The transcription factors and its implications in adaptive evolution

Víctor A. Zapata Trejo[1, *]

[1]Department of Marine Biology, Autonomous University of Yucatan, Campus of Biological and Agricultural Sciences, highway Merida-Xmatkuil Km. 15.5, Merida 97000, Mexico.

*Correspondence: A16018093@alumnos.uady.mx

Telephone number: +52 999 207 3228

ORCID: 0000-0002-8930-4195

Abstract The random to explain the emergence of variations in sequence of the alleles is the current scientific paradigm of evolutionary biology. Here is argued that interactions between instincts and transcription factors (TFs) are the main cause of emergence of such variations. Advances in epigenomics show that this molecular function plays an important role in the regulation of gene expression of all cells, both prokaryotes and eukaryotes, and indicates which specific genes should be transcribed and which should be translated. Under this context, the present work pretends evaluate the current evidence on the increase in mutation rate caused by transcription-associated mutational pressure in primordial germ cells due to presence of TFs also present in somatic cells involved in an instinct. In conclusion is established that adaptive evolution can understood as biological superposition of 4 functional states. This work that is added to the evolutionary theoretical framework contributes with an alternative causal understanding of adaptive evolution.

Keywords adaptive system, gametogenesis, stimuli, stochastic, robustness.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

New data, theoretical findings and approaches on niche construction, developmental plasticity and organism-environment feedback, suggest that is needed an alternative causal understanding of adaptive evolution[1-5]. Evolutionary biology currently accepts that DNA sequences transmitted in eukaryotic gametes are not affected by the experiences of individuals, as well as that variations in sequence of the alleles arise at random and not in response to any need of the organisms[6].

This paper evaluates the implications of transcription factors (TFs) in mutation rate of somatic cells and primordial germ cells and complement the hypothesis, as far as I am aware, initially proposed by Bernard[7] with premises that explains how gene expression in somatic cells caused by an instinct increase mutation rate of same genes in primordial germ cells before and during the gametogenesis[8]. Here is established that random does not exist and instead is offered a causal explanation that involves the study of interactions between instinct and TFs.

Transcription-associated mutational pressure

If we knew the number of molecules of the mRNAs that integrate the transcriptome of a certain cell type in any moment and the number of nucleotides of the genes that to the mRNAs correspond, would be possible calculate the stochastic minimum (1) for possible mutations considering the following: the probability of that a gene after finished the transcription has undergone at least 1 mutation during the interaction will always be $1/2$.

equation (1)

$$\text{stochastic minimum} = (1/s)(t)(1/2)$$

Where s is the number of nucleotides of a gene and t the number of molecules of mRNAs that to the gene correspond. The equation describes molecular behavior caused by transcription-associated mutational pressure and is offered as formal theoretical definition.

4 functional states for adaptive evolution

Keeping the above in mind, is argued that emergence of a point mutation in sequence of an allele is result of an increase in mutation rate due mainly by transcription-associated mutational pressure in genetic networks in demand[9, 10]. To understand how this molecular behavior affect adaptive evolution is necessary draw upon to the well-known principle of superposition: a state whose function can understood as sum of the independent functions of each state that

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

composes it. To this principle, for a biological system, is needed add that superposition is a reconfigurable event of dynamic and robust[11] nature (2), that is, a likely event of dynamic configuration in space.

equation (2)

$$f(a + b \ldots \pm n) = f(a) + f(b) \ldots \pm n$$

Where a y b are states and $\pm n$ represents the dynamic of states that can added or retired of the function. Thus, adaptive evolution can understand as biological superposition of 4 functional states, each one composed by a superposition of complex networks (3). The first is stochastic minimum, whose number is in function of gene expression (stochastic state). The following states can guide by the questions: what are mechanisms involved in mutagenesis? what mechanisms intervene in DNA repair, what are their signaling pathways and what TFs are involved? (cellular state); how does substitution of amino acids affect the structure and functionality of proteins and how influence in function of organism? (structural-functional state); how do organisms interact with ecosystem and what are states that define dynamic and robustness of interaction? (ecological state).

equation (3)

$$f(st \in c \subseteq sf \subset e) = f(st) \in f(c) \subseteq f(sf) \subset f(e)$$

Where st, c, sf, and e are the states stochastic, cellular, structural-functional and ecological, respectively.

