



DNA Loops inside genes affect the resulting protein

Mariana, Ruiz Velasco Leyva^{*}; Manjeet, Kumar; Mang Ching, Lai; Pooja, Bhat; Ana Belen, Solis Pinson; Alejandro Reyes; Stefan Kleinsorg; Kyung-Min, Noh; Toby J., Gibson; Judith B., Zaugg

Laboratorio Europeo de Biología Molecular (EMBL), Heidelberg, Alemania.

mariana.ruiz@embl.de

Nuclear organization is extremely complex and tightly regulated, as two meters of DNA need to fit into the microscopic nucleus. This packaging is mostly achieved with the help of several proteins, which form a fiber known as chromatin that compresses entire chromosomes. Surprisingly, this organization is not random and it allows certain parts of the genome to preferentially and specifically contact others at long distances through the formation of chromatin loops, therefore establishing tissue-specific regulatory landscapes.

In the last years, there has been major progress in elucidating the grammar of this fine nuclear organization. Mostly, a protein called CTCF serves as a stop signal during loop formation, indicating where a loop needs to anchor. When two CTCF molecules find each other in a defined direction, the loop stops elongating and the two distal genomic regions are brought into close physical proximity. These interactions have been mostly described to occur between regions residing very far apart in distance. However, with our study, we identified that CTCF-loops can also occur at smaller distances, in particular at the gene level.

Additionally, we observed that genes containing CTCF-loops were mostly participating in cell signaling and stress response, which made us wonder whether their presence was somehow affecting the way in which the genes are read and copied into RNA transcripts. We decided to test our observations by re-analyzing publicly available data for DNA, CTCF, and RNA from a human cohort. Indeed, we observed that the presence of CTCF-intragenic loops affects splicing (a process that is responsible for deciding which parts of the gene, i.e. exons, will be included into the final transcript).

Interestingly, we also observed that there was a large variability in the presence of CTCF-intragenic loops across the individuals from our cohort, which was explained by the existing variation in their DNA sequence. As ultimately the changes in splicing decisions lead to different versions of a protein, our finding suggests that humans may react differently to cellular stimuli, at least partially through the mechanism that we have identified.

Our study has several implications for understanding gene regulation in the context of 3D chromatin organization, as it is one of the first studies linking nuclear structure with a molecular process. It also opens exciting possibilities of how this additional layer of regulation can play a role in biological processes such as development. Finally, changes in CTCF binding patterns have been described in cancer and other pathologies. Now that we have identified a set of functional genomic regions, it will be relevant to study the consequences of the deregulation of intragenic loops as well as the extent of this source of molecular variation both in humans and other organisms.

1. M. Ruiz-Velasco *et al.*, CTCF-Mediated Chromatin Loops between Promoter and Gene Body Regulate Alternative Splicing across Individuals. *Cell Syst.* **5**, 628–637.e6 (2017).

2. M. Ruiz-Velasco, J. B. Zaugg, Structure meets function: How chromatin organisation conveys functionality. *Current Opinion in Systems Biology.* **1**, 129–136 (2017).