# Reservoir Computing capabilities of Gene Regulatory Networks

Marçal Gabaldà-Sagarra, Jordi Garcia-Ojalvo[*]
*Dynamical Systems Biology group, Department of Experimental and Health Sciences*
*Pompeu Fabra University (UPF), Barcelona*

Single cells have the potential and the necessity to process the information they receive from their environment. The interaction of the elements of their regulatory networks allows them to integrate different inputs and, even more, process temporal information. The global logic that these networks follow to perform such calculations is not well known yet.

Our hypothesis is that regulatory networks are able to function as a Reservoir Computing (RC) system. RC is based on a central recurrent reservoir of hidden nodes, analogous to a Recurrent Neural Network (RNN), with fading memory. In contrast to standard RNN, however, the output nodes are placed one or more linear layers of nodes that read the state of the recurrent part of the network in a strictly feed-forward manner[2]. The recurrences in the network would enhance its computational potential in general, and specially allow the processing of temporal information. On the other hand, the specific RC architecture would facilitate the learning of new environmental conditioning. Since, only the connections between the reservoir and the output layer need to be trained to learn a new task, it is easier to change a given behaviour without modifying the comportment of the whole network. This is especially relevant when the training happens through an evolutionary process and thus the weights are changed randomly. All in all, the approach based on RC simplifies notably the training of the RNN and makes it more meaningful as a biological model.

We analyzed gene regulatory networks already available in the bibliography, and found that their topologies are perfectly compatible with a RC functionality. The nodes in the networks are organised in three groups: a recurrent group, a small number of nodes upstream of the recurrent part and finally a major group of nodes downstream of the recurrent structure, which follow a strictly feed-forward organisation.
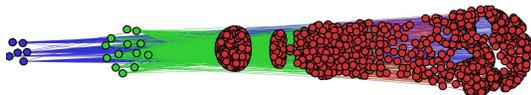


FIG. 1. Structure of the gene regulatory network of *Bacillus subtilis*. Green nodes form part of a recurrent structure or reservoir, while blue and red nodes form strictly feed-forward structures, upstream and downstream of the reservoir, respectively.

Further support of our hypothesis — i.e. that the regulatory networks adhere to a RC paradigm — was obtained confronting the systems to a prediction task. The dataset used in our prediction task benchmark was given by the *Nonlinear AutoRegressive Moving Average* (NARMA) function[1].

In general, modelling this system is difficult, due to the nonlinearity and potentially long memory.

The topologies of the recurrent parts of the of the gene regulatory networks analysed proved to be able to perform calculations comparable to other topologies that are known to perform well in RC, for a similar number of nodes.

Presumably, the differences in performance between the tested networks would be explained by the different number of nodes and cycles. A larger recurrent network would have more complex transient dynamics and a longer fading memory[2].

On the other hand, the dynamics of a gene acting as node are not as simple. In the first place, it can be activated to any degree of a continuous range of levels of expression between zero and an asymptotic maximum. Additionally, the biological processes that underlie expression —from the start of the transcription to the end of the translation— will introduce a delay in the activation of the nodes. These more complex dynamics of the genes can change the temporal dynamics of the reservoir, for example making its fading memory last longer. On top of that, it is important to note that the processing tasks of a biological network are remarkably different in nature and intensity to those addressed by non-standard computations in machine learning. Furthermore, in the former case results do not need to be as stably precise as in the latter.

These results suggest that the regulatory network core may be able to generate transient dynamics that encode the information obtained in the recent past from multiple inputs. Furthermore, plenty of nodes found downstream of the recurrent core, wich would function as output nodes, are reported to be involved in processes that are observable and meaningful as a learning layer of a RC. Some examples of those processes are biofilm formation, flagellum synthesis, protection against oxidative stresses, growth rate control, etc.

[*] jordi.g.ojalvo@upf.edu

[1] Atiya, A. F. & Parlos, A. G. (2000) IEEE Trans Neural Networks 11(3), 697-709

[2] Buonomano, D. V. & Maass, W. (2009) Nat Rev Neurosci 10(2), 113-25