Robust interactions that respond to external stimuli

The evolutionary theoretical framework explain that basic process of biological evolution is a process at the population level where adaptive evolution is due to shifting gene frequencies by natural selection, from an abundant pre-existing variation[12]. If basic process of biological evolution is a process at the population level, instincts of the population should be determinant for adaptation of organisms. Therefore, instincts must have a role in evolutionary process. However, instincts itself are phenotypes that responds to external stimuli of environment[13, 14], and according to theory of niche construction, this process of dynamic and robust interaction between instincts of a population and its environment can affect selection that acts on same population and on other species[15, 16].

Any function in prokaryotic and eukaryotic cells exerts a genetic demand in genome and there are stimuli that, depending of cell type, activate transcription of required genetic product. The cell recognizes effectively each stimulus

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

through TFs whose evolutionary function is closely related to cell diversity[17]. Therefore, for appear a certain network (cellular or structural-functional) associated to specific stimuli, is needed that cells locate in genome the operator sites of DNA sequences that to TFs correspond[18]. To regulate gene expression, cells use diverse mechanisms that can, for example, silence transcription or suppress translation of a genetic network even when stimuli associated to TFs are present[19, 20]. Some of these mechanisms are DNA methylation and MicroRNAs (miRNAs) that modulate gene expression at the post-transcriptional level[21, 22]. To interaction of these mechanisms with genome to control transcription and translation of a specific genetic network is called epigenomics function[23]. Therefore, epigenomics function have an important role in gene expression of specific genetic networks that cells need to perform its functions in presence of stimuli associated to TFs. All cells of a eukaryotic organism have the same genetic information, but each cell type that integrates it have a different epigenomics function[24].

In eukaryotic organisms there is evidence that show how exposure to a specific stimulus can regulate gene expression of cells involved[25-27]. For example, in zebra finch, the perception of singing is sexually dimorphic and implies that during spermatogenesis, oogenesis, embryogenesis and development the genes involved has sex-biased expression levels[28]. In male, miR-2954 gene associated with habituation of song has higher levels of expression than in female when both sexes are exposed to stimulus of new song although with frequent exposure of same song expression of miR-2954 is gradually inhibited. Therefore, if a new song is frequently repeated, genes that intervene in positive control of transcription associated to habituation of song will be susceptible to higher transcription rates[29].

In theory, operator sites in genes that respond to presence of their TFs are conserved in genetic information of all cells that make up a eukaryotic organism. So, in primordial germ cells is likely that TFs also present in somatic cells act in operator sites to which they are associated. Recently, has been proposed that type of regulator mechanism of gene expression of an instinct can affect likelihood of trait plasticity evolving[30]. Here is added to the above that external stimuli responsible of manifestation of an instinct can influence in transcription rate of somatic cells implicated, and by means of TFs, too in primordial germ cells. Consequently, external stimuli can increase mutation rate in genome, molecular evolution rate and affect likelihood of trait plasticity evolving.

Transcription-associated mutational pressure and GC-rich content

Not long ago was suggested that gene of platelet phosphofructokinase in songbirds has enriched its guanine and cytosine content due mainly by transcription-associated mutational pressure and it is important because the platelet

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

phosphofructokinase is a tissue enzyme that should not be transcribed during gametogenesis and yet it does partially in one or several steps of this process[31]. The compartmentalized of guanine and cytosine is one of defining characteristics to the eukaryotic genome[32]. This quality is associated with positive control of transcription. For example, empirical evidence in studies on pre-implantation in both human and mouse indicated that expression levels increased for those genes that were in regions with GC-rich content while than those that were in regions with GC-poor content decreased expression levels[33]. The GC-rich regions are also associated with negative control of transcription. For example, in exons of some homologous copies of MET1 gene family of hexaploid genome of wheat has been suggested that presence of GC-rich regions was caused by point mutations that induced pseudogenization and DNA methylation[34].

