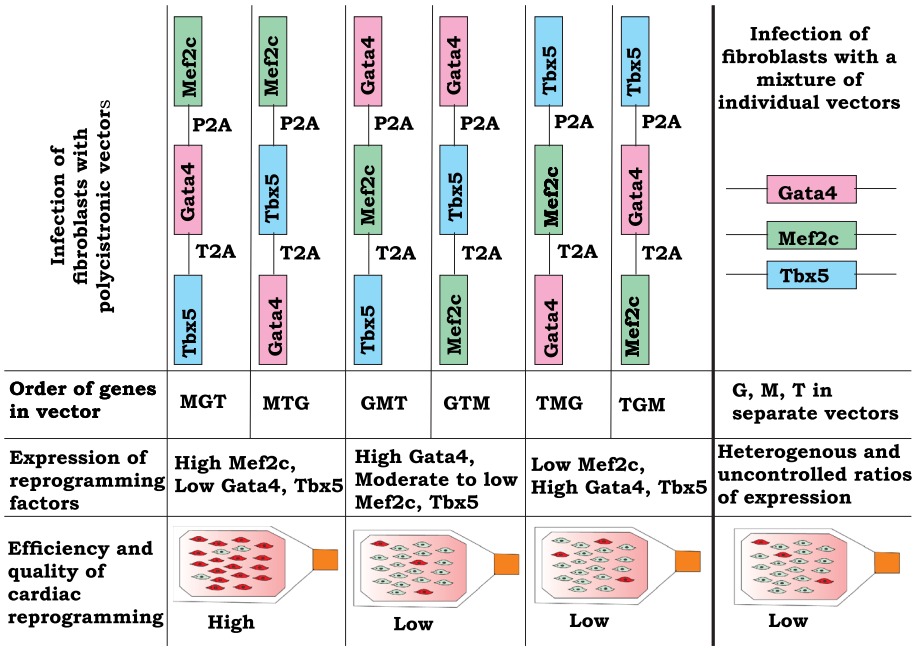

Figure 3. Strategies for direct cardiac reprogramming using polycistronic vectors and individual vectors.

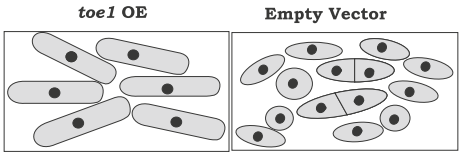

Figure 4. Toe1 transcription factor overexpression in fission yeast *Schizosaccharomyces pombe* produces elongated cells; OE, overexpression.

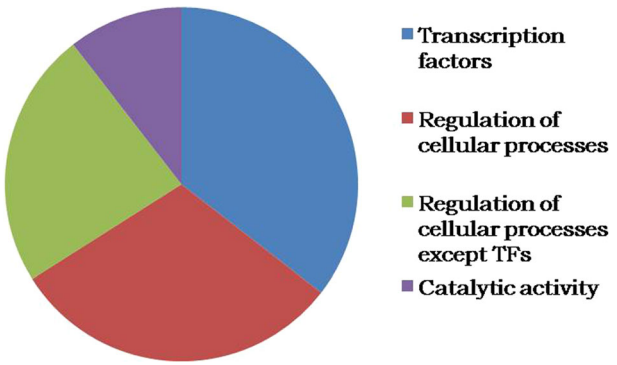

Figure 5. Classes of proteins involved in protein overexpression toxicity in *E. coli*. Proteins acting as transcription factors constitute the major class of proteins involved in protein overexpression toxicity.

alterations in the level of transcription factors provides a way to discover new genes and their analysis.

Transcription factor toxicity due to disturbed stoichiometry

It has been found that overexpression of a transcription factor or transcription activator that is known to be involved in promoting transcription starts working in a reverse way when overexpressed. This is not always due to direct repression of target genes activated by low levels of transcription factors by the activity of other target genes that are activated by high levels of transcription factors as already discussed, rather some other mechanisms are involved. The stoichiometry of proteins is highly maintained even in simple unicellular organisms like bacteria for normal functioning. Many protein related studies involve overexpression of proteins in *Escherichia coli*. It is also known that overexpression of majority of proteins in *E. coli* leads to growth inhibition. Among all these proteins, transcription factors form the major class (figure 5). Non-specific electrostatic interactions have been found to be involved behind toxicity caused by protein overexpression (Singh and Dash 2013).

