The Principle of Self-Organization in Multicellular Organisms—Fuzzy Regulation of Differentiation and Dedifferentiation

Zheming Cao * and Weidong Ding

Posted Date: 29 October 2024

doi: 10.20944/preprints202410.2255.v1

Keywords: gene regulation; embryonic development; cell differentiation; mathematical model

 doi:10.20944/preprints202410.2255.v1

Article

The Principle of Self-Organization in Multicellular Organisms—Fuzzy Regulation of Differentiation and Dedifferentiation

Zheming Cao * and Weidong Ding

Key Laboratory of Freshwater Fisheries and Germplasm Resources Utilization, Ministry of Agriculture, Freshwater Fisheries Research Center, Wuxi, China

* Correspondence: author, Email address: caozm@ffrc.cn

Abstract: The gene sets of the two cells produced by each division of the fertilized egg are different, and these differences gradually accumulate. When the similarity between a certain cell and a certain type of cell is below a threshold, the differentiation ability in that direction is lost. When all cells can only have one developmental direction, the expression of zygotic genes is turned on. When a stem cell differentiates, it first divides into two cells. One of two cells, similar to the next level cell, becomes the next level cell, while the other, similar to the mother cell, continues to become the other stem cell. When a cell undergoes dedifferentiation, it expresses a group of genes. When the gene set is similar to the upper level stem cell, the cell undergoes dedifferentiation and becomes the upper level stem cell.

Keywords: gene regulation; embryonic development; cell differentiation; mathematical model

---

Cell differentiation refers to the process in which cells of the same origin gradually produce cell groups with different morphological structures and functional characteristics, resulting in spatial differences and temporal differences between the same cell and its previous state. The essence of cell differentiation is the selective expression of the genome in time and space, which ultimately produces landmark proteins through the activation or deactivation of different gene expressions. In general, the process of cell differentiation is irreversible. However, under certain conditions, differentiated cells are also unstable, and their gene expression patterns can undergo reversible changes, returning to their undifferentiated state, a process called dedifferentiation. The essence of cell differentiation is the selective expression of genes, and due to the complex mechanism of cell differentiation, it is still a mystery to this day. In the author's previous article, it was mentioned that cell differentiation and dedifferentiation require the cell to undergo an unstable state in order to complete (1,2). However, this hypothesis does not address the direction of differentiation and dedifferentiation, so it cannot be used in reality. To solve this problem, this article explores the regulation of differentiation and dedifferentiation from another perspective.

1. Fuzzy Regulation of Fertilized Egg Development

We start analyzing from the fertilized egg. As the activation of zygotic genes generally does not begin until the late blastocyst stage, most fertilized eggs rely on RNA synthesized by the mother for growth and differentiation. The total number of genes involved is m, and the expression level of each gene is f, which is represented by a total set A. Due to the complexity of involving two variables, we will ignore the expression level in the following discussion and assume that differentiation is only related to the number and type of genes.

 doi:10.20944/preprints202410.2255.v1

At the blastocyst stage, the embryo completes the differentiation of all types of cells, each with a specific direction of differentiation. The genes involved in the production of each type of cell form a set {K1, K2, K3, ⋯, Kn}, Among them, Kn is the gene set involved in each type of cell.

A = ∪{K1, K2, K3, ⋯, Kn}

At the beginning of fertilization, due to the uneven distribution of maternal mRNA in the cytoplasm, each division will result in differences in the distribution of a certain number of genes (n) between the two cells.

The similarity between two cells in the two cell phase is: $x = \frac{m-n}{m}$,

We form two sets of genes expressed by these two cells, X1-1 and X1-2, which intersect with all Kn and the intersection is almost equal to Kn.

X1-1∩Kn≈Kn ; X1-2∩Kn≈Kn

The similarity between two groups of cells in the four cell phase is: $x = \left(\frac{m-n}{m}\right)^2$

We form four sets of genes expressed by these four cells, X2-1, X2-2, X2-3, X2-4, These four sets intersect with all Kn, and the intersection is almost equal to Kn.

X2-1, X2-2, X2-3, X2-4∩Kn≈Kn,

The similarity between the four groups of cells in the eight cell phase is: $x = \left(\frac{m-n}{m}\right)^3$

Similarly, among these eight cells, a certain cell may experience X3-n ∩ Kn ≈ Kn or X3-n ∩ Kn<Kn, the latter indicating that the cell has begun to break free from pluripotency.

The similarity between different groups of cells after L divisions is: $x = \left(\frac{m-n}{m}\right)^L$

Similarly, among these 2L cells, a certain cell may have a similarity with one of Kn that is below a threshold (represented by the Jaccard similarity coefficient D), indicating that the cell has already begun to differentiate into this type of cell.

