Did the modulation of expression noise shape the evolution of three dimensional genome organizations in eukaryotes?

Kuljeet Singh Sandhu

Keywords: chromatin interactions, expression noise, evolution, epigenetics, 3D genome organization, nuclear architecture

Introduction

Accumulating evidence suggests the prominent role of non-random three dimensional organization of chromatin in essential genomic functions in the eukaryotic nucleus. The earliest evidence comes from the observation that the random insertion of a gene in distinct genomic locations modulates its transcription distinctly. Long range cis and trans chromosomal interactions correlate strongly with the active or repressive epigenetic states of the gene. Moreover, the spatial positioning of genes in the nucleus matters to their expression. The nuclear core is associated with high transcription activity, while nuclear lamina interactions of chromosomal domains are associated with stable silencing of genes. Recent reports also suggest that most genes are not independent transcriptional units on their own; instead they interact with each other in an extensive manner. Despite these advances, it is yet not clear how the three dimensional organization of genome has evolved into a non-random conformation. Transcriptional activity (or inactivity) and coordination of genes are arguably the best studied constraints of eukaryotic genome organization. Here, I present a few lines of evidence, which suggest that the modulation of expression noise could have been one of the evolutionary constraints, which might have shaped the three dimensional conformation of eukaryotic genome. Highly varying expression levels of genes could be deleterious if it happens at essential gene loci and hence could have been an evolutionary constraint to shape the 3D conformation of genome in a manner to keep the expression noise of essential loci to the minimum. One way a locus could experience the variation in expression is via highly diverse chromatin interactions. Large number of chromatin interactions of a locus generally do not occur at the same time in the same cell due to spatial constraints in the nucleus, and rather represent the aggregation of diverse interactions across a population of cells. Highly varying interactions with respect to time and space might not allow a gene to maintain its constant expression level. I test this hypothesis using the genome wide inter-chromosomal (trans) chromatin interaction data in yeast.

Results

Using the trans chromatin interactions from two biological replicates, I calculated the normalized number of interactions, i.e., degree (δ), for each 3 kb region of the yeast genome (Material and Methods section). I had following key observations:

1. As apparent from the Figure 1A–D, genomic regions with low nucleosome occupancy, abundant expression and rapid rate of nascent transcription show significantly lower degree of trans interactions than the regions with moderate levels of nucleosome occupancy, gene expression and transcription rate (p = 3.8e-13, 1.3e-04, 3.2e-05 respectively), suggesting that the nucleosome depleted domains with highly transcribing genes, are selected against the high variation in trans chromatin interactions.

2. Similarly, regions with low gene activity or the repressed genes also exhibit lower degree (p = 8.8e-05), presumably to ensure their stable silencing or a delicate regulation of precise transcriptional level.

The pervasive role of distant chromatin interactions in transcriptional regulation is increasingly becoming evident. There is a possibility that the greater diversity in chromatin interactions of a genomic locus could contribute to stochastic variation in its gene expression. However, this issue has not been addressed. Here, I present a few lines of evidence, which suggest that the variation in trans chromatin interactions might occur at the cost of expression noise. Genomic regions with nucleosome depletion, abundant and rapid transcription and with essential gene clusters exhibit relatively fewer trans chromatin interactions in the nucleus. Moreover, loci with greater number of interactions tend to show higher expression noise. Based on these observations, I hypothesize that the three dimensional organization of eukaryotic genomes might have evolved under a selective pressure to minimize the expression noise of essential gene clusters in the nucleus.
These observations are also in line with a recent report which suggests that inter-chromosomal DNA interactions occur at the cost of increased expression noise. Since the expression noise data available so far is measured at protein level rather than gene transcription, there is a possibility that noise in gene transcription might have been partly compensated at post-transcriptional level by other regulatory mechanisms yet to be explored in the context, which might explain relatively weaker significance, observed in Figure 1F. Moreover, alterations in chromatin organization would only account for the intrinsic noise, while the expression noise data also includes extrinsic noise contributed by factors like fluctuating micro-environments of the cells.

