BIOENGINEERED
2016, VOL. 7, NO. 2, 57–59
http://dx.doi.org/10.1080/21655979.2016.1142635

Taylor & Francis
Taylor & Francis Group

COMMENTARY

OPEN ACCESS

Human genome regulation

Christian Bach and Prabir Patra

Department of Biomedical Engineering, University of Bridgeport, CT, USA

ARTICLE HISTORY Received 6 January 2016; Accepted 11 January 2016

KEYWORDS fractal globule; genetic information; genome origami; origin; transcription factor; wholeness

In the early 90s, Manfred Eigen lectured at Hoffman-La Roche and explained that a gene of 1,000 base pairs has 10^{605} combinations. This number suggests that the Human Genome cannot be randomly created in time.[1] In order to give human beings the power to control any disease including aging, we must, first, accept the definition that the Human Genome contains an infinite amount of information (nonlocal) that morphs into an infinite number of genes (local) as reflected in origami shapes and fractal[2,3] structures.[4-11] Based on this viewpoint, the Human Genome, which consists of 3 billion bps, is categorized into 2 parts. The expression (local) part consists of the 20,687 known protein-coding genes that occupy 2.94% of the Human Genome[12]; whereas, the remaining 97% is referred to as the unknown creation (nonlocal) part. The 20,687 functional genes are developed through unknown processes and mechanisms within the creation part. We postulate that nature not only uses the 1,000 gene-regulating proteins (transcription factor)[13] to regulate the 20,687 known genes, but also that a proportion of these genes are potentially involved in the unknown processes and mechanisms within the Human Genome's 97% creation part. From these 1,000 gene-regulating proteins, the zinc finger domains with 3 fingers recognize 9 base pairs that occur 10,000 times in the Human Genome. In addition, such domains identify specific 9 base pairs which vary under different conditions.[14] Therefore, it is plausible that 3-finger-domains are involved in the unknown processes and mechanisms of the creation part. Further research is necessary to explore how and why nature uses small, 3-finger-domains. If we understand how nature uses these 3-finger-domains to regulate genes, then we might begin to comprehend how nature controls the unknown parts of the Human Genome.[14] Since each single zinc finger binds to a variety of the 64 possible trinucleotide combinations, we can anticipate that a 3-finger-domain binds to numerous 9-base-pair sequences. When a zinc finger protein binds to numerous sequences, it is only active at the desired target locations; whereas, if the protein binds at off-target locations, then it is not active. In addition, when the protein is active at a target location, the binding must be reversible in a very timely manner to avoid over- or under-expression that lead to pathogenic side-effects.[13] Although, if the natural process only guides and binds the protein to a targeted location and further prevents the domain from binding at off-target locations, then the process becomes too complex and inefficient for nature to control. So, why then, does nature rely on 3-finger-domains? We propose that 3-finger-domains are versatile[15] and that the same domain can be used to bind in different conditions to other 9-base-pair sequences in order to execute diverse functions in both the 3% and the 97% sections of the Human Genome. In contrast, domains with 6 or more fingers are not versatile because such domains would bind to limited locations in the 3% section of the Human Genome. Moreover, an increased number of fingers in a domain decrease the control of the reversibility of the binding.[14] In order to explain how nature controls 3-finger-domains, we should develop methods that distinguish between active target sequences and inactive off-target sequences. Out of the 262,144 combinations of a 9 base pair sequence, the active target sequences are those where a domain induces gene expression in an in-vivo model or a human cell based assay. A more practical approach would be to test each finger with the 64 trinucleotide combinations in order to determine active

CONTACT Christian Bach cbach@bridgeport.edu 221 University Ave., Bridgeport, CT 06604 USA

9-base-pair sequences.[14] Each of these identified active target sequences occur 10,000 times in the Human Genome. This vast quantity of sequences must be mapped.[14] Then, the surrounding sequences should be analyzed to determine similarities, patterns, and structures that are clustered in certain areas in the Human Genome. David Bohm's theory states that nature is an infinite wholeness with an implicate order created neither randomly nor planned by outside forces.[16] His theory supports our premise that the Human Genome infinitely creates from within itself. Genes instantaneously appear[1,17-19] and nature directs their placement.[15] Once we discover how these underlying processes and mechanisms operate such as how nature regulates 3 finger domains, then researchers can identify techniques to regulate the Human Genome.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank members of the Center of Functional Nanotechnology at BNL for providing scientific insights and technological support. Grateful acknowledgement is given to Camy Deck for her valuable input and extraordinary patient.

