

Evolution of correlation between traits in a sexually recombining population within a single selective environment

Simon Pirkel, Andreas Rohrbach

June 2020



Department of Bioengineering Sciences, Bioprocess Informatics

Contents

1 Abstract	2
2 Introduction	2
3 Methods	2
3.1 Selective environment	2
3.2 Gene regulation network	2
3.3 Genotype and phenotype	3
3.4 Ontogenetic development	3
3.5 Selection and recombination	3
4 Results	4
5 Analysis and Discussion	4
6 Appendix	5

1 Abstract

Watson et al. investigated the evolution of phenotypic correlations and "developmental memory". In one of their experiments, an evolutionary model was simulated, containing a single representative individual undergoing individual mutations. This simplification is legit under the assumption of SSWM (strong selection, weak mutation). To verify their results in a more realistic model of evolution involving sexual recombination, the experiment was replicated with a more sophisticated fitness evaluation and simple sexual recombination.

As in the experiment 1 in the work of Watson et al., the evolved gene regulation networks showed striking similarities to the matrix postulated using Hebb's rule.

2 Introduction

The phenotypic variants we observe are not purely a product of their genes, but of multiple factors, with developmental processes being one of them [1]. In the work of Watson et al., the evolution of a network of recurrent nonlinear ontogenetic interactions was investigated [2]. Such a network could represent a gene regulation network. In their work, multiple experiments were conducted to better understand the evolution of such networks. In the first experiment, the basic effect of selection on interaction coefficients as a function of a single selective environment was assessed. This experiment was conducted under the assumptions of SSWM (strong selection, weak mutation), under which it is sufficient to model the evolution of just one representative individual undergoing individual mutations.

One of Watson et al.'s results is that correlations in their single gene regulation network evolve according to Hebb's rule [3], a simple associative neural learning mechanism, often summarized as "Cells that fire together wire together" [4]. Translated to genes and traits, it would state: if evolution selects a correlation of multiple gene states, those gene states will also become developmentally correlated (see also Pavlicev et al. 2011 [5]). This could be summarized as "traits that are selected together correlate together" [2].

We aim to verify their findings in a population context with sexual recombination. Since mutation and sexual recombination generate greater genetic variation, this could lead to correlations evolving not in accordance with Hebb's rule. However, we postulate that Hebb's rule is driven not by mutation and recombination, but by selection. If we reproduce Watson et al.'s experiment 1 (single selective environment) for a network population with mutation and recombination, and our adaptive theory is correct, we predict that Hebb's rule will apply to the evolved networks.

3 Methods

The population is modeled with individuals defined by their genotype, phenotype and gene regulation network. The evolution process is simulated by generating a new set of individuals through sexual recombination based on their previously developed adult phenotype, which is used to calculate the fitness. Furthermore the recombined individuals experiences a mutation before developing to their adult state.

3.1 Selective environment

The selective environment is modeled as in the work of Watson et al. as a column vector S with $nTrait$ elements. Each element contains a value of the set $\{-1; 1\}$ which specifies the optimum trait expression level.

3.2 Gene regulation network

As in the work of Watson et al., the gene regulation network (grn) is modeled as a two dimensional matrix B , containing $nTrait * nTrait$ elements, with $nTrait$ being the number of genes and the number of corresponding traits as well. Each element of the matrix contains the correlation coefficient between two genes/traits, and is in the range $[-1; +1]$, where 0 corresponds to no correlation at all, a negative

coefficient indicating anticorrelation (or negative correlation), and a positive coefficient indicating positive correlation.

For the evaluation of the evolved gene regulation networks, the matrices were compared to the Hebb matrix H calculated with Hebb's rule. the elements of the Hebb matrix are derived from the selective environment S , and would normally contain only entries of the set $\{-1; +1\}$, but is scaled with a factor r , which represents the mean magnitude of the correlation coefficients.

$$H_{ij} = r * S_i * S_j \quad (1)$$

3.3 Genotype and phenotype

The genotype G and phenotype P are represented by column vectors with $nTrait$ elements. Each element contains a value in the range $[-1; 1]$ which represents the expression of the traits.