As the number of cell divisions increases, different cells gradually lose different differentiation possibilities. If a cell can only have a Jaccard similarity coefficient with a certain Kn that exceeds a threshold, and its Jaccard similarity coefficient with all other Kns is below the threshold, it can only develop in this direction, and the zygotic gene may also be activated.

When any Kn finds only cells with a Jaccard similarity coefficient exceeding a threshold and a Jaccard similarity coefficient below the threshold with all other Kns, embryonic development is essentially complete, and it can activate the zygote gene and enter organ development.

How organisms arrange the gene set of fertilized eggs during evolution, as well as the size and capacity of subsets of different types of cells, and the differences produced during each division, determine the initial state and subsequent development of the vast majority of traits.

2. Fuzzy Regulation of Cell Differentiation

We form a set A of all genes involved in a type of stem cell, and different subsets An, An ⊆ A of genes involved in different individual cells in this type of cell are formed

Due to the fact that cell differentiation is inevitably accompanied by cell division, the two cells formed by division form two subsets: An1+An2=An

All the genes involved in his lower level cells form another set B, and the genes involved in different individual cells in this type of cell form different subsets Bn, Bn ⊆ B

To differentiate into the next level of cells, this type of cell must shut down some genes, which turn off a series of new genes to form a set C1; We also need to activate some genes that initiate a series of new genes to form a set C2.

The Jaccard similarity coefficient is a commonly used method for evaluating the similarity between two sets, where the similarity between these two cells and the next level cells (D1, D2) is:

$D1 = \frac{|(An1-C1+C2)\cap Bn|}{|(An1-C1+C2)\cup Bn|}$; $D2 = \frac{|(An2-C1+C2)\cap Bn|}{|(An2-C1+C2)\cup Bn|}$;

If one of D1 and D2 is greater than a certain threshold, the cell differentiation is successful.

The similarity between these two cells and stem cells (D1 ‘, D2’) is

$D1' = \frac{|(An1-C1+C2)\cap An|}{|(An1-C1+C2)\cup An|}$; $D2' = \frac{|(An2-C1+C2)\cap An|}{|(An2-C1+C2)\cup An|}$

 doi:10.20944/preprints202410.2255.v1

If either D1 ‘or D2’ exceeds a certain threshold, the cell returns and continues to become a stem cell.

If neither intersection meets the criteria, this type of cell becomes somewhat similar to a stem cell but not a stem cell, and somewhat similar to the next level cell but not the next level cell. This situation should occur frequently, but in real experiments, we cannot distinguish such cells at all, so there are no relevant reports. There are three final outcomes for this type of cell: 1, continue to adjust and dedifferentiate into stem cells; 2, continue adjusting to enter the next level of cells; 3. If the adjustments are not in place, it will eventually lead to apoptosis.

If all genes involved in the expression of any specific type of cell are a set, and when the upper stem cell differentiates into the lower stem cell, in addition to experiencing an unstable state (1,2), it also needs to go through a process involving many gene regulation modes. There are already many patterns of this process, such as the network regulation mode, which all have a common disadvantage that they do not match the results of real experiments. From single-cell transcriptome analysis or transcriptome analysis of cells of the same type from different individuals, a consensus is that the genes involved in the transcriptome of the same tissue from different individuals and different cells from the same tissue are quite different. This phenomenon can be easily explained by the above theory, because during the differentiation process of stem cells, the gene set of the remaining stem cells has changed from the original stem cells. When these stem cells differentiate again, the differentiated cells will inevitably be different from the previous cells, and this deviation can be called gene set deviation. This also explains how human aging occurs and how the speed of aging is highly correlated with deviations from the gene set. If analyzed using existing precise regulatory models, this phenomenon cannot be explained, and such precise regulation may result in a wide range of mutated cells, making it impossible for normal individuals to appear.

3. Fuzzy Regulation of Cell Dedifferentiation

We form a set A of all genes involved in a certain type of cell, and different subsets An, $An \subseteq A$ of genes involved in different cells in this type of cell are formed.

All the genes involved in his upper level stem cells form another set B, and the genes involved in different cells in this type of stem cells form different subsets Bn, $Bn \subseteq B$

In order for these types of cells to differentiate into higher-level stem cells, they must activate certain genes that initiate a series of new genes to form a set C1. They also need to shut down some genes, which shut down a series of new genes to form a set C2.

