



OPINION ARTICLE

Tripartite genome of all species [version 1; referees: 1 approved, 2 approved with reservations]

MengPing Long^{1,2*}, TaoBo Hu^{1,3,4*}

¹Division of Life Science, Hong Kong University of Science and Technology, Hong Kong SAR, China

²Department of Neurology, Xiangya Hospital, Central South University, Changsha, China

³Applied Genomics Center, Hong Kong University of Science and Technology, Hong Kong SAR, China

⁴Department of Oncology, The Second Xiangya Hospital, Central South University, Changsha, China

* Equal contributors

v1 First published: 19 Feb 2016, 5:195 (doi: [10.12688/f1000research.8008.1](https://doi.org/10.12688/f1000research.8008.1))
Latest published: 19 Feb 2016, 5:195 (doi: [10.12688/f1000research.8008.1](https://doi.org/10.12688/f1000research.8008.1))

Abstract

Neutral theory has dominated the molecular evolution field for more than half a century, but it has been severely challenged by the recently emerged Maximum Genetic Diversity (MGD) theory. However, based on our recent work of tripartite human genome architecture, we found that MGD theory may have overlooked the regulatory but variable genomic regions that increase with species complexity. Here we propose a new molecular evolution theory named Increasing Functional Variation (IFV) hypothesis. According to the IFV hypothesis, the genome of all species is divided into three regions that are 'functional and invariable', 'functional and variable' and 'non-functional and variable'. While the 'non-functional and variable' region decreases as species become more complex, the other two regions increase.

Open Peer Review

Referee Status:

	Invited Referees		
	1	2	3
version 1 published 19 Feb 2016	 report	 report	 report

- Shailay Dogra**, Vishuo BioMedical Pte Ltd Singapore
- Rahul Banerjee**, Saha Institute of Nuclear Physics India
- Alfredo Pulvirenti**, University of Catania Italy

Discuss this article

Comments (0)

Corresponding authors: MengPing Long (mplong@connect.ust.hk), TaoBo Hu (thuac@connect.ust.hk)

How to cite this article: Long M and Hu T. **Tripartite genome of all species [version 1; referees: 1 approved, 2 approved with reservations]** *F1000Research* 2016, 5:195 (doi: [10.12688/f1000research.8008.1](https://doi.org/10.12688/f1000research.8008.1))

Copyright: © 2016 Long M and Hu T. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: MPL & TBH are recipients of MD-PhD scholarships from Hong Kong University of Science and Technology. This work was supported by the grant to TBH from Central South University, China (2282013bks118).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare there is no conflict of interest.

First published: 19 Feb 2016, 5:195 (doi: [10.12688/f1000research.8008.1](https://doi.org/10.12688/f1000research.8008.1))

Introduction

The structure and function of the genome have been a major question that all researchers want to solve. The current popular view of the genomic structure is represented by the neutral theory. The neutral theory states that the majority of the genome is variable and neutral¹. The variable property of these genomic regions would not change as the complexity of species increases (Figure 1).

“Three Engenders The Myriad Things”

