Modelling Evolution of Gene Networks

Francisco Quevedo Camargo, Ard Louis

July 4, 2014

Abstract

Ever since the dawn of the molecular revolution in biology, it has been knowo that the number of genotypes in a given system usually exceeds the number of correspondent phenotypes by orders of magnitude. One naturally wonders how these are distributed: if all phenotypes are obtained by approximately the same fraction of genotypes, or if ones take a bigger share than others. Recent evidence suggests that this division is highly biased - that is, some phenotypes cluster a high proportion of genotypes.

It has been known for quite some time that there are small mutational biases, that is certain types of mutations are more likely to occur than others, but only more recently it was shown that some phenotypes can be orders of magnitude more likely to arise from random mutations than other phenotypes.

In this project, this idea was explored in the context of gene networks. We find differences of many orders of magnitude in models for the cell-cycle, and we relate this bias to concepts from Algorithmic Information Theory, and linking easily obtainable phenotypes to low complexity. We also study a simple model of mutational bias, and verify how random variation can outweigh natural selection.

1 Introduction

In papers published as early as 1932 [28], it has become clear that evolution explores hyper-astronomically large spaces. As a simple example, considering the average protein is composed of approximately 470 amino acids, each one of them being one of 20 possible aminoacids, the number of possible proteins with that length would be $20^{470} \approx 10^{600}$ possible proteins. As a comparison, 10^{80} is the estimated number of atoms in the visible universe. From that perspective, it would seem very unlikely for mutation to have any role in evolution, simply becaiuse because the spaces it explores would be too vast. This is sometimes called Hoyle's paradox [15].

One should bear in mind then that many genotypes usually express the same phenotype - a fact that is named in different publications as a phenotype's *abundance* [6], *degeneracy level* [2], *genotype set size* [25] or *neutral network size* [26],[27]. Since different phenotypes are expressed by different numbers of genotypes, many mutations can yield the same effect. This bias has several effects, such as evolvability (how quickly an organism can adapt) and robustness (how much mutation it can take without changing its phenotype) [27]. A population with a single robust phenotype can, therefore, have a diverse gene pool, hence being able to reach a high number of different phenotypes.

1.1 Bias and complexity

A simple and clear example of this bias is the degeneracy in the codon-aminoacid map. What Francis Crick described initially as a "frozen accident" [7] does not seem to be frozen, nor an accident: not only does the genetic code seem to be somehow optimised, but it is also still evolving [14], [22]-[16]. Even though it is not a big bias, the code degeneracy is correlated with amino acid abundancy in genomes. It is natural to ask which came first - the bias in the codon-aminoacid map or the bias in their abundance in genomes? Did the proteins adapt to the genotype-phenotype (GP) map or did the GP map change as genes required more or less of that aminoacid? There is a long discussion on that issue [11], and stronger arguments seem to support the first view: that the different frequencies of codons are a result of how some of them are obtained by a higher number of codons than others. Although this discrepancy in number is not strong (as this number of codons goes only from 1 to 6), it is quite remarkable that such a small bias has an effect on the codon distribution.

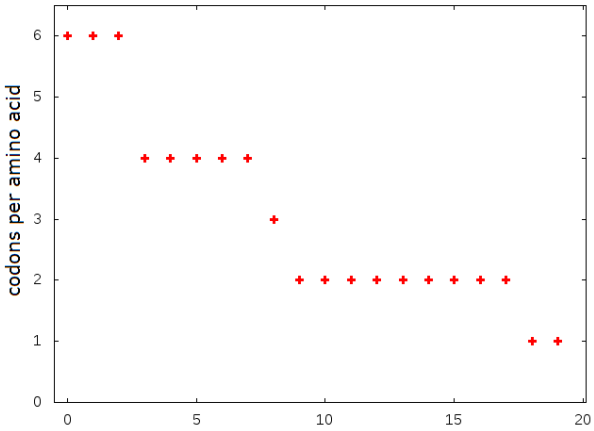

Figure 1: The different numbers of codons corresponding to each amino acid shows bias in the genetic code.

Another system with a visible bias is the mapping between RNA single strands and their secondary structures. Just like in the protein example, the number of possible strands grows exponentially with their length, while the number of phenotypes does not grow that much. For example, for RNA molecules of 20 bases, the $4^{20} \approx 10^{12}$ possible RNA strands correspond to 11219 different RNA secondary structures, some of which are obtained by 10^{10} sequences, while others are reached only by one or two. The GSS (genotype set size) bias is clear and evident, and the structures present in nature seem to have a higher than average GSS [10] - that suggests that their existence would be justified not only by a higher fitness of that secondary structure, but also by the fact that it would be easier to be obtained. In fact, it suggests that maybe selection by itself is not that necessary for the fixation of a certain phenotype. This idea will be explored in further sections.

