

ГЕТЕТИКА GENETICS



Invited Editorial

DOI: 10.18413/2658-6533-2020-6-1-0-1

Henry H. Heng

**New data collection priority: focusing on genome-based
bioinformation**

Center for Molecular Medicine and Genomics, and Department of Pathology,
Wayne State University School of Medicine, Detroit, MI, United States,
3226 Scott Hall, 540 E. Canfield, Detroit, MI 48201, USA
Corresponding author: Henry H. Heng (hheng@med.wayne.edu)

Abstract

Genetic research used to be data-driven under the framework of well-established gene theory. With the advance of various-omics technologies, large-scale data generation has become routine. However, data analyses have unexpectedly become extremely challenging due to the highly heterogeneous nature of bio-data. Furthermore, these diverse data sets seem to be inconsistent with some key predictions of gene theory. To briefly address this new reality, this editorial suggests a new genome theory to explain the new facts. By briefly comparing cytogenetics and gene sequencing research, the importance and key principles of chromosomal coded bioinformation is highlighted in the context of evolutionary studies and disease research.

Keywords: genome chaos; genome theory; karyotype coding; missing heritability; non clonal chromosome aberrations; system inheritance; two-phases of cancer evolution

For citation: Heng HH. New data collection priority: focusing on genome-based bioinformation. Research Results in Biomedicine. 2020;6(1):5-8. DOI: 10.18413/2658-6533-2020-6-1-0-1

While classical cytogenetic analyses were technically challenging (e.g. extensive training and practice are required to perform different banding methods and precisely identify altered chromosomes and chromosomal fragments), the theoretical perspective motivating it was rather simple: chromosomes are the carriers of genes, and clonal chromosomal aberrations are the focus for analyses [1, 2]. Molecular cytogenetics has certainly made cytogenetic analyses much easier. Although the higher order of chromosomal structure is still unsolved, karyotype analyses are working

so far, especially with the help of various FISH technologies, including SKY and M-FISH.

The discovery of CNVs (copy number variations) generated high excitement, and array CGH (comparative genomic hybridization) has become routine in cytogenetic labs. For years, there have been discussions and predictions on whether array CGH should replace classical karyotype analyses [1, 3]. One of the key rationales of such discussion is simple: array CGH provides much higher resolution, and in the era of molecular biology,

higher molecular resolution is considered higher quality, more valuable information. For many researchers, only molecular sequencing approaches can reveal any “mechanistic” understanding: cytogenetic profiles are merely descriptive [4]. No wonder “sequencing everything” has become a new trend for data collection in biological research.

Ironically, the surprising results of various large scale-omics studies, including the Human Genome Project, Personal Genome Project, and Cancer Genome Project, have forcefully challenged the molecular rationale that focuses on “molecular part characterization” rather than system behavior at more suitable levels of genomic-systems.

For example, the “highly predictable” relationship between genes and phenotypes has become increasingly complicated, even more so with high degrees of uncertainty. Most complex traits often cannot be explained by polygenic models as the missing heritability seems to be caused by the emergent system behavior of a closely connected yet adaptive genomic network within dynamic environments. Such emergent properties are hard to link to a limited number of either dominant genomic or environmental factors, even with whole genome scanning methodologies based on huge numbers of samples. A system is much more than the collection of all parts.

Similarly, a more specific puzzle comes from the current Cancer Genome Project. Its goal, based on the gene mutation theory of cancer, is to identify a key limited number of common cancer gene mutations. Presumably, by sequencing many cancer samples for each cancer type, the pattern of driver genes will emerge from the background of heterogeneity or “genomic noise.” Quite the opposite is true, however. What was detected was a high degree of heterogeneity (at multiple genomic and nongenomic, including epigenetic, levels), which reflects the stochastic nature of genomic changes in most cancers [1, 2].

While it is disappointing to fail the initial goal of identifying common mutation drivers, the sequencing data did confirm the importance of cytogenetic findings that chromosomal changes are the common drivers of

cancer evolution. First, it confirms the two phases of cancer evolution (punctuated and stepwise) in most cancer types [5, 6]; Second, it endorses the importance of Genome Chaos in evolutionary phase transition [7-9]; Third, it supports the significance of heterogeneity in cancer, as previously ignored “genomic noise” actually represents evolutionary potential and can be used as a new biomarker to monitor system instability [1-2]; and Fourth, it validates the conclusion that profiles of chromosomal abnormalities have much better clinical prediction value than gene mutation data [10].