One more mechanism has been described through which such toxicity is caused by high levels of transcription factors where overexpression of a strong activator that can activate its target genes only by forming complex with a general transcription factor of RNA polymerase II may inhibit the expression of other genes that require the same general transcription factor but some other activator that form a weaker complex with the same general transcription factor. This mechanism has been called 'squelching' or in more precise term- the 'transcriptional squelching' (figure 6). The most common example of transcriptional squelching is the inactivation of many genes in yeast by overexpression of Gal4 protein that contains a strong activation domain (apart from a DNA binding and a dimerization domain) that can bind with many coactivators and general transcription factors of RNA polymerase II. In other words, high levels of Gal4 protein titrates the level of general transcription factors by forming strong complexes thereby sequestering them away from the access of other activator proteins thereby inhibiting (or competitively inhibiting) gene expression (Gill and Ptashne 1988; Ptashne 1988; Latchman 2008; Vachon *et al.* 2013).

Disturbed stoichiometry of many transcription factors may lead to many diseases in human beings. For example, overexpression of transcription factors like signal transducer and activator of transcription 3 (STAT3), signal transducer and activator of transcription 5 (STAT5), nuclear factor kappa B (subunits- c-REL, RELA, NFKB1, NFKB2 and BCL2), β-catenin, Notch, v-jun avian sarcoma virus 17 oncogene homolog (c-JUN), glioma associated oncogene (GLI), etc. are associated with many forms of cancer (Darnell 2002). Nuclear factor kappa B (NF-κB) overexpression is also known to be associated with most chronic inflammatory diseases (Barnes and Karin 1997).

Artificial transcription factors in regulation of transcription factor stoichiometry

Various artificial transcription factors (ATFs), especially the ones containing zinc-fingers have been used to generate phenotypic variations in yeast and mammalian cells (Bae *et al.* 2003; Park *et al.* 2003). For example, zinc finger protein 226 (ZFP226), which is a Cys2-His2 zinc finger based artificial trascription factor has been demonstrated to selectively activate tumour suppressor kidney and brain expressed protein (KIRBA), which is a key regulator of the hippo pathway (Schelleckes *et al.* 2018). Similarly, there is another ATF which can selectively regulate the pro-apoptotic bax gene (Falke *et al.* 2003). Potentials of nanomaterials in regulating gene expression are also being explored (Chun *et al.* 2018). Recently, a nanotechnology based artificial transcription factor named NanoScript has been developed with similar potential. It was constructed by tethering synthetic transcription factors with functional peptide domains which mimic the individual TF domains, onto gold nanoparticles to generate a tunable gene regulatory platform with high specificity and nuclear localization potential (Patel *et al.* 2014). N-(4-chloro-3-(trifluoromethyl) phenyl)-2-ethoxybenzamide (CTB) is known to stimulate pathway leading to activation of Sox9, which is considered the master

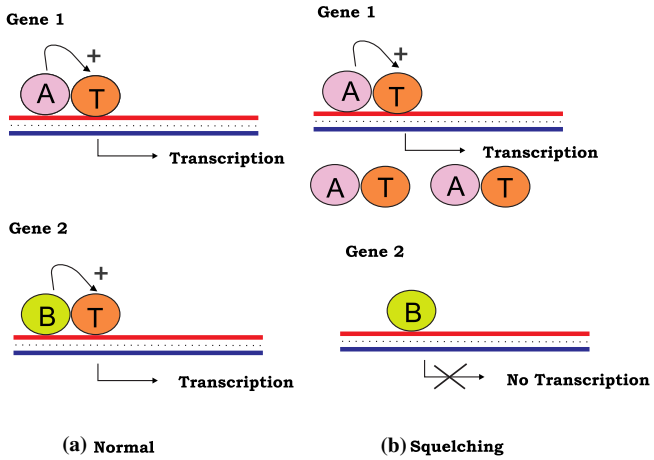

Figure 6. Transcriptional squelching. (a) Two activator molecules A and B activates transcription of genes 1 and 2 respectively only when complexed with a general transcription factor T. (b) If the concentration of activator molecule A is very high, it leads to formation of large number of complexes with general transcription factor T, making it unavailable to B so that only gene 1 is transcribed and gene 2 remains inactive. This phenomenon is called transcriptional squelching.

