

Evolution of Pleiotropy and Epistasis in a Gene Regulatory Network

Priyanka Mehra¹ and Arend Hintze^{1,2}

¹Dalarna University, Department of MicroData Analytics, Dalarna, 79188, Sweden

²Michigan State University, BEACON Center for the Study of Evolution in Action, East Lansing, 48824, United States of America
ahz@du.se

Abstract

The study of gene regulatory networks (GRNs) is fundamental to the understanding of evolutionary dynamics and artificial life modeling. This paper presents an integration of a GRN into the NK-fitness landscape model and explores the impact of sparsity on epistasis and pleiotropy. As sparsity augments, gene interactions diminish, expectedly leading to a reduction in both epistasis and pleiotropy. Our findings corroborate the model's response to such perturbations, demonstrating its potential for investigating a range of GRN adaptations within the NK-fitness landscape framework.

Keywords: epistasis, pleiotropy, gene regulatory network, evolution.

Introduction

Computational models have demonstrated their value in the study of evolutionary dynamics (Nowak and Sigmund, 2004; Hallinan and Wiles, 2004). In particular, the NK-fitness model has established its utility due to its feature that allows the researcher to adjust the ruggedness of the landscape (Altenberg, 1996), making it an effective benchmark for evaluating the efficiency of genetic algorithms (Hill and O'Riordan, 2001; Merz and Freisleben, 1998). This model utilizes a simple vector of binary numbers that evolves due to mutations. The NK landscape model enables rapid fitness determination for each vector, facilitating their adaptation within the virtual fitness landscape. The original model uses the factor K to set the ruggedness; however, it simultaneously determines the degree of pleiotropy and epistasis. Essentially, the vector provides a method of directly encoding a phenotype (Pitzer and Affenzeller, 2012), although the model can be easily modified to incorporate indirect encoding (Mehra and Hintze, 2022b). Previous versions substituted the phenotype vector with a genotype vector and a mapping matrix, the dot product of which forms the phenotype vector. This modification allows pleiotropy and epistasis to evolve independently through mutations in the mapping matrix.

Epistasis (ϵ) quantifies the extent of gene interaction, while pleiotropy (π) represents the number of functions a gene participates in. The modified, indirectly encoded

model, enables the study of the evolution of ϵ and π for individual genes and for the entire system (Mehra and Hintze, 2022b). A variety of indirect encodings can be conceived (Ronald, 1997; Kassahun et al., 2007), such as NEAT and HyperNEAT encoding for artificial neural networks (Stanley and Miikkulainen, 2002; Gauci and Stanley, 2010; Stanley and Miikkulainen, 2004; Meli et al., 2021), grammars (Miller et al., 1989), models based on L-systems (Jacob et al., 1994; Gruau, 1994; Lee et al., 2005; de Campos et al., 2015), and other alternatives like developmental encodings (Harding and Miller, 2006; Komosiński and Rotaru-Varga, 2001).

However, our focus here is on an indirect encoding grounded in gene regulatory networks, a concept found in biology. Previous studies have explored gene regulatory networks (Spirov and Holloway, 2016; Cussat-Blanc et al., 2019), assessing their specific levels of ϵ or π in natural organisms (Cussat-Blanc et al., 2015), or how they adapt over time to particular landscapes (Fragata et al., 2019). Therefore, this paper presents preliminary results on the incorporation of a basic gene regulatory network into the NK-fitness landscape framework. We also demonstrate the evolution of ϵ or π in relation to the sparsity of the gene regulatory interactions.

Gene regulatory networks

Gene regulatory network was defined by a mapping matrix C which contains a single number (between -1.0 and 1.0) describing the strength of interaction for all N genes. Gene expression levels (G_i) start at $\frac{1}{N}$ for each gene (i), and are updated for 100 discrete time steps as follows:

$$q_i = \sum_{j=0}^N G_j C_{i,j} \quad (1)$$

$$Q_i = \frac{2.0}{1.0 + e^{-\alpha q_i}} - 1.0 \quad (2)$$

$$G'_i = G_i - \beta(G_i - Q_i) \quad (3)$$

The parameters $\alpha = 1.0$ define the temperature of the

system, $\beta = 0.2$ is a dampening factor, q_i is the binary phenotype vector, and $N = 20$ is the genome size.

