



# Formation of Cell-Type-Associated Chromatin Configurations: A Hypothesis

Gao-De Li

Chinese Acupuncture Clinic, Liverpool, UK

Email: gaode\_li@yahoo.co.uk

**How to cite this paper:** Li, G.-D. (2019) Formation of Cell-Type-Associated Chromatin Configurations: A Hypothesis. *Open Access Library Journal*, 6: e5246. <https://doi.org/10.4236/oalib.1105246>

**Received:** February 14, 2019

**Accepted:** February 25, 2019

**Published:** February 28, 2019

Copyright © 2019 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

More than 30 years ago, we published a hypothesis in which the 3-dimensional structure of entire chromatin fibres in a eukaryotic cell is named as chromatin configuration that determines the pattern of gene expression in each type of cell. Based on this hypothesis, cell differentiation is the process in which different types of cells are endowed with different chromatin configurations. Therefore, this chromatin configuration could also be named as cell-type-associated chromatin configuration. However, how this cell-type-associated chromatin configuration is formed during cell differentiation remains unknown. In this paper, we hypothesize that a group of unknown proteins might be involved in the formation of cell-type-associated chromatin configurations during cell differentiation.

## Subject Areas

Cell Biology, Molecular Biology, Genomics

## Keywords

Cell-Type-Associated Chromatin Configuration, 3-Dimensional Genome Architecture, Cell Differentiation, Gene Regulation, Rivet Protein, *Plasmodium falciparum* Chloroquine Resistance Marker Protein (Pfcrrmp)

## 1. Introduction

Thirty-three years ago, we proposed that chromatin configuration is associated with gene activity (expression) and cell differentiation. Based on these assumptions, we hypothesized that abnormal chromatin configuration might be the cause of oncogenesis [1] [2]. In recent years, more and more research findings have strongly supported our hypothesis [3] [4] [5]. We also speculated that cell

differentiation is the process in which different types of cells are endowed with different chromatin configurations. Therefore, this type of chromatin configuration could also be named as cell-type-associated chromatin configuration. Since there are about 200 different types of cells in our human body, it is reasonable to think that human genome has a potential for forming about 200 different cell-type-associated chromatin configurations. Undoubtedly, this is true in embryonic stem cell. But in a somatic or differentiated cell, the same genome can only form one cell-type-associated chromatin configuration, which is the reason why a somatic cell can only carry out a specialised job. Why in a somatic or differentiated cell, the genome that contains the same DNA as embryonic stem cell can only form one cell-type-associated chromatin configuration and has lost the ability to form the rest ones? No one can answer this question at present.

We think that cell-type-associated chromatin configuration is the highest-level 3-dimensional (3D) genome architecture which sets up a foundational framework for gene expression pattern in each type of cell. Currently, the techniques used in 3D genome architecture study mainly focus on chromatin interactions [6] [7], which belongs to regional or low-level 3D genome architecture study and is within the framework set up by cell-type-associated chromatin configuration. Conceivably, just based on this kind of 3D genome architecture study, it is impossible to reveal the mystery of how cell-type-associated chromatin configuration is formed during cell differentiation. In order to theoretically resolve this issue, we propose a hypothesis in this paper.

## 2. The Hypothesis

A group of unknown proteins might be involved in the formation of cell-type-associated chromatin configurations during cell differentiation. These proteins could be named as rivet proteins which can fasten chromatin fibres up to form cell-type-associated chromatin configuration that sets up a foundational framework for gene expression pattern in each type of cell. Different distribution patterns of these rivet proteins in the 3D genome architectures could form a large number of cell-type-associated chromatin configurations, resulting in many different types of cells.