regulator of chondrogenesis. Modification of Sox9-specific NanoScript with CTB has been reported to promote chondrogenesis in stem cells (Patel *et al.* 2015). There are many such examples of ATFs which have been designed to selectively activate or suppress gene expression. ATFs can prove to be a robust tool to dissect transcriptional networks, thereby opening a new approach to study functional genomics. These may also act as a new class of therapeutics as many diseases including cancer are associated with fault in the gene expression patterns (Ansari and Mapp 2002). ATFs have potential to act as regulatory molecules with huge number of medical applications. The potential abilities of these factors in regulating cell state dynamics and reprogramming of cell fate cannot be ruled out (Heiderscheit *et al.* 2018).

CRISPR mediated alteration of gene expression and transcription factor stoichiometry

Repurposing clustered regularly interspaced short palindromic repeat (CRISPR) proteins for regulating transcription factor stoichiometry proves to be a robust tool with huge therapeutic potentials in diseases occurring due to transcriptional anomalies. These tools are simple to use and highly scalable. This technology has already been used for precise gene editing and epigenome modifications (Hilton *et al.* 2015). CRISPR proteins function by recruiting a modified *Cas* (CRISPR-associated) protein, which is a sequence specific endonuclease, to target sequences using a short stretch of RNA called guide RNA (gRNA). Cas9 is the most widely used Cas with CRISPR (Baliou *et al.* 2018). CRISPR–Cas9 system has been used to target genes associated with binding to specific transcription factors or the genes directly involved in coding transcription factors to produce alterations in the gene expression patterns by affecting transcription factor stoichiometry (Sugano and Nishihama 2018). Cell specific knockdown has also been achieved using this technology (Babaei *et al.* 2019). Apart from this approach, CRISPR-Cas9 based synthetic transcription factors have also been developed with engineered catalytically dead Cas proteins with to alter gene expression (Nihongaki *et al.* 2015; Pandelakis *et al.* 2020). In some of these experiments, synthetic promoters have been constructed and employed to alter phenotypes in yeast cells using CRISPR-Cas9 based synthetic transcription factors (Machens *et al.* 2017). Optogenetic platforms like photoactivatable transcription systems based on CRISPR-Cas9 module have also been developed (Nihongaki *et al.* 2015; Sato 2020). For instance, neuronal differentiation in iPSCs has been achieved recently by upregulating neuronal differentiation 1 (NEUROD1) transcription factor expression using similar photoactivatable transcription system (Nihongaki *et al.* 2017). CRISPR-based synthetic transcription factors can be used for the optimization of gene expression as well as for the spatiotemporal control of gene expression. Therefore, CRISPR–Cas system has emerged as a robust tool with huge potential in the field of transcriptional therapeutics (Pandelakis *et al.* 2020).

Conclusion

The stoichiometry of transcription factors is strictly maintained by cells to maximize their chances of growth, development and survival thereby maximizing the overall fitness of the organism. Disturbances in this stoichiometric balance may have disastrous consequences like cancer. As already explained, many transcription factors are specific to a particular cell type and they act as a master regulator of the fate of that particular cell type, i.e. they are essential for differentiation as well as maintenance of the differentiated

state of that cell type. The methodologies developed to control the levels of transcription factors in a cell, as already discussed, can be used for therapeutic purposes like cell therapy and regenerative medicine. Using such methodologies various disease models can be generated with controlled expression levels of transcription factors that would be more authentic and reproducible. Emerging technologies like artificial transcription factors, nanotechnology-based transcriptional tools and CRISPR-based transcription modules with capabilities of precise regulation of gene expression patterns hold huge potentials in the field of transcriptional therapeutics. In the very beginning of this review, there is quoted, a famous statement of Francois Jacob (1965) which states that environment has 'elective' effects on the functions of genome, it cannot be 'instructive'. Environmental signals are sensed by the cells and to cope up with the changing conditions, which may be favourable or unfavourable to the cell, metabolic changes occur which help the cells to keep integrity of its organization intact under different conditions. Genes are the sole reserves of the information which guide all the metabolic pathways whether it is anabolic or catabolic in response to any signal. For instance, environment cannot have 'instructive' effects on metabolism rather it provides signals, which produce changes only in accordance of the available genomic information. So genome (which contains all the genes) is the sole instructor and decision maker. All the transcription factors, activators, structural proteins, enzymes involved in various metabolic pathways; and products and intermediates of these metabolic pathways; all are dependent on the genomic information for their synthesis, degradation and recycling. If we say that a particular protein, e.g., a transcription factor is regulating the function of a particular gene, it means that genome's own product is regulating the functions of genome. As transcription factors have direct access to the genome, change in its stoichiometry directly affects genetic and cellular functions.