**Discussion**

The recent boom in the proximity ligation based techniques in combination with next generation sequencing has enabled the mapping of all-to-all chromatin interactions in the nucleus. Interestingly, highly expressed housekeeping genes are shown to converge to transcriptionally active multi-gene complexes for efficient and coordinated transcription. Such coordination of transcription might get disturbed if the locus involved
stochastically moves in or out of active compartment or, in worse case, mixes into a repressed compartment. Therefore, these loci would be expected to be relatively stable at transcription foci sites, than the ones which are not essential and could tolerate the noisy expression. Indeed, highly varying expression level has been proposed to be a facilitator of adaptive gene expression evolution for certain genes and hence could be advantageous in certain contexts. However, importantly, essential genes tend to show low expression noise since noisy expression of essential genes would be deleterious for the organism. Moreover, in order to minimize the expression noise, essential genes have evolved to locate in clusters along the chromosome and harbor consistently open state of chromatin, as shown by others. I show that essential gene clusters might maintain lower expression noise by limiting their total number of interactions, possibly by restricting the global mobility of the loci involved. Essential gene clusters might be tethered to the nuclear sites of abundant and rapid transcription. The hypothesis is strongly supported by the observation that nucleosome depleted, highly and rapidly transcribed regions have lower degree of trans chromatin interactions, but relatively higher degree of cis chromatin interactions (Fig.1A–C, Figs. S2 and S3). It is also notable that the maximal extent of cis interactions per locus is far lower (3.4-fold) than that of trans interactions, suggesting the intra-territorial spatial constraints selectively restricting the mobility of genomic domains of high gene activity. Similarly, the regions with low (or no) gene activity also show lower degree of trans interactions hinting at maintenance of their stable repressive state or low and delicate transcription via limiting the mobility of the locus. Indeed, attachment with sub-nuclear structures like nucleoli or nuclear periphery, which are associated with transcriptional repression, significantly limits the mobility of chromatin domains. Conversely, the regions with moderate levels of expression show relatively higher degree of trans interactions and would be susceptible to higher expression noise. Higher expression noise of moderately expressed genes has been observed elsewhere too. Furthermore, some reports suggest that certain chromatin interactions could also be deleterious or mutagenic. By extrapolation, a region having higher number of trans chromatin interactions, would be under the risk of experiencing such erroneous interactions. Restricting the total number of interactions might also minimize the risk to such mutagenic interactions. Indeed, DNA double strand break signal strongly associates with the regions of higher degree of chromatin interactions (Fig. S4). Therefore, I propose that the nonrandom three dimensional organization of eukaryotic genome might have evolved under the selection constraint to reduce the expression noise or the risk to erroneous or mutagenic interactions of essential gene loci. Having said that, I do not rule out the other evolutionary constraints like coordinated expression of functionally related genes, as being reported elsewhere. More data sets in future would help scrutinize the hypothesis.

**Material and Methods**

Data sets. Genome wide chromatin interaction data was taken from Duan et al. Biological replicate data, generated using two distinct restriction enzymes (EcoRI and HindIII), were used to calculate the normalized degree of chromatin interactions. Nucleosome occupancy data was taken from Kaplan et al. and RNA-Seq data set was taken from Nagalakshmni et al. Data for nascent transcription rate was taken from Pelchano et al. List of essential genes was taken from Saccharomyces Genome Deletion Project (http://www-sequence.stanford.edu/group/yeast_deletion_project/downloads.html). Expression noise data was originally generated by Newman et al. and scaled from 0 to 1 by Li et al. DNA double strand break data for Dicer3 mutant was taken from Buhler et al. Most of the experimental data sets taken are generated on yeast cells grown inYPD growth media in the respective studies.

