
**DESARROLLO DE UN MÉTODO PARA LA INFERENCIA DE REDES
DE REGULACIÓN DE GENES TIEMPO Y LUGAR ESPECÍFICO:
APLICACIÓN Y ESTUDIOS COMPARATIVOS EN D. MELANOGASTER"**

**LEANDRO MURGAS SAAVEDRA
INGENIERO EN BIOINFORMÁTICA**

RESUMEN

La regulación de la expresión génica es uno de los factores determinantes en el desarrollo y mantenimiento de la vida en todos los organismos. Esta regulación se lleva a cabo principalmente por medio de la acción de factores de transcripción (TFs), aunque otros factores también están involucrados. Notablemente, cualquier avance en el conocimiento sobre la regulación de la expresión es clave para entender el funcionamiento de los distintos organismos a nivel molecular. Este conocimiento permite entender los procesos que desencadenan distintas enfermedades y permitiendo el desarrollo de nuevas dianas terapéuticas. Es por ello que para el estudio de este tipo de interacciones se han desarrollado las Redes de regulación de genes (GRNs). Estas redes contienen información sobre los diversos factores que se encuentran involucrados en la regulación de la expresión de distintos genes. Dada la gran abundancia de datos experimentales sobre los distintos factores que intervienen en la regulación de la expresión génica y el poco conocimiento específico sobre esta regulación en los distintos tejidos y tipos celulares que forman los organismos en estadios concretos de su desarrollo, la creación de nuevos métodos computacionales para integrar toda esta información con el objetivo de inferir redes sitio y condición específicas es de suma importancia para estudios futuros. En este trabajo se realizó un método para la inferencia de GRNs tiempo y lugar específicas, que después se aplicó al estudio del desarrollo embrionario de *Drosophila melanogaster*. El método integra datos experimentales de múltiple origen incluyendo expresión génica, estructura de la cromatina, TFBSs, entre otros. Mediante este método se crearon un total de 12 GRNs divididas en dos tipos: Caso I, aquellas que consideran una regulación si el extremo 3' del sitio de unión de un TF se encuentra dentro de una distancia ΔD al sitio de inicio de la transcripción (TSS) de un gen; y las de Caso II, que consideran que existe una regulación si el TF se une desde el inicio de la región "D" hasta el final del gen. Cada caso fue generado con el valor de ΔD 1500 bases y en tiempos que variaban en 4 horas desde la hora 0 a la 24 abarcando el desarrollo completo del embrión

de *D. melanogaster*. Finalmente, las GRNs se caracterizaron y analizaron mediante técnicas basadas en graphlets, pudiendo de esta forma identificar elementos concretos cuya relación con el resto de la red varían.

ABSTRACT

The regulation of gene expression is one of the determining factors in the development and maintenance of life in all organisms. This regulation is carried out mainly through the action of transcription factors (TFs), although other factors are also involved. Notably, any advance in knowledge on the regulation of the expression is key to understanding the functioning of the various organisms at the molecular level. This knowledge also has direct application in the understanding of the processes that trigger different diseases and allowing the development of new therapeutic targets. That is why the study of such interactions have required the development of Gene Regulatory Networks (GRNs). These networks containing information on the various factors that are involved in regulating the expression of different genes. Given the abundance of experimental data on the various factors involved in the regulation of gene expression and the little specific knowledge of this regulation in different tissues and cell types that form organisms in specific stages of development, the development of new computational methods to integrate all this information aiming inference networking site and specific condition is of crucial relevance for future studies. In this work a method was developed for the inference of GRNs of time and specific location which was later applied to the study of embryonic development of *Drosophila melanogaster*. The method integrates experimental data of multiple origin including gene expression, chromatin structure, TFBSs, among others. Employing this method a total of 12 GRNs were created of two different types were created: Case I, which consider a regulation and the 3' end the binding site of a TF is within a distance D to the transcription start site (TSS) of a gene, and case II, which consider that there is a regulation and the TF for the beginning of the region. Each case was generated with the value of " D " 1500 bases and in times that vary in 4 hours from the hour 0 to the 24 covering the complete development of the embryo of *D. melanogaster*. Finally, GRNs obtained was be characterized and studied by graphlets based techniques, thus, being able to identify specific elements whose relationship with the rest of the network vary.