References

Ansari A. Z. and Mapp A. K. 2002 Modular design of artificial transcription factors. *Curr. Opin. Chem. Biol.* **6**, 765–772.

Babaei M., Liu Y., Wuerzberger-Davis S. M., McCaslin E. Z., DiRusso C. J., Yeo A. T. *et al*. 2019 CRISPR/Cas9-based editing of a sensitive transcriptional regulatory element to achieve cell type-specific knockdown of the NEMO scaffold protein. *PLoS One* **14**, e0222588.

Bae K. H., Kwon Y. D., Shin H. C., Hwang M. S., Ryu E. H., Park K. S. *et al*. 2003 Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat. Biotechnol.* **21**, 275–280.

Baliou S., Adamaki M., Kyriakopoulos A. M., Spandidos D. A., Panayiotidis M., Christodoulou I. *et al*. 2018 Role of the CRISPR system in controlling gene transcription and monitoring cell fate (Review). *Mol. Med. Rep.* **17**, 1421–1427.

Barnes P. J. and Karin M. 1997 Nuclear factor-kappaB: a pivotal transcription factor in chronic inflammatory diseases. *N. Engl. J. Med.* **336**, 1066–1071.

Benevento M., Tonge P. D., Puri M. C., Hussein S. M., Cloonan N., Wood D. L. *et al*. 2014 Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks. *Nat. Commun.* **5**, 5613.

Briggs R. and King T. J. 1952 Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc. Natl. Acad. Sci. USA* **38**, 455–463.

Carreira S., Goodall J., Denat L., Rodriguez M., Nuciforo P., Hoek K. S. *et al*. 2006 Mitf regulation of Dia1 controls melanoma proliferation and invasiveness. *Genes Dev.* **20**, 3426–3439.

Cheli Y., Giuliano S., Botton T., Rocchi S., Hofman V., Hofman P. *et al*. 2011 Mitf is the key molecular switch between mouse or human melanoma initiating cells and their differentiated progeny. *Oncogene* **30**, 2307–2318.

Chua G., Morris Q. D., Sopko R., Robinson M. D., Ryan O., Chan E. T. *et al*. 2006 Identifying transcription factor functions and targets by phenotypic activation. *Proc. Natl. Acad. Sci. USA* **103**, 12045–12050.

Chun S. H., Yuk J. S. and Um S. H. 2018 Regulation of cellular gene expression by nanomaterials. *Nano Converg.* **5**, 34.

Clancy J. L., Patel H. R., Hussein S. M., Tonge P. D., Cloonan N., Corso A. J. *et al*. 2014 Small RNA changes en route to distinct cellular states of induced pluripotency. *Nat. Commun.* **5**, 5522.

Cozar-Castellano I. and Stewart A. F. 2005 Molecular engineering human hepatocytes into pancreatic beta cells for diabetes therapy. *Proc. Natl. Acad. Sci. USA* **102**, 7781–7782.

Darnell J. E. Jr 2002 Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2**, 740–749.

Darnell R. B. 2020 The genetic control of stoichiometry underlying autism. *Ann. Rev. Neurosci.* **43**, 509–533.

Falke D., Fisher M., Ye D. and Juliano R. L. 2003 Design of artificial transcription factors to selectively regulate the pro-apoptotic bax gene. *Nucleic Acids Res.* **31**, e10.

Garraway L. A. and Lander E. S. 2013 Lessons from the cancer genome. *Cell* **153**, 17–37.

Garraway L. A., Widlund H. R., Rubin M. A., Getz G., Berger A. J., Ramaswamy S. *et al*. 2005 Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122.

