DOI 10.18699/SBB-2022-059

Insights into the 3D-genome organization in malaria mosquitoes

Lukyanchikova V.[1, 2, 3, 4]*, Nuriddinov M.[3, 4], Belokopytova P.[3, 4], Sharakhov I.[1, 2], Fishman V.[3, 4]

[1] *Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, USA*

[2] *Fralin Life Science Institute, Virginia Polytechnic Institute and State University, Blacksburg, USA*

[3] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[4] *Novosibirsk State University, Novosibirsk, Russia*

* *varvara_lukyanchikova@mail.ru*

Key words: 3D genome organization, long-distance chromatin loops, Hi-C technology, mosquito genomics, chromatin, malaria mosquitoes

Anopheles mosquitoes are the only carriers of malaria disease across the world. Their quick spreading all over the globe and devastating effect on human health raises the dramatic necessity to investigate their genomes in details, especially knowing that gene expression and regulation is tightly linked with 3D chromatin interactions. Comparative studies performed in vertebrates species revealed that genome architecture is evolutionarily conserved and could be explained by the dynamic interplay between processes of cohesin-mediated loop extrusion and chromatin compartmentalization. In insects, comprehensive analyses and cross-species comparisons of genome architecture have to date focused on *Drosophila* species mostly. Using Hi-C technology, we characterized genomic landscape of chromatin in five species of *Anopheles* mosquitoes: *An. coluzzii*, *An. merus*, *An. stephensi*, *An. atroparvus*, and *An. albimanus*. We comprehensively profiled all levels of chromatin organization using Hi-C, ChIP-seq, RNA-seq, and FISH methodologies, and studied connections between the distribution of epigenetic marks and 3D-interactions to provide mechanistic explanations underlying observed chromatin structures (Lukyanchikova et al., 2022). We showed that most 3D-structures can be explained by principles described previously for *Drosophila* species. However, at the level of individual loci, we identified specific, extremely long-ranged looping interactions, which are strikingly conserved across *Anopheles* genera for ~100 million years. We detected specific looping interactions, sometimes spanning several dozens of megabases (Mb) (up to 31-Mb) in all studied *Anopheles* species, and validated those interactions by independent FISH approach. Anchors of those loops were represented by large (~200-300-kb) loci, interacting significantly more than expected at their genomic distance (top 1–2 % of all interactions). We compared the gene content located within the loops anchors and, surprisingly, all species in comparison shared orthologous genes. Hi-C contact maps obtained for adults and embryos of *An. merus* demonstrated that the long-distance loops are present at both developmental stages, indicating that they are not developmental-stage specific. Analysis of gene expression profiles within loop anchors resulted in finding several genes with moderate expression. In accordance with this observation, cePC1 analysis showed that loop anchors were enriched in B-compartment regions, so their formation could not be explained by the clustering of active chromatin. Analysis of ChIP-seq results showed moderate H3K27me3 and H2AK119Ubiq enrichment within our loop anchors while some other regions highly enriched in H3K27me3 and H2AK119Ubiq histone marks and located between the loop anchors were not involved in long distant looping. That fact indicates that long-distance chromatin loops are probably not formed as a result of Polycomb proteins activity (PRC1 and PRC2) but by other molecular mechanism.

Acknowledgements: The research was supported by RSF No. 22-14-00247 grant.