Funding

Research was carried out in part at the Center for Functional Nanomaterials, Proposal no. 30097, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract no. DE-AC02-98CH10886.

References

[1] Eigen M. The origin of genetic information: viruses as models. Gene 1993; 135(1–2):37-47; PMID:8276276

[2] Mirny LA. The fractal globule as a model of chromatin architecture in the cell. Chromosome Res 2011; 19(1):37-51; PMID:21274616; http://dx.doi.org/10.1007/s10577-010-9177-0

[3] McNally JG, Mazza D., Fractal geometry in the nucleus. EMBO J 2010; 29(1):2-3; PMID:20051993; http://dx.doi.org/10.1038/emboj.2009.375

[4] Bach C, Belardo S. The General Theory of Information: Origin of Truth and Hope2012, North Charleston, SC: CreateSpace at Amazon.com.

[5] Demaine ED, O'Rourke J. Geometric Folding Algorithms: Linkages, Origami, Polyhedra2007, Boston: Cambridge University Press

[6] Cremer T, Cremer M, Hübner B, Strickfaden H, Smeets D, Popken J, Sterr M, Markaki Y, Rippe K, Cremer C. The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. FEBS Letters 2015; 589(20 Pt A):2931-43; PMID:26028501

[7] Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 2012; 148(3):458-72; PMID:22265598; http://dx.doi.org/10.1016/j.cell.2012.01.010

[8] Thévenin A, Ein-Dor L, Ozery-Flato M, Shamir R. Functional gene groups are concentrated within chromosomes, among chromosomes and in the nuclear space of the human genome. Nucleic Acids Res 2014; 42(15):9854-61; PMID:25056310; http://dx.doi.org/10.1093/nar/gku667

[9] Jost D, Carrivain P, Cavalli G, Vaillant C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. Nucleic Acids Res 2014; 42(15):9553-61; PMID:25092923; http://dx.doi.org/10.1093/nar/gku698

[10] Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009; 326(5950):289-93; PMID:19815776; http://dx.doi.org/10.1126/science.1181369

[11] Bach C, Patra PK, Pallis JM, Sherman WB, Bajwa H. Strategy for naturelike designer transcription factors with reduced toxicity. IEEE Trans Comput Biol Bioinformat 2013; 10(5):1340-3; PMID:24384718; http://dx.doi.org/10.1109/TCBB.2013.107

[12] Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. Nature 2012; 489(7414):57-74.

[13] Bellefroid E. The human genome contains hundreds of genes coding for finger proteins of the Kruppel type. DNA 1989; 8(6):377-87; PMID:2505992; http://dx.doi.org/10.1089/dna.1.1989.8.377

[14] Bach C, Sherman W, Pallis J, Patra P, Bajwa H. Evaluation of novel design strategies for developing zinc finger nucleases tools for treating human diseases. Biotechnol Res Int 2014; 2014(Article ID 970595):1-27; PMID:24808958; http://dx.doi.org/10.1155/2014/970595

[15] Weirauch MT, Hughes TR. Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. Trends Genet 2010; 26(2):66-74; PMID:20083321; http://dx.doi.org/10.1016/j.tig.2009.12.002

[16] Bohm D. Wholeness and the implicate order 2002, New York: Routledge.

[17] Huntley S, Baggott DM, Hamilton AT, Tran-Gyamfi M, Yang S, Kim J, Gordon L, Branscomb E, Stubbs L. A comprehensive catalog of human KRAB-associated zinc

finger genes: insights into the evolutionary history of a large family of transcriptional repressors. Genome Res 2006; 16(5):669-77; PMID:16606702; http://dx.doi.org/10.1101/gr.4842106

[18] Emerson RO, Thomas JH, Adaptive evolution in zinc finger transcription factors. PLoS Genet 2009; 5(1): e1000325; PMID:19119423; http://dx.doi.org/10.1371/journal.pgen.1000325

[19] Kellis M, Birren BW, Lander ES, Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 2004; 428(6983):617-24; PMID:15004568; http://dx.doi.org/10.1038/nature02424