Gill G. and Ptashne M. 1988 Negative effect of the transcriptional activator GAL4. *Nature* **334**, 721–724.

Goldberg A. D., Allis C. D. and Bernstein E. 2007 Epigenetics: a landscape takes shape. *Cell* **128**, 635–638.

Gonzalez D. H. 2016 *Plant transcription factors: evolutionary, structural and functional aspects*, pp. 35–72. Academic Press, Cambridge.

Graf T. 2011 Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell.* **9**, 504–516.

Gurdon J. B., Elsdale T. R. and Fischberg M. 1958 Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* **182**, 64–65.

Heiderscheit E. A., Eguchi A., Spurgat M. C. and Ansari A. Z. 2018 Reprogramming cell fate with artificial transcription factors. *FEBS Lett.* **592**, 888–900.

Hilton I. B., D'Ippolito A. M., Vockley C. M., Thakore P. I., Crawford G. E., Reddy T. E. *et al*. 2015 Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517.

Hoek K. S. and Goding C. R. 2010 Cancer stem cells versus phenotype-switching in melanoma. *Pigment Cell Melanoma Res.* **23**, 746–759.

Hussein S. M., Puri M. C., Tonge P. D., Benevento M., Corso A. J., Clancy J. L. *et al*. 2014 Genome-wide characterization of the routes to pluripotency. *Nature* **516**, 198–206.

Ieda M., Fu J. D., Delgado-Olguin P., Vedantham V., Hayashi Y., Bruneau B. G. and Srivastava D. 2010 Direct reprogramming of

fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375–386.

Jacob Francois 1965 Genetics of bacterial cell. *Nobel Lecture*.

Kim J. H., Lee S. R., Li L. H., Park H. J., Park J. H., Lee K. Y. *et al*. 2011 High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**, e18556.

Latchman D. S. 2008 *Eukaryotic transcription factors, chapter 5 activation of gene expression by transcription factors*, pp. 161-228. Academic Press, Cambridge.

Lee D. S., Shin J. Y., Tonge P. D., Puri M. C., Lee S., Park H. *et al*. 2014 An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. *Nat. Commun.* **5**, 5619.

Lefebvre V. and Dvir-Ginzberg M. 2017 SOX9 and the many facets of its regulation in the chondrocyte lineage. *Connect. Tissue Res.* **58**, 2–14.

Levy C., Khaled M. and Fisher D. E. 2006 MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol. Med.* **12**, 406–414.

Lin J. Y. and Fisher D. E. 2007 Melanocyte biology and skin pigmentation. *Nature* **445**, 843–850.

Machens F., Balazadeh S., Mueller-Roeber B. and Messerschmidt K. 2017 Synthetic promoters and transcription factors for heterologous protein expression in *Saccharomyces cerevisiae*. *Front. Bioeng. Biotechnol.* **5**, 63.

Morris S. A. and Daley G. Q. 2013 A blueprint for engineering cell fate: current technologies to reprogram cell identity. *Cell Res.* **23**, 33–48.

Muraoka N. and Ieda M. 2015 Stoichiometry of transcription factors is critical for cardiac reprogramming. *Cir. Res.* **116**, 216–218.

Nihongaki Y., Furuhata Y., Otabe T., Hasegawa S., Yoshimoto K. and Sato M. 2017 CRISPR-Cas9-based photoactivatable transcription systems to induce neuronal differentiation. *Nat. Methods* **14**, 963–966.

Nihongaki Y., Yamamoto S., Kawano F., Suzuki H. and Sato M. 2015 CRISPR-Cas9-based photoactivatable transcription system. *Chem. Biol.* **22**, 169–174.

Pandelakis M., Delgado E. and Ebrahimkhani M. R. 2020 CRISPR-based synthetic transcription factors in vivo: the future of therapeutic cellular programming. *Cell Syst.* **10**, 1–14.

Park K. S., Lee D. K., Lee H., Lee Y., Jang Y. S., Kim Y. H. *et al*. 2003 Phenotypic alteration of eukaryotic cells using randomized libraries of artificial transcription factors. *Nat. Biotechnol.* **21**, 1208–1214.