From the above is concluded that exons located in GC-rich regions present evidence of evolutionary changes in sequences associated with control both positive and negative of transcription. If is considered the epigenomics function, is also concluded that transcription-associated point mutations had to occur in primordial germ cells[35], and if in they transcription-associated mutational pressure responds to an increase in genetic demand as in somatic cells[36], could deducted that currently exons located in areas with GC-rich content were under transcription-associated mutational pressure[37], just what has been proposed for platelet phosphofructokinase in songbirds.

In silico genomic evidence already has confirmed the relation between expression levels, transcription-associated mutational pressure and increase observed in mutation rate in yeast as well as the relation between increase observed in mutation rate in germline cells and expression levels in humans[38]. Therefore, scientific solidity of hypothesis initially proposed by Bernard[7] and complemented by this paper is demonstrated.

Biological superposition as a principle to understand adaptive evolution

The relation between TFs and transcription-associated mutational pressure is really superposition of two functional states: the cellular and stochastic, respectively. The other two superpositions correspond to the instinct inside of structural-functional state and to the niche construction inside of ecological state. This approach explains robustness of an adaptive system with their environment where perturbance caused by point mutations is restricted by superposition of complex networks inside ecological state and such mutations are optimal only if cause the upgrade in flow of energy and information[39]. This is where Dawkins's concepts on The Extended Phenotype[40] and The Selfish Gene[41] take on special importance.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

The restriction by superposition of complex networks is also reason of why neutral mutations define the evolution of genome, contrary to what happens at the organism level, where adaptive changes are strictly optimal and not neutral. A neutral mutation can change composition of genome but not affect superposition of networks. As result, evolution rate of genome is higher than observed at the organism level[42].

Three axioms to address the origin of life

In a biological sense: the entropy, number of probable configurations in space for a certain state, decreases as complexity of networks increase. Likewise, the superposition decreases as robustness of networks increase.

One of questions that forces us to consider this way of see natural world is: how does information behave in universe and what are laws that govern its activity?

What is possible add to theory about flow of information is that increases when, in a determinate space, superposition of complex networks does it, and that information to decoded and encoded by organism must travel in packages called information units.

Conclusion

The superposition demonstrated among transcription-associated mutational pressure (stochastic state), TFs (cellular state), instinct (structural-functional state) and the niche construction (ecological state) indicates that gene expression of somatic cells caused by an instinct due to external stimuli can increase transcription-associated mutational pressure in spermatogonias and oogonias by presence of TFs that act in operator sites of same genetic networks that is involved in instinct. Consequently, this system of non-direct interaction between primordial germ cells and environment mediated by TFs have repercussions in developmental plasticity, adaptive evolution, molecular evolution rate and represent a contribution to the evolutionary theoretical framework.

References

1. Gilbert, S.F., Bosch, T.C.G. & Ledón-Rettig, C. Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat Rev Genet*. **16**, 611-622 (2015). doi:10.1038/nrg3982

2. Laland, K.N. *et al*. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc R Soc B*. **282**, 20151019 (2015). doi:10.1098/rspb.2015.1019

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

154 3. Noble, D. Evolution viewed from physics, physiology and medicine. *Interface Focus* **7**, 20160159 (2017).
155 doi:10.1098/rsfs.2016.0159

156 4. Saltz, J.B. & Nuzhdin, S.V. Genetic variation in niche construction: implications for development and evolutionary
157 genetics. *Trends Ecol Evol*. **29**, 8-14 (2014). https://doi.org/10.1016/j.tree.2013.09.011

158 5. Svensson, E.I. On Reciprocal Causation in the Evolutionary Process. *Evolutionary Biology*. **45**, 1-14 (2017).
159 https://doi.org/10.1007/s11692-017-9431-x

160 6. Futuyma, D.J. Evolutionary biology today and the call for an extended synthesis. *Interface Focus*. **7**, 20160145
161 (2017). doi:10.1098/rsfs.2016.0145

162 7. Bernard, D.D. Transcriptional bias: a non-Lamarckian mechanism for substrate-induced mutations. *PNAS*. **86**,
163 5005-5009 (1989). https://doi.org/10.1073/pnas.86.13.5005