Why is cytogenetic data more powerful than individual gene mutation data? The genome theory of cancer and organismal evolution offers a clear explanation: the chromosome is not just the carrier of genes, but also the genomic organizer at a higher level of organization. More specifically, the chromosomal sets encode a new type of bio-information above genes. In the perspective of systems biology, the network of gene interaction is the key. However, what defines the gene interaction is unknown. The genome theory proposes that the physical platform of the gene interaction is provided by the karyotype coding, where the physical relationship of genes/regulatory elements along and among chromosomes can form action domains within the 3D nuclei. When the karyotype coding is changed by chromosomal numerical or/and structural changes, new action domains will form, leading to changes of the gene interaction relationship, often at the global level [11]. Those variants form the genomic basis for different types of diseases when interacting with dynamic environments [12].

Knowing the ultimate importance of genome-encoded information (system inheritance) and its relevance to biological processes, from evolutionary mechanisms to implications in disease formation, a new research priority needs to be established in the field of genomics and evolution, as well as molecular medicine [13]. First, a new framework embracing the multiple levels of genomics needs to be established by accepting

the fact that not all levels of organization are equal. Long-ignored chromosomal level profiling might be more useful than DNA sequencing, especially during the macro-cellular evolutionary phase. Second, efforts are needed to build technical platforms suitable for chromosomal profiling [14-19]. There is much to be done: a) Systematic identification of different types of chromosomal numerical/structural changes, many of which have been considered “noise” in previous studies due to the lack of clonality; b) Understanding the relationship among various subtypes of chromosomal abnormality (from aneuploidy to polyploidy, from simple translocations to chaotic genome) and comparing their contributions to disease processes; c) Acceptance of using Non-Clonal Chromosomal aberrations or NCCAs to study genome instability and unify different molecular mechanisms; d) Profiling both individual cells and population dynamics in watch evolution in action experiments. Third, new knowledge of genome-based genomics and evolution needs to be applied into disease studies, especially when dealing with complex issues such as genomic mosaicism, phase transitions of somatic evolution, and the interaction between genomics and environment [20-22].

It is thus timely to write this editorial to remind readers of the importance of data collection at the chromosomal level in current biomedicine. Fortunately, the genome theory has outlined the key rationales and conceptual frameworks of genome-based genomic and evolutionary ideas, based on genome-defined bioinformation principles [1].

No conflict of interest was recorded with respect to this article.

References

1. Heng HH. Genome Chaos: Rethinking Genetics, Evolution, and Molecular Medicine. Cambridge, MA, USA: Academic Press Elsevier; 2019.
2. Heng HH. Debating Cancer: The Paradox in Cancer Research. Singapore: World Scientific Publishing Co; 2015.
3. Liehr T. "Classical cytogenetics" is not equal to "banding cytogenetics". Mol Cytogenet. 2017;10:3. DOI: 10.1186/s13039-017-0305-9
4. Heng HH, Regan SM, Liu G, et al. Why it is crucial to analyze non clonal chromosome aberrations or NCCAs? Mol Cytogenet. 2016;9:15. DOI: 10.1186/s13039-016-0223-2
5. Heng HH, Stevens JB, Liu G, et al. Stochastic cancer progression driven by non-clonal chromosome aberrations. J Cell Physiol. 2006 May 10;208(2):461-72. DOI: <https://doi.org/10.1002/jcp.20685>
6. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. Nature. 2011 March 11;472:90-94. DOI: <https://doi.org/10.1038/nature09807>
7. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011 January 7;144(1):27-40. DOI: <https://doi.org/10.1016/j.cell.2010.11.055>
8. Erenpreisa JE, Ivanov A, Dekena G, et al. Arrest in metaphase and anatomy of mitotic catastrophe: mild heat shock in two human osteosarcoma cell lines. Cell. Biol. Int. 2000 Feb;24(2):61-70. DOI: <https://doi.org/10.1006/cbir.1999.0466>
9. Liu G, Stevens JB, Horne SD, et al. Genome chaos: survival strategy during crisis. Cell Cycle. 2014 Dec 3;13(4):528-37. DOI: <https://doi.org/10.4161/cc.27378>
10. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the evolution of non-small-cell lung Cancer. N Engl J Med. 2017 June 1;376(22):2109-21. DOI: 10.1056/NEJMoa1616288
11. Ye CJ, Stilgenbauer L, Moy A, et al. What Is Karyotype Coding and Why Is Genomic Topology Important for Cancer and Evolution? Front Genet, 2019 November 1;10:1082. DOI: <https://doi.org/10.3389/fgene.2019.01082>
12. Ye CJ, Regan S, Liu G, et al. Understanding aneuploidy in cancer through the lens of system inheritance, fuzzy inheritance and emergence of new genome systems. Molecular cytogenetics. 2018 May 10;11(1):31. DOI: <https://doi.org/10.1186/s13039-018-0376-2>
13. Ye CJ, Sharpe Z, Alemara S, et al. Micronuclei and Genome Chaos: Changing the System Inheritance. Genes (Basel). 2019 May 13;10(5):E366. DOI: 10.3390/genes10050366
14. Rangel N, Forero-Castro M, Rondón-Lagos M. New insights in the cytogenetic practice: Karyotypic chaos, non-clonal chromosomal alterations and chromosomal instability in human cancer and therapy response. Genes. 2017 June