Analysis. First, the yeast genome was divided into bins of 3 kb (equivalent to the average size of EcoRI/HindIII fragments in the yeast genome). The centers of the restriction fragments were mapped to the 3 kb bins. Data sets for biological replicates (EcoRI and HindIII) were, then, quantile normalized. The quality control plots are shown in the Figure S5A. For each of the biological replicates, undirected network of trans chromatin interactions among all 3 kb bins was constructed using igraph library (http://igraph.wikidot.com/) on R-package (http://www.r-project.org/). Number of interactions (degree) for each 3 kb region was calculated using ‘degree’ function on igraph. Geometric mean of the two replicates was used for further analysis. Binning the data and calculation of geometric mean from normalized EcoRI and HindIII data sets also make the analysis robust against possible biases caused by the length and the local GC content of the distal ends of restriction fragments in individual data sets (Fig. S5B and C). Rolling mean of 10 consecutive windows was taken for plotting and visualization purpose. The yeast genes (S288C genome) were assigned to 3 kb bins by mapping the location of transcription start sites (TSS) to the bins. Mean nucleosome occupancy, gene expression and nascent transcription rate were calculated for each 3 kb window. The cut-off values for degree, nucleosome occupancy, gene expression and transcription rate are elaborated in the Figure S6. More liberal or stringent cut-offs do not alter the conclusions of the article (Fig. S7). To calculate the essential gene density, neighboring four windows of 3 kb (12 kb in total) were considered. High and moderate density of essential genes was demarcated as ≥ 3 and 1–3 genes in the 12 kb region respectively. A detailed degree distribution for the distinct representations of essential genes is given in the Figure S1. Since expression noise data are not available for all the yeast genes, I considered maximum noise value available in a 3 kb bin. The key findings of the study were also scrutinized against the possible sequence mapping biases by removing the telomeric (15 kb) and centromeric (15 kb either side) regions from the analysis (Fig. S8A–D). Moreover, yeast chromosomes are known to be tethered at centromeres and, therefore, show abundant trans interactions at sites proximal to the centromeres (± 50 kb, Fig. S8E). I reassessed the major observations of the study by excluding the regions ± 50 kb to the centromere. The analysis consistently adheres to the hypothesis (Fig. S8F–I). Similarly, the key observations of the study are also robust against the variations in the two replicates, as measured using Shannon’s entropy for normalized EcoRI and
Supplemental Material

No potential conflicts of interest were disclosed.

References

1. Göndör A, Ohlsson R. Chromosome cross-talk in three dimensions. Nature 2009; 462:58-64. PMID:19890323; http://dx.doi.org/10.1038/nature08497.

2. Baldi P, Bello R. Evolution of chromosome organization driven by selection for reduced gene expression noise. Nat Genet 2007; 39:945-9; PMID:17660811; http://dx.doi.org/10.1038/ng2071.

3. Nandi AK, Roginski RS, Gregg RG, Smithies O, Skoultchi AI. Regulated expression of genes inserted at the human chromosomal beta-globin locus by homologous recombination. Proc Natl Acad Sci USA 1988; 85:3943-9; PMID:3375244; http://dx.doi.org/10.1073/pnas.85.11.3845.

4. Driel R. Spatial relationship between transcription sites and chromosome territories. J Cell Biol 1999; 147:13-24. PMID:10508851; http://dx.doi.org/10.1083/jcb.147.1.13.

5. Ghosh L, Papi L, Brown E, Medvedin W, Faust MB, Talbert W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 2008; 453:948-51; PMID:18463634; http://dx.doi.org/10.1038/nature06947.

6. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, et al. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 2009; 462:58-64. PMID:19890323; http://dx.doi.org/10.1038/nature08497.

7. Collins F, Oki N, Parry A, Froud M, Shimamura K. Stochastic gene expression in a single cell. Science 2002; 297:1566-9; PMID:12043669; http://dx.doi.org/10.1126/science.1070919.

8. Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. J Cell Biol 2009; 5:299; PMID:19690568; http://dx.doi.org/10.1083/jcb.200811103.

