Gene Regulatory Networks behind simple patterning phenotypes: circuit analysis

Sebastian Espejo Agudelo

Contents

1 Introduction

During development, animal embryos rely on differential expression of some genes to create early patterns that define cell fate. This process of early pattering ends up giving rise to differentiated tissues and body structure in the adult. One of the most studied mechanisms of differentiation is the gene regulatory network (GRN) model (Fig. 1A) that interpret the concentration of molecules in the cellular environment, producing a clear pattern of gene expression (Fig. 1B).

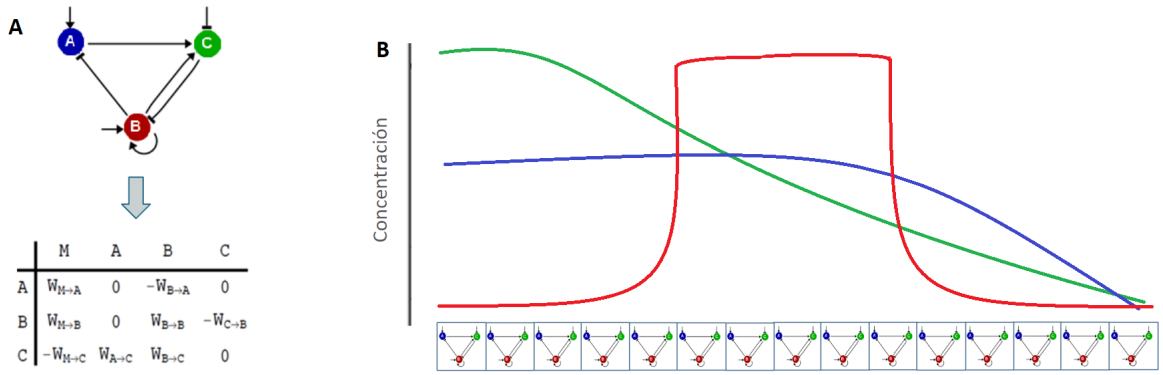


Figure 1. A) Exemplar of a gene regulatory network and its representation as a matrix of interactions, where negative interactions represent inhibitions, positive interactions represent activations and zeros, no interactions. Arrows coming from outside represent the interpretation of the external molecule B) Representation of the model. The array of GRN represent a tissue and one of the genes (B:red) only expresses in the middle of it as a result of the gene interactions.

All the GRNS were obtained by Arboleda et al with the following conditions: they are composed in 3 genes, they are embedded in an array of identical circuits representing an embryonic tissue (Fig 1B.), at least one of the genes is able to receive an external signal and, at least one of the genes is going to express only in the middle of the tissue. They used an evolutionary algorithm that recreates events of mutation and replication to obtain a population of 2061 GRNs able to complete this phenotypical task. The evolutionary algorithm also selected parameters of diffusion and degradation of this gene products that are disregarded in the current study. However, these parameters are necessary to create the dynamical system of gene expression that leads to the stationary levels of expression of the three genes represented in Fig 1B.

The purpose of this project is to analyze this large population of GRNS in terms of diversity, complexity and similarity. All the analyses were done in R, and GRN representation was done in Mathematica.

2 Data set up

Data was obtained from the author Github repository. Every GRN was defined by 18 parameters, 12 parameters of interaction between elements of the circuit (see Fig 1A) and, 6 parameters of diffusion and degradation. The latter set of parameters were disregarded for all the GRNs.

Important elements were built for the following analysis, including the data frame of topologies, being topology the type of interaction (inhibition or activation) of the GRNS without considering specific values of interaction. Also, the matrix of unique topologies, since the evolutionary algorithm obtained same topologies with different interaction strength.