## 3. Detailed Explanation of the Hypothesis

Formation of cell-type-associated chromatin configuration involves various factors, such as DNA, histones, and non-histone proteins. However, the key proteins in building cell-type-associated chromatin configurations are still missing. We speculated that a group of unknown proteins might function as removable rivets to fasten chromatin fibres up in the 3D genome architecture to form various cell-type-associated chromatin configurations, for example, the genome in human embryonic stem cell can form about 200 cell-type-associated chromatin configurations, resulting in about 200 types of cells. These proteins could be named as rivet proteins, each of which might function as a fastener or co-work

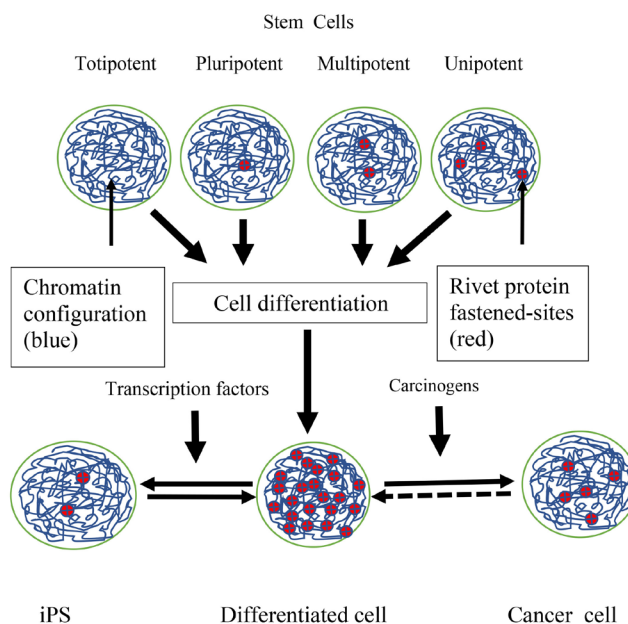
with other proteins to form a complex-fastener. The unoccupied rivet protein fastened-sites could be named as rivet holes which are the structures formed by DNA and related proteins during cell differentiation. The numbers and distributions of rivet holes in the 3D genome architectures of different types of cells are different, resulting in different cell-type-associated chromatin configurations. Therefore, it is reasonable to conclude that different types of cells have different rivet protein fastened-patterns in their 3D genome architectures which correspond directly to their cell-type-associated chromatin configurations.

Cell differentiation could be considered as the process in which rivet proteins are gradually put into rivet holes in the 3D genome architectures of cells that are undergoing differentiation. Usually, once cell differentiation is completed the rivet protein fastened-pattern in a differentiated cell's 3D genome architecture will not change throughout its life time. However, the rivet proteins are removable during cell cycle progression. Presumably, when DNA replication begins, or when chromatin fibres are being packaged into chromosomes, the rivet proteins will be gradually removed from the rivet holes and will be gradually put back into the holes when chromosomes begin to uncoil into chromatin fibres in daughter cells. Since every single chromosome is only part of a cell's 3D genome, its 3D structure is not equivalent to the 3D genome architecture or cell-type-associated chromatin configuration described in this paper and thus has nothing to do with gene regulations. In a word, gene expression regulated by the dynamic 3D genome architecture or chromatin configuration during cell cycle progression begins at the point when chromosomes begin to uncoil into chromatin fibres and finishes at the point when chromatin fibres are repackaged into chromosomes [2].

Theoretically, less rivet protein fastened-sites in a cell's 3D genome architecture mean more flexibility in the cell's chromatin configuration, resulting in more gene activities. Therefore, stem cells might have no or less rivet protein fastened-sites in their 3D genome architectures, whereas highly differentiated cells might have the largest number of rivet protein fastened-sites in their 3D genome architectures. In a word, the number of rivet protein fastened-sites in a cell's 3D genome architecture determines the cell's potency (**Figure 1**).

Cell-type-associated chromatin configuration is the highest-level 3D genome architecture, which sets up the foundational framework for gene expression pattern in each type of cell. Recently, the 3D genome architecture studies have found some special chromatin structures called topologically associated domains (TADs) and sub TADs in which chromatin fibres interact more frequently than surrounding regions, indicating they are involved in gene regulations [8] [9]. We think that both TADs and sub TADs belong to low-level or regional chromatin configurations because their activities are still within the framework set up by cell-type-associated chromatin configuration.