Each generation, for every individual, the genotype and gene regulation network is mutated by adding a value in a predefined range to one element. The same mutational parameters as in the work of Watson et al. were used. The genotype is mutating faster within the range $[-0.1; 0.1]$ than the gene regulation network, which mutational increment is 1/15th of the genotype mutation range $([-0.0067; 0.0067])$.

3.4 Ontogenetic development

In every generation the phenotypes have to be developed through time, influenced by the gene regulation network. The starting point is defined by the genotype: $P(0) = G$.

After defining the initial phenotype, the ontogenetic development of individuals is modeled by simulating through developmental time steps (set at $nDevSteps = 10$). In each timestep the phenotype experiences the influence of the gene regulation network, calculated by the same formula as used by Watson et al.:

$$P(t+1) = P(t) + \tau_1 \sigma(B * P(t)) - \tau_2 P(t) \quad (2)$$

As in the work of Watson et al., the sigmoid function of choice was the \tanh function.

3.5 Selection and recombination

Under the assumptions of strong selection, weak mutation, it was sufficient for Watson et al. to implement a fairly simple fitness evaluation of the phenotypes ($fitness = 1 + P * S$), since only one locus at a time is polymorphic, and either the current mutation is accepted, or not.

For a simulation in a population context, this dichotomous evaluation is not sufficient and a slightly more sophisticated fitness evaluation is necessary. The adult phenotypes are evaluated according to their similarity or deviation from the "target" phenotype S , representing the selective environment:

$$eval = \sum_{i=1}^{nTrait} (P_i - S_i) * (P_i - S_i) \quad (3)$$

This standard deviation of this evaluation is then used to calculate the fitness values of the individuals:

$$fitness = 1 + \frac{mean(eval) - eval}{std(eval)} \quad (4)$$

This column vector containing the fitness values for each individual is then normalised. Based on this normalised fitness vector, the parental individuals are selected through Stochastic Universal Sampling, as described by Baker (1987) [6].

The selected individuals are then randomly divided into maternal and paternal individuals needed for sexual recombination. The genotype and the gene regulation network of corresponding parents are then crossed over (with crossover probability = 0.4) to generate the genotype and gene regulation network of the following generation. The crossover point in the genotype corresponds to the crossover point of the gene regulation network.

4 Results

For the parameters $nGen = 500$, $nPop = 20$, and $nTrait = 8$, the simulation produced the result in figure 1 (can be found in the appendix).

5 Analysis and Discussion

As it could be seen in fig. 1, the evolved gene regulation networks have a striking similarity to the Hebb matrix. In our view, this confirms the results of Watson et al.'s experiment 1 in a more realistic model of evolution involving sexual recombination.

It should be noted, that the magnitudes of the correlation coefficients are significantly lower (approx. $1/10$) than in the results of Watson et al.. A reason for this is probably the reduced number of generations in this simulation ($5 * 10^2$, compared to Watson et al.: $2 * 10^5$). The decision to run the simulation over less generations was made because it was observed that the quality of the gene regulation networks (similarity to Hebb matrix) decreased after a certain amount of generations (approx. 500 - 600, depending on population size).

When investigating the reasons for this decrease of quality, it was found that the developing individuals reached maximum trait expression levels $\{-1; +1\}$ already before being fully developed (at approx. generation 600 with $nPop = 100$). Since mutation and sexual recombination in a population context generate greater genetic variation than in Watson et al.'s setting with one representative individual, fit individuals emerge quicker. But not only fitter individuals emerge quicker, but also "fit" genotypes, so that the influence of the gene regulation network on the fitness could be considerably less relevant than in the model of Watson et al.. Therefore a direct comparison of results for the same number of generations is not sensible. A possible way to further increase the quality of the evolved gene regulation networks could be increasing the selective pressure acting on the population.

Furthermore the choice of the relative mutational increment range of genotype and gene regulation network mutations ($1/15$) from Watson et al. seems to be tuned to their specific experiment design. The validity to choose the same parameters for a sexual recombination model is therefore questionable.