3 Analysis

3.1 Topological considerations

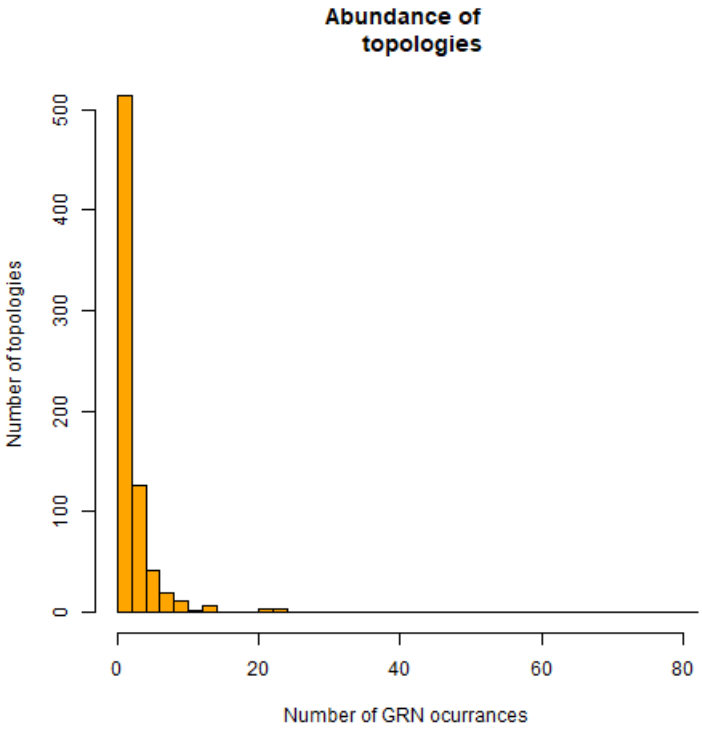
Every topology is represented several times in the original population because for every topology there are up to six equal configurations or permutations that were not considered as equal by the evolutionary algorithm that produce the original population. Fig. 2 shows an example of two different circuits with the same topology.

Fig 2. Example of two GRNs with identical topology

In "02 topological considerations.R" this is solved by comparing the population of topologies with the possible permutations and then generating a data frame with the single topologies without any of its permutations.

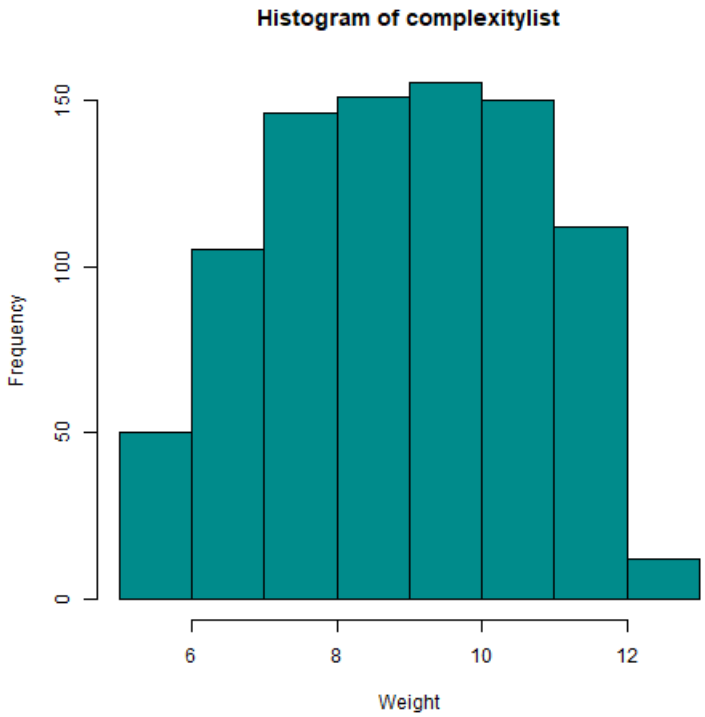
3.2 Abundance analysis

To understand how well represented are every type of topology in the sample of GRNs, I created an abundance



histogram where

3.3 Complexity analysis



3.3.1 Similarity analysis

[].