The Jaccard similarity coefficient D between the new cell and the previous cell is:

$$D = \frac{|\ (An - C1 + C2)\ \cap Bn\ |}{|\ (An - C1 + C2)\ \cup Bn\ |}$$

If D is greater than a threshold, cell dedifferentiation is successful. If it does not meet the criteria, dedifferentiation is not successful. These types of cells do not know their specific functions and are somewhat similar to higher-level stem cells but not higher-level stem cells.

The widely reported result regarding dedifferentiation is IPS cells. In 2006, a laboratory led by Shinya Yamanaka at Kyoto University in Japan introduced four transcription factors, Oct3/4, Sox2, c-Myc, and Klf4, into mouse embryonic or skin fibroblasts and found that they could induce transformation. The resulting iPS cells were highly similar to embryonic stem cells in terms of morphology, gene and protein expression, epigenetic modification status, cell doubling ability, embryoid and teratogenic ability, differentiation ability, etc (3,4). In the following research, it was found that many combinations of other transcription factors can also cause specific cells to dedifferentiate into a state similar to embryonic stem cells (5,6).

There are two issues during the formation of IPS cells. Firstly, the methods used by researchers to generate induced pluripotent stem cells are inefficient. According to current methods, when introducing four transcription factors into adult cells such as skin cells, only a few iPS cells can be obtained using thousands of skin cells. During the induction process of the second iPS stem cells, a large number of cell clones appear, which resemble stem cells in appearance, growth rate, and other aspects, but do not have the gene expression and function that stem cells should have.

 doi:10.20944/preprints202410.2255.v1

The reason of the phenomenon is very simple , that is, if the Jaccard similarity coefficient D is greater than a threshold, dedifferentiation is successful, but the proportion is very low. If the Jaccard similarity coefficient D is less than a threshold, dedifferentiation is not successful, and a series of cell clones resembling stem cells appear. If the set C changes, the Jaccard similarity coefficient D also changes. How to increase the cloning ratio can be achieved by changing the number and types of transcription factors.

Discussion

The human genome has approximately about 20000 genes, including over 1800 transcription factor gene loci that encode over 3500 isoforms of transcription factors (7). Except for house-keeping genes, every type of cell has an unique expression profile. We don't know exactly how many types of cells humans have, but it should be much lower than the number of regulatory factors. This raises the question of why so many regulatory factors are needed. It should be noted that if different combinations of regulatory factors can also regulate differentiation or dedifferentiation, so the types of regulation that 1800 transcription factors can produce are almost astronomical. Many people have proposed the network regulation of traits, that is said a certain trait is jointly regulated by a large number of genes. The biggest problem with this hypothesis is the development of animal embryos, because animal embryos rely on mRNA provided by the mother before zygotic gene activation from fertilized eggs to zygotes, and there is no regulatory problem.

The same gene should have two variables, whether it is expressed and the expression level. We overlooked the issue of expression level in the article. The expression level undoubtedly has an extremely important impact on differentiation, and these two variables should be given equal weight. This article only ignores the discussion of expression level for the sake of easier understanding. A comprehensive analysis of fuzzy regulation may require the use of AI's analytical capabilities.

During differentiation or dedifferentiation, it is worth discussing whether each gene has the same weight in similarity calculation. From the results of many experiments, we believe that the weight of different genes varies in most tissue differentiation. A small number of tissues differentiate with similar gene weights, which may be the case during embryonic development.

Gene knockout is now widely used to study the function of a gene. In practice, gene knockout is difficult to achieve corresponding results. In most cases, the experimental organism with a gene knockout does not show any mutations, and in some cases, the experimental organism with a gene knockout shows very low frequency mutations. The explanation for this is that organisms have compensatory mechanisms that can counteract the effects of knocking out genes. The author believes that in the process of differentiation and dedifferentiation, most traits are vaguely regulated, and the role of a single gene is actually very limited. Therefore, knocking out a certain gene does not result in mutant traits, and the so-called compensation mechanism does not exist. It is possible that the gene set involved in such traits includes the genes involved in the compensation mechanism. As described in 2, we can see that differentiation is closely related to both An and Bn, and whether the elimination of C1 by gene knockout can produce results depends entirely on An and Bn. When the similarity coefficient D is below a certain threshold, the result of unsuccessful differentiation can occur, and mutants can also appear.

It is a very common phenomenon for the same gene to appear in many tissues. The housekeeping genes are generally present in any intersection, but they only affect cellular instability ( )and do not affect differentiation or dedifferentiation. Many regulatory factors also appear in different tissues, and they generally have two functions: 1. Because cell differentiation and dedifferentiation require a certain level of complexity, the emergence of these genes is to maintain cell complexity; 2. They are also actual participants in differentiation and dedifferentiation, and they can enter characteristic subsets of many types of cells.