Patel S., Jung D., Yin P. T., Carlton P., Yamamoto M., Bando T. *et al*. 2014 NanoScript: a nanoparticle-based artificial transcription factor for effective gene regulation. *ACS Nano.* **8**, 8959–8967.

Patel S., Pongkulapa T., Yin P. T., Pandian G. N., Rathnam C., Bando T. *et al*. 2015 Integrating epigenetic modulators into NanoScript for enhanced chondrogenesis of stem cells. *J. Am. Chem. Soc.* **137**, 4598–4601.

Ploper D., Taelman V. F., Robert L., Perez B. S., Titz B., Chen H. W. *et al*. 2015 MITF drives endolysosomal biogenesis and potentiates Wnt signaling in melanoma cells. *Proc. Natl Acad. Sci. USA* **112**, E420–E429.

Ptashne M. 1988 How eukaryotic transcriptional activators work. *Nature* **335**, 683–689.

Radcliffe P. A. and Mitrophanous K. A. 2004 Mutiple gene products from a single vector: 'self-cleaving' 2A peptides. *Gene Therapy* **11**, 1673–1674.

Sánchez Alvarado A. and Yamanaka S. 2014 Rethinking differentiation: stem cells, regeneration, and plasticity. *Cell* **157**, 110–119.

Sato M. 2020 Yakugaku zasshi [Manipulating living systems by light]. *J. Pharm. Soc. Jap.* **140**, 993–1000.

Schelleckes K., Schmitz B., Lenders M., Mewes M., Brand S. and Brand E. 2018 ZFP226 is a novel artificial transcription factor for selective activation of tumor suppressor KIBRA. *Sci. Rep.* **8**, 4230.

Schmidt R. and Plath K. 2012 The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome Biol.* **13**, 251.

Singh G. P. and Dash D. 2013 Electrostatic mis-interactions cause overexpression toxicity of proteins in E. coli. *PLoS One* **8**, e64893.

Steingrímsson E., Copeland N. G. and Jenkins N. A. 2004 Melanocytes and the microphthalmia transcription factor network. *Annu. Rev. Genet.* **38**, 365–411.

Strub T., Giuliano S., Ye T., Bonet C., Keime C., Kobi D. *et al*. 2011 Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. *Oncogene* **30**, 2319–2332.

Sugano S. S. and Nishihama R. 2018 CRISPR/Cas9-based genome editing of transcription factor genes in *Marchantia polymorpha*. *Methods Mol. Biol.* **1830**, 109–126.

Szymczak A. L., Workman C. J., Wang Y., Vignali K. M., Dilioglou S., Vanin E. F. *et al*. 2004 Correction of multi-gene deficiency in vivo using a single "self-cleaving" 2A peptide-based retroviral vector. *Nat. Biotechnol.* **22**, 589–594.

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

Tonge P. D., Corso A. J., Monetti C., Hussein S. M., Puri M. C., Michael I. P. *et al*. 2014 Divergent reprogramming routes lead to alternative stem-cell states. *Nature* **516**, 192–197.

Vachon L., Wood J., Kwon E. J., Laderoute A., Chatfield-Reed K., Karagiannis J. *et al*. 2013 Functional characterization of fission yeast transcription factors by overexpression analysis. *Genetics* **194**, 873–884.

Wang L., Liu Z., Yin C., Asfour H., Chen O., Li Y. *et al*. 2015 Stoichiometry of Gata4, Mef2c, and Tbx5 influences the efficiency and quality of induced cardiac myocyte reprogramming. *Circ. Res.* **116**, 237–244.

Weintraub H., Tapscott S. J., Davis R. L., Thayer M. J., Adam M. A., Lassar A. B. *et al*. 1989 Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci. USA* **86**, 5434–5438.

Wilmut I., Schnieke A. E., McWhir J., Kind A. J. and Campbell K. H. 1997 Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813.

Wu J. and Izpisua Belmonte J. C. 2014 Stem cells: A designer's guide to pluripotency. *Nature* **516**, 172–173.

Yang R., Zheng Y., Li L., Liu S., Burrows M., Wei Z. *et al*. 2014 Direct conversion of mouse and human fibroblasts to functional melanocytes by defined factors. *Nat. Commun.* **5**, 5807.

Corresponding editor: Upendra Nongthomba