Minor low-level chromatin configuration change caused by minor DNA or related protein alterations could affect regional 3D genome architecture but



**Figure 1.** Cell-type-associated chromatin configuration (blue) and number of rivet protein fastened-sites (red) in different types of cells.

might not change the cell-type-associated chromatin configuration. Cell-type-associated chromatin configuration change is a fundamental change, which changes one cell type to another (cell-type transition) and thus could be named as chromatin configuration transition. A normal cell becoming a cancer cell is due to chromatin configuration transition caused by long-term exposure to carcinogens, *i.e.*, transition of a normal cell-type-associated chromatin configuration to a cancer-associated chromatin configuration (CACC) [2]. Induced pluripotent stem cells (iPS or iPSC) is also the result of chromatin configuration transition which is induced by transcription factors [10]. As mentioned above, less rivet protein fastened-sites in a cell's 3D genome architecture mean more potency in the cell's functions. Therefore, the rivet protein fastened-sites in both cancer and iPS cells are much less than those in the cells from which they are derived (Figure 1). In theory, changing CACC to normal cell's chromatin configuration is possible if we can find out right inducing molecules which might shape the rivet protein fastened-landscape in the 3D genome architecture of cancer cell to elicit chromatin configuration transition.

#### 4. Idea for Experimental Test of the Hypothesis

To support this hypothesis, the first thing we need to do is to identify rivet proteins and then to investigate their roles in the formation of cell-type-associated chromatin configuration. The reason why these rivet proteins are still missing is perhaps because rivet proteins are very big in molecular weight (MW) and very small in quantity. If a protein is too big, for example, MW is above 400 - 500 kDa, it is difficult to separate it in routine electrophoresis gels and thus the existence of this protein could be neglected. Compared to genome architectural pro-

tein CTCF and cohesin which are abundant in the boundaries of TADs [8] [9], the rivet proteins are very rare, which is another reason why they are missing.

In order to find out these rivet proteins, an idea is presented here: first, searching the human genome DNA sequence databases and collecting the DNA sequences of various unknown genes which encode big nuclear proteins (MW > 400 kDa) containing DNA-binding and/or histone-binding domains; second, based on the DNA sequences of these genes, designing and producing fusion proteins; third, producing antibodies using these fusion proteins as antigens; fourth, testing these antibodies to see which one can recognise the proteins that are involved in the formation of the 3D genome architecture.

We once discovered a big unknown nuclear protein (MW  $\approx$  430 kDa) in *Plasmodium falciparum*, which is named as *P. falciparum* chloroquine (CQ) resistance marker protein (Pfcrrmp) [11] [12]. The protein contains DNA binding domains and DNMT1-RFD domain which is a potent histone H3 binding domain [13]. Therefore, Pfcrrmp is most likely to be one of genome architectural proteins, probably functioning as a rivet protein involved in the formation of cell-type-associated chromatin configuration in *P. falciparum*. Since the genetic alterations in Pfcrrmp's gene are closely associated with CQ resistance phenotype [14], it is reasonable to speculate that the mutant Pfcrrmp might have no chance to be put back into its original rivet holes in the genome architecture, resulting in alterations in the 3D genome architecture of malaria parasites. It is this aberrant 3D genome architecture that determines the CQ resistance-associated gene expression pattern which determines parasite's CQ resistance phenotype. In a word, CQ-sensitive parasite becoming CQ-resistant parasite is a kind of cell-type transition resulting from chromatin configuration transition.