- 3;8(6):155. DOI: <https://doi.org/10.3390/genes8060155>
15. Frias S, Ramos S, Salas C, et al. Non-clonal chromosome aberrations and genome chaos in somatic and germ cells from patients and survivors of hodgkin lymphoma. *Genes (Basel)*. 2019 January 10;10(1):37. DOI: <https://doi.org/10.3390/genes10010037>
16. Liehr T. Cytogenetically visible copy number variations (CG-CNVs) in banding and molecular cytogenetics of human; about heteromorphisms and euchromatic variants. *Mol Cytogenet*. 2016 January 22;9:5. DOI: <https://doi.org/10.1186/s13039-016-0216-1>
17. Niederwieser C, Nicolet D, Carroll AJ, et al. Chromosome abnormalities at onset of complete remission are associated with worse outcome in patients with acute myeloid leukemia and an abnormal karyotype at diagnosis: CALGB 8461 (Alliance) *Haematologica*. 2016;101:1516-1523. DOI: [10.3324/haematol.2016.149542](https://doi.org/10.3324/haematol.2016.149542)
18. Zhang S, Mercado-Uribe I, Xing Z, et al. Generation of cancer stem-like cells through the formation of polyploid giant cancer cells. *Oncogene*. 2014;33:116-128. DOI: <https://doi.org/10.1038/onc.2013.96>
19. Heng HH, Liu G, Stevens JB, et al. Karyotype heterogeneity and unclassified chromosomal abnormalities. *Cytogenet Genome Res*. 2013 April;139(3):144-157. DOI: <https://doi.org/10.1159/000348682>
20. Iourov IY, Vorsanova SG, Yurov YB. Chromosomal mosaicism goes global. *Mol. Cytogenet*. 2008 November 25;1:26. DOI: <https://doi.org/10.1186/1755-8166-1-26>
21. Iourov IY, Vorsanova SG, Yurov YB, et al. Ontogenetic and Pathogenetic Views on Somatic Chromosomal Mosaicism. *Genes (Basel)*. 2019 May 19;10(5):E379. DOI: <https://doi.org/10.3390/genes10050379>
22. Heng HH, Regan S, Ye CJ. Genotype, environment, and evolutionary mechanism of diseases. *Environ Dis [Internet]*. 2016 [cited 2020 Feb 29];1:14-23. Available from: <http://www.environmentmed.org/text.asp?2016/1/1/14/180332>

Received 15 January 2020

Revised 17 February 2020

Accepted 20 February 2020

Information about the authors

Henry H. Heng, PhD, Professor, Head of Laboratory, Center for Molecular Medicine and Genomics, and Department of Pathology, Wayne State University School of Medicine, Detroit, MI, United States, E-mail: hheng@med.wayne.edu, ORCID: 0000-0003-1475-9547.