6 Appendix

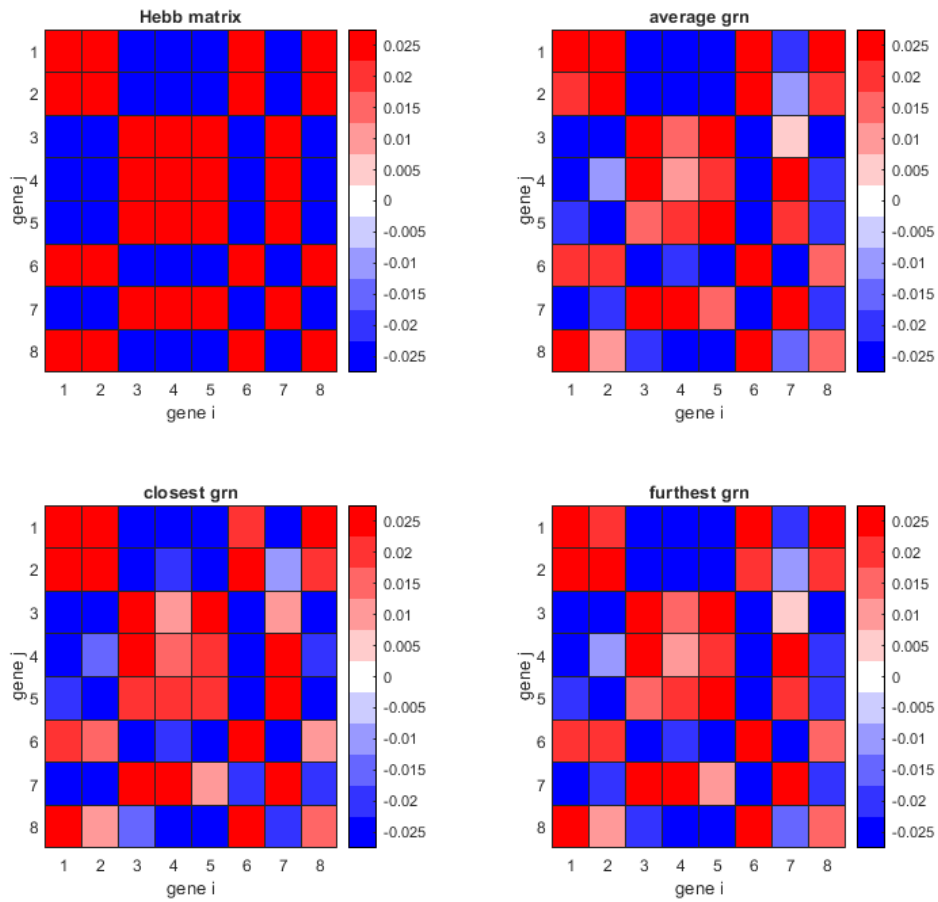


Figure 1: Representation of gene regulation networks. Top left: Matrix calculated through Hebb's rule. Top right: Average gene regulation network, arithmetic mean of all gene regulation networks. Bottom left: Closest gene regulation network, most similar to Hebb matrix. Bottom right: Furthest gene regulation network, least similar to Hebb matrix.

References

- [1] K. N. Laland, T. Uller, M. W. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka, and J. Odling-Smee, "The extended evolutionary synthesis: its structure, assumptions and predictions.," *Proceedings of The Royal Society B: Biological Sciences*, vol. 282, no. 1813, pp. 20151019–20151019, 2015.
- [2] R. A. Watson, G. P. Wagner, M. Pavlicev, D. M. Weinreich, and R. Mills, "The evolution of phenotypic correlations and developmental memory:," *Evolution*, vol. 68, no. 4, pp. 1124–1138, 2014.
- [3] "Hebb, d. o. the organization of behavior: A neuropsychological theory. new york: John wiley and sons, inc., 1949. 335 p.," *Science Education*, vol. 34, no. 5, pp. 336–337, 1950.
- [4] C. J. Shatz, "The developing brain.," *Scientific American*, vol. 267, no. 3, pp. 60–67, 1992.

- [5] M. Pavlicev, J. M. Cheverud, and G. P. Wagner, "Evolution of adaptive phenotypic variation patterns by direct selection for evolvability," *Proceedings of The Royal Society B: Biological Sciences*, vol. 278, no. 1713, pp. 1903–1912, 2011.
- [6] J. E. Baker, "Reducing bias and inefficiency in the selection algorithm," in *Proceedings of the Second International Conference on Genetic Algorithms and their application*, pp. 14–21, 1987.