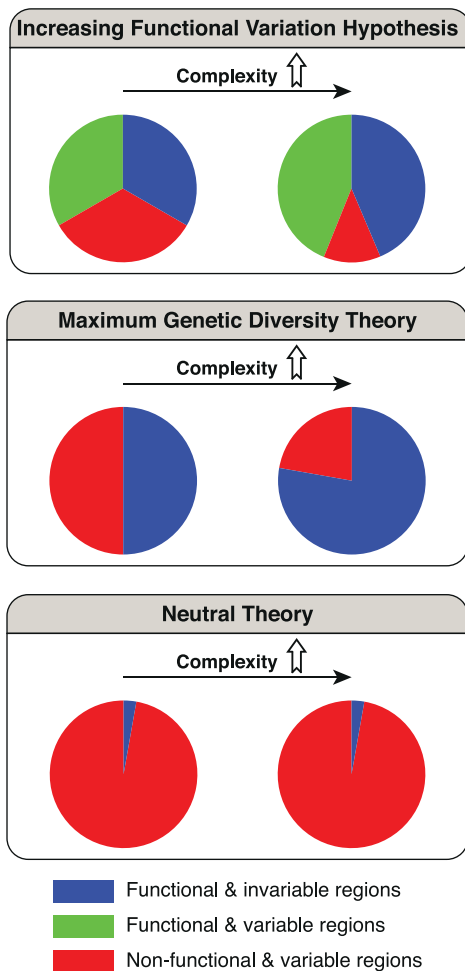


Figure 1. Comparison of IFV, MGD, and neutral theory. While the neutral theory and MGD theory analyze genome structure as bipartite, the IFV hypothesis adds an additional region which is the variable and functional gene regulatory region. As species complexity increases, the variable region of the genome would stay as variable according to neutral theory. While in MGD theory, as species complexity increases there would be less variable region. Unlike MGD theory, IFV hypothesis states that the functional variable region which contains gene regulatory elements would also increase with species complexity.

While in recent years, another theory called Maximum Genetic Diversity (MGD) provided unprecedented insights into the genome structure²⁻⁵. The MGD theory originated from blasting some conserved proteins such as cytochrome C and hemoglobin of different species. By computing the changeable sites of each species³, Huang found that more complex species have less changeable sites in certain regions of the genome. Thus, MGD theory states that as the complexity of species increases, the genome would have more invariable regions and less variable regions (Figure 1).

IFV hypothesis

Here we proposed the Increasing Functional Variation (IFV) hypothesis inspired by both the MGD theory² and our recent work on human genome architecture⁶. Recently, based on co-localization of various genomic features we divided the human genome into three parts, referred to as gene enriched (Genic) zones, gene regulatory elements enriched (Proximal) zones and non-functional features enriched (Distal) zones⁶. We regard the Genic zones as mainly functional and invariable, and the Distal zones as mainly non-functional and variable. The Proximal zones that compose 31% of human genome contain the majority of gene regulatory elements including transcriptional factor binding sites (TFBSs) and are at the same time enriched with conserved indels. These features make Proximal zones functional and variable. It has been proven that as the complexity of species increase, there would be more gene regulatory region in the genome. Based on these two points, we propose that as the complexity of species increases, this variable part of the genome which contains functional regulatory elements would also increase. We call it the Increasing Functional Variation (IFV) hypothesis. Besides the variable gene regulatory region, the other part of the genome can be divided into two parts, the functional and invariable region and the non-functional and variable region. The alteration of these two parts with species complexity can be explained by MGD theory (Figure 1). What the MGD theory lacks and IFV hypothesis complements is the existence of the variable and functional gene regulatory region in the genome. And according to the IFV hypothesis, as species complexity increases, the variable part of the genome would not simply decrease as stated by MGD theory. The differences between IFV hypothesis and MGD theory have been illustrated in Table 1.

Table 1. Comparison of IFV hypothesis and MGD theory.

	IFV	MGD
Genome architecture	Tripartite	Bipartite
Types of variable region	Two	One
Alteration of variable region as species complexity increases	Functional variable region increases while non-functional variable region decreases	Decreases

Conclusions

MGD theory has refuted the idea stated by the neutral theory that the majority of the genome is neutral and variable among all species. Instead, it proved that the variable region of the genome would decrease as species become more complex.

However, MGD theory has its own limitation as pointed out by Ho shortly after the publication of MGD. As Ho has mentioned in her book⁷, more complex species have more sequence diversity, which is needed for precise regulation of local somatic expression. Ho also stated that although MGD theory solved the paradoxes in molecular evolution, the diversity of complex species at somatic level can't be explained by it. Our recent study⁶ on human genome architecture discovered not only variable but also functional regions of the human genome. In an attempt to provide a more comprehensive view of genome structure and molecular evolution, we developed the IFV hypothesis based on our discovery of the variable property of the gene regulatory region.