164 8. Sandip, P., Million-Weaver, S., Chattopadhyay, S., Sokurenko, E. & Merrikh, H. Accelerated gene evolution
165 through replication–transcription conflicts. *Nature*. **495**, 512-515 (2013). https://doi.org/10.1038/nature11989

166 9. Chen, X. & Zhang, J. No Gene-Specific Optimization of Mutation Rate in *Escherichia coli*. *Mol Biol Evol*. **30**,
167 1559-1562 (2013). https://doi.org/10.1093/molbev/mst060

168 10. Cui, P., Lin, Q., Ding, F., Hu, S. & Yu, J. The Transcript-centric Mutations in Human Genomes. *GPB*. **10**, 11-22
169 (2012). https://doi.org/10.1016/S1672-0229(11)60029-6

170 11. Kitano, H. Biological robustness. *Nat Rev Genet*. **5**, 826-837 (2004). https://doi.org/10.1038/nrg1471

171 12. Stoltzfus, A. Why we don't want another "Synthesis". *Biol Direct*. **12**, 23 (2017). https://doi.org/10.1186/s13062-
172 017-0194-1

173 13. Tinbergen, N. *The Study of Instinct* (Oxford University Press, 1989).

174 14. Yan, H. *et al*. Eusocial insects as emerging models for behavioural epigenetics. *Nat Rev Genet*. **15**, 677-688 (2014).
175 https://doi.org/10.1038/nrg3787

176 15. Constant, A., Ramstead, M.J.D., Veissière, S.P.L., Campbell, J.O. & Friston, K.J. A variational approach to niche
177 construction. *J R Soc Interface*. **15**, 20170685 (2018). doi:10.1098/rsif.2017.0685

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

16. Jänes, H., Herkül, K. & Kotta, J. Environmental niche separation between native and non-native benthic invertebrate species: Case study of the northern Baltic Sea. *Marine Environmental Research*. **131**, 123-133 (2017). https://doi.org/10.1016/j.marenvres.2017.08.001

17. Konstantinides, N. *et al*. Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. *Cell*. **174**, 622-635 (2018). https://doi.org/10.1016/j.cell.2018.05.021

18. Getz, L.J. & Thomas, N.A. The Transcriptional Regulator HlyU Positively Regulates Expression of exsA, Leading to Type III Secretion System 1 Activation in *Vibrio parahaemolyticus*. *J Bacteriol*. doi:10.1128/JB.00653-17 (2018).

19. Liew, Y.J., *et al*. Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral. *Sci Adv*. **4**, eaar8028 (2018). doi:10.1126/sciadv.aar8028

20. Nusrin, S. *et al*. Regulation of CYP11B1 and CYP11B2 steroidogenic genes by hypoxia-inducible miR-10b in H295R cells. *Marine Pollution Bulletin*. **85**, 344-351 (2014). https://doi.org/10.1016/j.marpolbul.2014.04.002

21. Ni, P. *et al*. Methylation divergence of invasive *Ciona* ascidians: Significant population structure and local environmental influence. *Ecol Evol*. https://doi.org/10.1002/ece3.4504 (2018).

22. Zhang, X. *et al*. Identification and characterization of microRNAs involved in ascidian larval metamorphosis. *BMC Genom*. **19**, 168 (2018). https://doi.org/10.1186/s12864-018-4566-4

23. Carlberg, C. & Molnár, F. *Human Epigenomics* (Springer, 2018). https://doi.org/10.1007/978-981-10-7614-5

24. Kenchanmane, R.S.K. & Niederhuth, C.E. Epigenetic Diversity and Application to Breeding. In: Christophe, M. (ed) *Advances in Botanical Research* (Academic Press, 2018). https://doi.org/10.1016/bs.abr.2018.08.001

25. Gunaratne, P.H. *et al*. Song exposure regulates known and novel microRNAs in the zebra finch auditory forebrain. *BMC Genom*. **12**, 277 (2011). https://doi.org/10.1186/1471-2164-12-277

26. McQuaid, J.B. *et al*. Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature*. **555**, 534-537 (2018). https://doi.org/10.1038/nature25982

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

27. Nilsson, B., Jepsen, P.M., Bucklin, A. & Hansen, B.W. Environmental Stress Responses and Experimental Handling Artifacts of a Model Organism, the Copepod *Acartia tonsa* (Dana). *Front Mar Sci*. https://doi.org/10.3389/fmars.2018.00156 (2018).