## 5. Conclusion

Cell differentiation is still a big mystery in the development of multicellular organisms. More than 30 years ago, we proposed that cell differentiation is the process in which different types of cells are endowed with different chromatin configurations [1]. In this paper, we defined this chromatin configuration as cell-type-associated chromatin configuration which sets up a fundamental framework for gene expression pattern in each type of cell. The hypothesis proposed in this paper is an attempt to understand the mechanism by which the cell-type-associated chromatin configuration is formed. To support this hypothesis, identification of rivet proteins is urgently needed.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Li, G.D. (1986) Abnormal Chromatin Configuration and Oncogenesis. *Medicine and Philosophy*, 7, 12-14. (In Chinese)

- [2] Li, G.D. (2019) Further Thoughts on Abnormal Chromatin Configuration and Oncogenesis. *Open Access Library Journal*, **6**, e5185. <https://doi.org/10.4236/oalib.1105185>
- [3] Corces, M.R. and Corces, V.G. (2016) The Three-Dimensional Cancer Genome. *Current Opinion in Genetics & Development*, **36**, 1-7. <https://doi.org/10.1016/j.gde.2016.01.002>
- [4] Flavahan, W.A., Drier, Y., Liao, B.B., Gillespie, S.M., Venteicher, A.S., Stem-Mer-Rachamimov, A.O., Suva, M.L. and Bernstein, B.E. (2016) Insulator Dysfunction and Oncogene Activation in IDH Mutant Gliomas. *Nature*, **529**, 110-114. <https://doi.org/10.1038/nature16490>
- [5] Taberlay, P.C., Achinger-Kawecka, J., Lun, A.T., Buske, F.A., Sabir, K., Gould, C.M., Zotenko, E., Bert, S.A., Giles, K.A., Bauer, D.C., Smyth, G.K., Stirzaker, C., O'Donoghue, S.I. and Clark, S.J. (2016) Three-Dimensional Disorganization of the Cancer Genome Occurs Coincident with Long-Range Genetic and Epigenetic Alterations. *Genome Research*, **26**, 719-731. <https://doi.org/10.1101/gr.201517.115>
- [6] Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing Chromosome Conformation. *Science*, **295**, 1306-1311. <https://doi.org/10.1126/science.1067799>
- [7] Risca, V.I. and Greenleaf, W.J. (2015) Unraveling the 3D Genome: Genomics Tools for Multiscale Exploration. *Trends in Genetics*, **31**, 357-372. <https://doi.org/10.1016/j.tig.2015.03.010>
- [8] Song, S.H. and Kim, T.Y. (2017) CTCF, Cohesin, and Chromatin in Human Cancer. *Genomics & Informatics*, **15**, 114-122. <https://doi.org/10.5808/GI.2017.15.4.114>
- [9] Norton, H.K. and Phillips-Cremins, J.E. (2017) Crossed Wires: 3D Genome Misfolding in Human Disease. *The Journal of Cell Biology*, **216**, 3441-3452. <https://doi.org/10.1083/jcb.201611001>
- [10] 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. <https://doi.org/10.1016/j.cell.2006.07.024>
- [11] Li, G.D. (2006) Nucleus May Be the Key Site of Chloroquine Antimalarial Action and Resistance Development. *Medical Hypotheses*, **67**, 323-326. <https://doi.org/10.1016/j.mehy.2006.02.008>
- [12] Li, G.D. (2007) *Plasmodium falciparum* Chloroquine Resistance Marker Protein (Pfcrrmp) May Be a Chloroquine Target Protein in Nucleus. *Medical Hypotheses*, **68**, 332-334. <https://doi.org/10.1016/j.mehy.2006.07.016>
- [13] Misaki, T., Yamaguchi, L., Sun, J., Orii, M., Nishiyama, A. and Nakanishi, M. (2016) The Replication Foci Targeting Sequence (RFTS) of DNMT1 Functions as a Potent Histone H3 Binding Domain Regulated by Autoinhibition. *Biochemical and Biophysical Research Communications*, **470**, 741-747. <https://doi.org/10.1016/j.bbrc.2016.01.029>
- [14] Li, G.D. (2008) Pfcrrmp May Play a Key Role in Chloroquine Antimalarial Action and Resistance Development. *Medical Hypotheses and Research*, **4**, 69-73.