Why would we develop this tripartite model of genome architecture across all species? As the ancient Chinese philosopher *Lao Tzu*

stated in *Tao Te Ching* thousands of years ago that “*three engenders the myriad things*”, which means “three” is the root of all things. If the truth of the universe is universal, we believe that the consistency between our tripartite genome architecture of all species and *Lao Tzu's* philosophical thinking is not a coincidence.

Author contributions

TBH & MPL jointly conducted this work and wrote the manuscript.

Competing interests

The authors declare there is no conflict of interest.

Grant information

MPL & TBH are recipients of MD-PhD scholarship from Hong Kong University of Science and Technology. This work was supported by the grant to TBH from Central South University, China (2282013bks118).

I confirm that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Nei M, Suzuki Y, Nozawa M: **The neutral theory of molecular evolution in the genomic era.** *Annu Rev Genomics Hum Genet.* 2010; **11**: 265–89.
[PubMed Abstract](#) | [Publisher Full Text](#)
2. Hu T, Long M, Yuan D, *et al.*: **The genetic equidistance result: misreading by the molecular clock and neutral theory and reinterpretation nearly half of a century later.** *Sci China Life Sci.* 2013; **56**(3): 254–61.
[PubMed Abstract](#) | [Publisher Full Text](#)
3. Huang S: **The Overlap Feature of the Genetic Equidistance Result—A Fundamental Biological Phenomenon Overlooked for Nearly Half of a Century.** *Biological Theory.* 2010; **5**(1): 40–52.
[Reference Source](#)
4. Huang S: **Inverse relationship between genetic diversity and epigenetic complexity.** Preprint available at Nature Proceedings. 2009.
[Publisher Full Text](#)
5. Huang S: **The Genetic Equidistance Result of Molecular Evolution is Independent of Mutation Rates.** *J Comput Sci Syst Biol.* 2008; **1**: 92–102.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Ng SK, Hu T, Long X, *et al.*: **Feature co-localization landscape of the human genome.** *Sci Rep.* 2016; **6**: 20650.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Ho MW: **Development and Evolution Revisited.** In: Hood KE, Halpern CT, Greenberg G, Lerner RM, editors. *Handbook of Developmental Science, Behavior, and Genetics.* Blackwell Publishing Ltd; 2010.
[Publisher Full Text](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 16 June 2016

doi:10.5256/f1000research.8617.r14116



Alfredo Pulvirenti

Department of Clinical and Experimental Medicine, University of Catania, Catania, Italy

The paper, in its current version, presents several major limitations.

The authors should provide more details about the theories.

Furthermore, the rationale behind their IFV hypothesis should be given in a deeper way. Examples supporting the conclusions are strongly needed. The authors should also highlight the limitations and the opportunities of this hypothesis.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.

Referee Report 03 May 2016

doi:10.5256/f1000research.8617.r13461



Rahul Banerjee

Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, Kolkata, India

The manuscript by Long and Hu is of considerable interest in the field of evolution of genomes. However the authors could consider the following points to improve the quality of the manuscript :

1. Is there any numerical measure of species complexity ?
2. On page 2, column 2, line 6

'Huang found that more complex species have less changeable sites in certain regions of the genomes.'

It would be nice if some more details could be included with regard to these sites.

3. The authors should be more elaborate as to why they considered the Genic zones to be mainly functional and invariable and the Distal zones as mainly non-functional and variable, and the Proximal zones functional and variable.
4. The note about Lao Tzu was also quite attractive and added some philosophical charm to the scientific note.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Competing Interests: No competing interests were disclosed.

Referee Report 29 March 2016

doi:[10.5256/f1000research.8617.r12943](https://doi.org/10.5256/f1000research.8617.r12943)



Shaillay Dogra

Vishuo BioMedical Pte Ltd, Singapore, Singapore

It would be nicer for the readers if there was more explanation and discussion in the text on the various different theories, what they imply, their pros and cons etc. Currently, it requires some background reading on these topics for the not so well-informed reader.

A detailed discussion of the basis of the IVF hypothesis, the basis of the proposals, its criticism etc. is suggested for a greater understanding of the author's ideas.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Competing Interests: No competing interests were disclosed.