Transcriptome analysis is a commonly used method in genetics to compare transcriptional differences between different treatment groups or mutation groups and normal control tissues. The analysis of transcriptome results is often full of contradictions, with common contradictions being that the differences between individuals in the normal control group are even greater than those in

 doi:10.20944/preprints202410.2255.v1

the experimental group; There are too many related genes identified through transcriptome analysis to be selected; The analysis of transcriptome results often leads to completely unrelated conclusions when applying related genes to other individuals. We have already mentioned that this is a phenomenon caused by the deviation of gene sets resulting from differentiation. The deviation of the gene set generated per unit time from different organisms, individuals, organs of the same organism, and even the same tissue of different individuals of the same organism is different. It determines the aging of the animal, the aging lesions of organs, and the occurrence of cancer, and is an important systematic characteristic value.

The arrangement of different gene sets involved in embryonic differentiation or differentiation and dedifferentiation of all tissues in this mathematical model requires precise calculation and experimental verification, and there is basically no literature available for specific information. The gene set of embryonic differentiation in each organism is its characteristic value, which determines what kind of organs and appearance the organism will have. In evolution, organisms have already accurately calculated what kind of gene set the fertilized egg should have, and what kind of gene set different types of cells should have in order to accurately reach the end point of embryonic differentiation in each division. And the gene sets for differentiation and dedifferentiation will not be too complex, because after embryonic differentiation, all types of cells do not have so many genes to choose from, because most of their genome genes have been closed and are no longer in the gene set, so the characteristics of their gene set are not complicated. These can all be calculated using pure mathematical methods, but the settlement results also need to be verified through experiments.

The occurrence of cancer is closely related to differentiation and dedifferentiation. From our above analysis, we can see that the process of cell differentiation will produce some abnormal cells, which may re-enter stem cells through dedifferentiation. However, the gene set of these cells has changed compared to the original stem cells, so the proportion of dedifferentiated cells entering stem cells will not reach 100%. Those cells that fail to dedifferentiate will become a type of cells of neither fish nor fowl, which are likely the source of cancer. During a person's growth process, stem cells continuously differentiate into the next level of cells, while the remaining stem cells slowly change their gene set. The probability of these stem cells differentiating incorrectly increases with age, leading to an increasing number of atypical cells that can significantly increase the risk of cancer in the elderly.

Before the fertilized egg divides into the 8-cell stage, each cell has the ability to develop into a complete embryo, while cells after the 8-cell stage have little ability to develop into a complete embryo. This phenomenon can be easily understood using our previous hypothesis. After the eight cell stage, the gene subsets contained in each cell of embryonic cells may not have sufficient intersection with a certain type of cell after individual development, resulting in the inability to form corresponding organs differentiated from the development of that type of cell.

Some stem cells can appear in the blood of infants, and it is now widely believed that these cells have great differentiation ability and can differentiate into most organs. But we believe that these types of cells are mostly cells of neither fish nor fowl with unclear differentiation directions that appear during embryonic development. The gene set of these cells is incomplete, making differentiation difficult, and they are prone to evolve into another type of cells of neither fish nor fowl during the differentiation process, ultimately leading to cancer. Therefore, their preservation value is not significant.

The last question is, how are the subsets of genes that are added or closed during differentiation and dedifferentiation generated? For any cell, it absorbs negative entropy, increases its volume and the content of organic matter such as proteins, and the content of transcription factors is constantly increasing. When the concentration of transcription factors exceeds a certain threshold, they not only bind to the main binding site, but also to the secondary binding site at a certain frequency, which opens up some new genes or closes some genes. Although these secondary binding sites are also fixed, they have a certain degree of randomness. This randomness also contributes to the diversity of biological cells and the deviation of gene sets.

 doi:10.20944/preprints202410.2255.v1

References

1. Cao Z, Ding W, Qiang J, Wang M, Zhu H. Use of the catastrophe model to preliminarily explore the principles of multicellular organization. Anat Physiol. 2023. 13(1): 1-5.
2. Cao Z, Ding W, Preliminary Examination of the Principle of Self-Organization in Multicellular Organisms: How Traits are formed, Hereditary Genet, Vol.12 Iss. 3 No:1000252
3. Pei XT. Stem Cell biology. Science Press. 2003.
4. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. cell. 2006;126(4):663-76.
5. Yakubov E, Rechavi G, Rozenblatt S, Givol D. Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of fourtranscription factors. Biochemical and biophysical research communications. 2010;394(1):189-93.
6. Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, et al. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell stem cell. 2011;8(6):633-8.
7. Amaral, P., Carbonell-Sala, S., De La Vega, F.M. et al. The status of the human gene catalogue. Nature 622, 41–47 (2023).