Expression levels are binned, with values above 0.0 becoming 1 and others 0, forming the phenotype vector. Regular selection follows, based on the NK-fitness landscape (Mehra and Hintze, 2022b), with point mutations at a rate of 0.001 per site. Mutated sites take a new value from a uniform distribution $[-1, 1]$. Sparsity is introduced through a probability S , setting seeding or mutating matrix values to 0.0 or uniform $[-1, 1]$ with $1 - S$. For each parameter K (0, 1, 3, 5, and 7) and sparsity S (0.0, 0.5, 0.8, 0.9, 0.95, and 0.99), 500 replicates were run for 50,000 generations. The representative organism was selected after 45,000 generations from a reconstructed line of descent Mehra and Hintze (2022b).

Each gene's contribution to each phenotypic trait was determined by setting expression levels to -1.0 and running 100 updates of the GRN. An interaction matrix was derived from these results, indicating contribution (1) or no interaction (0). This matrix quantifies ϵ and π (Mehra and Hintze, 2022b).

Results

When comparing evolved GRNs and their degree of ϵ and π with those from a random (unevolved) GRN, we make two observations. First, the sparsity of the GRN dominates the effect of K for ϵ and π alike (see Figures 1 and 2). With an increase in sparsity, the mean degree of ϵ and π drop (see Figure 3). However, while ϵ is reduced consequently, we find a small increase in π for the sparsest networks.

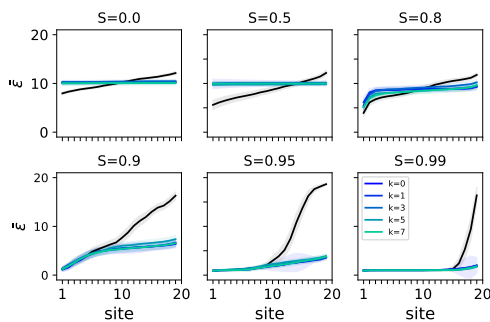


Figure 1: Epistasis in differently rugged landscapes. Average epistasis (y-axis) for all genes (x-axis, sorted by epistasis). Results from different K s are shown in different blue tones (see legend). 95% confidence intervals are shown as shadows. The black line is the random expectation.

At the same time, we would have expected ϵ and π to also drop with an increase in K (Mehra and Hintze, 2022b), but that effect can not be reconstructed for all levels of S .

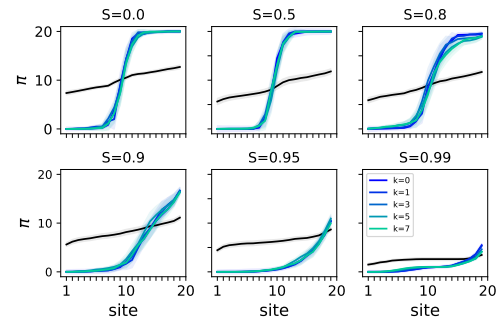


Figure 2: Pleiotropy in differently rugged landscapes. Average pleiotropy (y-axis) for all genes (x-axis, sorted by epistasis). Results from different K s are shown in different blue tones (see legend). 95% confidence intervals are shown as shadows. The black line is the random expectation.

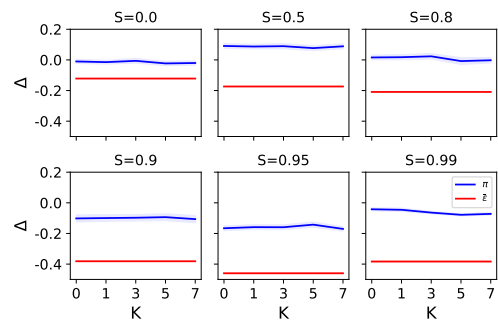


Figure 3: Distance (Δ y-axis) between the measured epistasis (ϵ red) or pleiotropy (π blue) for different K s (x-axis) and the random expectation (see Mehra and Hintze (2022b) for the distance measure). Shadows indicate the 95% confidence interval respectively.

Discussion

The integration of a GRN into the NK-fitness landscape as a form of indirect encoding was straightforward and computationally efficient. However, analyzing ϵ and π through a perturbation analysis requires testing all genes, which introduces additional computational cost. We examined the impact of sparsity on the evolution of ϵ and π in a GRN.

As anticipated, an increase in sparsity led to fewer gene interactions, enhancing their independence. A fully connected matrix results in extensive gene interaction, yielding high ϵ and π . The model responded appropriately to sparsity-induced perturbation, showing a decrease in ϵ and π . The reason for the somewhat reduced ϵ and π with increasing K , as seen in other indirect encoding models (Mehra and Hintze, 2022b), remains unclear. Replicating this phenomenon in another model strengthens the validity of previous findings. Future work includes exploring additional GRN adaptations within the NK-fitness landscape (Mehra and Hintze, 2022a).