9. Nomen JB, Ghosnianaghchi I, Brandi J, Beucher DK, Nofs M, Daffis R, et al. Single-cell proteomic analyses of A. constrictus reveal the architecture of biological noise. Nature 2010; 461:212-5; PMID:20695522; http://dx.doi.org/10.1038/nature08453.

10. McCallum-Arden A, van Bohemen NL, Williams L, Ibrahim M, Regnier T, Stallings A, et al. Combinatorial mapping of long-range interactions reveals folding principles of the human genome. Science 2009; 326:209-33; PMID:19137570; http://dx.doi.org/10.1126/science.1183140.

11. Liberek-Adamowicz AW, van Bohemen NL, Williams L, Ibrahim M, Regnier T, Stallings A, et al. Combinatorial mapping of long-range interactions reveals folding principles of the human genome. Science 2009; 326:209-33; PMID:19137570; http://dx.doi.org/10.1126/science.1183140.

12. Fullwood MJ, Liu MH, Pan TE, Liu J, Xu H, Mohamed TH, et al. An estrogen-receptor-alpha-bound human chromatin interactome. Nature 2009; 462:58-64. PMID:19890323; http://dx.doi.org/10.1038/nature08497.

13. Zhang Z, Qin W, Zhang J. Robust selection for reduced gene expression noise in yeast. Mol Syst Biol 2010; 6:64; PMID:20983323; http://dx.doi.org/10.1038/mstpm.2010.64.

14. Skoultchi AI. Regulated expression of genes inserted at the human chromosomal beta-globin locus by homologous recombination. Proc Natl Acad Sci USA 1988; 85:3943-9; PMID:3375244; http://dx.doi.org/10.1073/pnas.85.11.3845.

15. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science 2002; 297:1566-9; PMID:12043669; http://dx.doi.org/10.1126/science.1070919.

16. Lin C, Yang L, Tian A, He L, Ke R, Olbrich K, et al. Nuclear receptor-induced chromosome protrusion and DNA breaks underlie specific translocations in cancer. Cell 2009; 139:1069-83; PMID:19962179; http://dx.doi.org/10.1016/j.cell.2009.11.030.

17. Babu MM, Janga SC, de Santiago I, Pombo A. Endocytic gene regulation in three dimensions and its impact on genome evolution. Curr Opin Genet Dev 2008; 18:71-7; PMID:18076286; http://dx.doi.org/10.1016/j.gde.2008.10.002.

18. Kaplan D, Moore JK, Froud M, Shimamura K, Guacci V, Tel-Or D, Field Y, et al. The DNA-encoded chromatin-organization network of a eukaryotic genome. Nature 2009; 458:362-6. PMID:19092803; http://dx.doi.org/10.1038/nature07687.

19. Nagalakshmi U, Wang Z, Wasson K, Sho C, Raha D, Guttman M, et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 2008; 320:103-6. PMID:18451266; http://dx.doi.org/10.1126/science.1158481.

20. Perelman V, Chao S, Pires-Oliveira J. A complete set of yeast transcription sites for gene-grain. PLoS One 2014; 9:51940. PMID:25101038; http://dx.doi.org/10.1371/journal.pone.0091342.

21. Li J, Ma R, Vinnicombe JF, Je R, Xie X, Zhang Z. Exploring the determination of stochastic gene expression in Saccharomyces cerevisiae for genome-wide prediction of expression noise. Proc Natl Acad Sci USA 2013; 110:167-72; PMID:24049180; http://dx.doi.org/10.1073/pnas.1219030110.

22. Boldt C, Bodey V, Libbrecht M. Mapping vesicular single-stranded DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae. PLoS Biol. 2007; 5:534. PMID:17672067; http://dx.doi.org/10.1371/journal.pbio.0050024.

Acknowledgments

K.S.S. acknowledges the financial support from A*STAR Singapore Joint Council Office (JCO) grant (12302EG016).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

www.landesbioscience.com/journals/nucleus/article/20263

www.landesbioscience.com

www.landesbioscience.com