28. Luo, G.Z. *et al*. Genome-wide annotation and analysis of zebra finch microRNA repertoire reveal sex-biased expression. *BMC Genom*. **13**, 727 (2012). https://doi.org/10.1186/1471-2164-13-727

29. Lin, Y.C., Balakrishnan, C.N. & Clayton, D.F. Functional genomic analysis and neuroanatomical localization of miR-2954, a song-responsive sex-linked microRNA in the zebra finch. *Front Neurosci*. **8**, 409 (2014). https://doi.org/10.3389/fnins.2014.00409

30. Rittschof, C.C. & Hughes, K.A. Advancing behavioural genomics by considering timescale. *Nat Commun*. **9**, 489 (2018). doi:10.1038/s41467-018-02971-0

31. Khrustalev, V.V., Barkovsky, E.V., Khrustaleva, T.A. & Lelevich, S.V. Intragenic isochores (intrachores) in the platelet phosphofructokinase gene of Passeriform birds. *Gene*. **546**, 16-24 (2014). https://doi.org/10.1016/j.gene.2014.05.045

32. Costantini, M., Alvarez-Valin, F., Costantini, S., Cammarano, R. & Bernardi, G. Compositional patterns in the genomes of unicellular eukaryotes. *BMC Genom*. **14**, 755 (2013). https://doi.org/10.1186/1471-2164-14-755

33. Barton, C., Iliopoulos, C.S., Pissis, S.P. & Arhondakis, S. Transcriptome activity of isochores during preimplantation process in human and mouse. *FEBS Lett*. **590**, 2297-2306 (2016). https://doi.org/10.1002/1873-3468.12245

34. Thomas, M. *et al*. Evolutionary history of Methyltransferase 1 genes in hexaploid wheat. *BMC Genom*. **15**, 922 (2014). https://doi.org/10.1186/1471-2164-15-922

35. Percharde, M., Wong, P. & Ramalho-Santos, M. Global Hypertranscription in the Mouse Embryonic Germline. *Cell Rep*. **19**, 1987-1996 (2017). https://doi.org/10.1016/j.celrep.2017.05.036

36. Bergman, J., Betancourt, A.J. & Vogl, C. Transcription-Associated Compositional Skews in Drosophila Genes. *Genome Biol Evol*. **10**, 269-275 (2018). https://doi.org/10.1093/gbe/evx200

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26862v8 | CC BY 4.0 Open Access | rec: 17 Dec 2018, publ: 17 Dec 2018

37. Touchon, M., Nicolay, S., Arneodo, A., d'Aubenton-Carafa, Y. & Thermes, C. Transcription-coupled TA and GC strand asymmetries in the human genome. *FEBS Lett*. **555**, 579-582 (2003). https://doi.org/10.1016/S0014-5793(03)01306-1

38. Park, C., Qian, W. & Zhang, J. Genomic evidence for elevated mutation rates in highly expressed genes. *EMBO rep*. **13**, 1123-1129 (2012). doi:10.1038/embor.2012.165

39. Chen, B.S. & Wu, W.S. Underlying Principles of Natural Selection in Network Evolution: Systems Biology Approach. *Evol Bioinform*. **3**, 245-262 (2007). https://doi.org/10.1177/117693430700300010

40. Dawkins, R. *The Extended Phenotype* (Oxford University Press, 2016).

41. Dawkins, R. *The Selfish Gene* (Oxford University Press, 2016).

42. Zhang, J. Neutral Theory and Phenotypic Evolution. *Mol Biol Evol*. **35**, 1327–1331 (2018). https://doi.org/10.1093/molbev/msy065

Acknowledgements

The author is grateful with the friends, colleagues and academics who contributed to gradual evolution of this paper.

To my mother, the woman who makes my education possible.

Competing interests

The author(s) declare no competing interests.