References

- Altenberg, L. (1996). B2. 7.2 nk fitness landscapes. *Evolution*, 2.
- Cussat-Blanc, S., Harrington, K., and Banzhaf, W. (2019). Artificial gene regulatory networks—a review. *Artificial life*, 24(4):296–328.
- Cussat-Blanc, S., Harrington, K., and Pollack, J. (2015). Gene regulatory network evolution through augmenting topologies. *IEEE Transactions on Evolutionary Computation*, 19(6):823–837.
- de Campos, L. M. L., de Oliveira, R. C. L., and Roisenberg, M. (2015). Evolving artificial neural networks through l-system and evolutionary computation. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9. IEEE.
- Fragata, I., Blanckaert, A., Louro, M. A. D., Liberles, D. A., and Bank, C. (2019). Evolution in the light of fitness landscape theory. *Trends in ecology & evolution*, 34(1):69–82.
- Gauci, J. and Stanley, K. O. (2010). Indirect encoding of neural networks for scalable go. In *Parallel Problem Solving from Nature, PPSN XI: 11th International Conference, Kraków, Poland, September 11-15, 2010, Proceedings, Part I 1*, pages 354–363. Springer.
- Gruau, F. (1994). Automatic definition of modular neural networks. *Adaptive behavior*, 3(2):151–183.
- Hallinan, J. and Wiles, J. (2004). Evolving genetic regulatory networks using an artificial genome. In *Proceedings of the second conference on Asia-Pacific bioinformatics-Volume 29*, pages 291–296. Citeseer.
- Harding, S. and Miller, J. F. (2006). The dead state: A comparison between developmental and direct encodings. In *Proc. GECCO Workshop on Complexity Through Development and Self-Organizing Representations*.
- Hill, S. and O’Riordan, C. (2001). Genetic algorithms, their operators and the nk model. *Preprint, National University of Ireland, Galway*.
- Jacob, C., Lindenmayer, A., and Rozenberg, G. (1994). Genetic l-system programming. In *PPSN*, pages 334–343.
- Kassahun, Y., Edgington, M., Metzen, J. H., Sommer, G., and Kirchner, F. (2007). A common genetic encoding for both direct and indirect encodings of networks. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 1029–1036.
- Komosiński, M. and Rotaru-Varga, A. (2001). Comparison of different genotype encodings for simulated three-dimensional agents. *Artificial Life*, 7(4):395–418.
- Lee, D.-W., Kong, S. G., and Sim, K.-B. (2005). Evolvable neural networks based on developmental models for mobile robot navigation. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, volume 1, pages 337–342. IEEE.
- Mehra, P. and Hintze, A. (2022a). Evolutionary dynamics in the nk treadmill fitness landscape. *Available at SSRN 4209350*.
- Mehra, P. and Hintze, A. (2022b). An extension to the nk fitness landscape model to study pleiotropy, epistasis, and ruggedness independently. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1259–1267. IEEE.
- Meli, C., Nezval, V., Oplatkova, Z. K., Buttigieg, V., and Staines, A. S. (2021). A study of direct and indirect encoding in phenotype-genotype relationships. In *Artificial Intelligence and Soft Computing: 20th International Conference, ICAISC 2021, Virtual Event, June 21–23, 2021, Proceedings, Part II 20*, pages 290–301. Springer.
- Merz, P. and Freisleben, B. (1998). On the effectiveness of evolutionary search in high-dimensional nk-landscapes. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98TH8360)*, pages 741–745. IEEE.
- Miller, G. F., Todd, P. M., and Hegde, S. U. (1989). Designing neural networks using genetic algorithms. In *ICGA*, volume 89, pages 379–384.
- Nowak, M. A. and Sigmund, K. (2004). Evolutionary dynamics of biological games. *science*, 303(5659):793–799.
- Pitzer, E. and Affenzeller, M. (2012). A comprehensive survey on fitness landscape analysis. *Recent advances in intelligent engineering systems*, pages 161–191.
- Ronald, S. (1997). Robust encodings in genetic algorithms: A survey of encoding issues. In *Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC’97)*, pages 43–48. IEEE.
- Spirov, A. and Holloway, D. (2016). Using evolutionary algorithms to study the evolution of gene regulatory networks controlling biological development. *Evolutionary Computation in Gene Regulatory Network Research*, pages 240–268.
- Stanley, K. O. and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evolutionary computation*, 10(2):99–127.
- Stanley, K. O. and Miikkulainen, R. (2004). Competitive coevolution through evolutionary complexification. *Journal of artificial intelligence research*, 21:63–100.