Despite being very interesting, the phenotypic bias is not the only bias present in the system. RNA strings can be represented by dots and brackets, the dots being the nucleotides that do not form any bonds and the (matching) brackets being the

ones that do, as in Fig. 2. Those dot-bracket strings can then be converted to binary strings that can have their Kolmogorov complexity measured by standard compression techniques such as the Lempel-Ziv algorithm for compression [19].

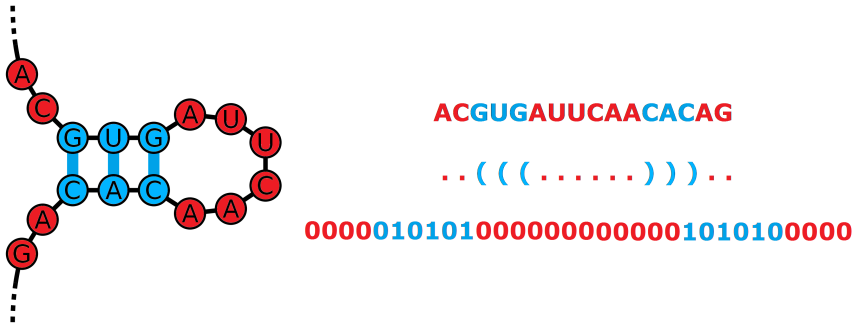

Figure 2: A RNA structure can be converted to a sequence of dots and brackets, and then to a binary string. By applying compression algorithms to measure that string's complexity, one can measure the complexity of a RNA structure.

When one applies this measure to all the secondary structures formed by the 20-bases-long RNA strands, an interesting trend appears: not only some RNA structures have a higher GSS than others, but those ones also seem to have a smaller complexity. That measure comes from a computation performed with the binary strings obtained as described, but it would relate to being simpler to the eye. In fact, this result relates to a theorem in algorithmic information theory described by Leonid Levin [20], expressed as the following:

$$P_U(x) = 2^{-K(x)+O(1)} \tag{1}$$

Levin's coding theorem states that, for a universal Turing machine U fed with random programs, the probability of an output x being generated decreases exponentially with its Kolmogorov complexity $K(x)$. It is true that the map between sequence and secondary structure is not a universal Turing machine, but the resemblance is striking, and it has been recently argued that the coding theorem can be generalised to systems like these GP maps [10]. The question of whether the most common phenotypes are also the less complex ones - in other words, the existence of a "simplicity bias" - can be posed to other systems, and the following section is devoted to that.

2 A simple model for mutational bias in population genetics

In section 1.1, it was mentioned that the predominance of a certain phenotype, other than being a mere result of its higher fitness, could be a byproduct of its GSS. One could then devise a situation in which two phenotypes compete for fixation, one with a higher fitness, the other one with a higher GSS. Yampolsky and Stoltzfus

[29] debate a toy model along those lines: given three genotypes, ab, Ab and aB with respective fitnesses 1, $1+s_1$ and $1+s_2$, with point mutations occurring from a to A with rate μ_1 and from b to B with rate μ_2, the ratio between the number of aB and Ab individuals should depend on μ_1, μ_2, s_1 and s_2. The authors of the paper, however, do not give the problem any analytical treatment. In this section, I apply some methods alternative to the ones used in their work.

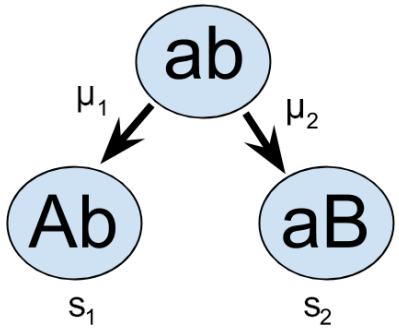

Figure 3: The Yampolsky-Stoltzfus toy model for mutational bias.

2.1 Markov chain

Assuming that N, the number of individuals in the population, is large, one can look at the population in discrete times, and treat the stochastic effects as transition rates. In that sense, the state of the system could be described by a vector $g(t) \in \mathbf{R}^3$ where g_0, g_1 and g_2 are, respectively, the numbers of individuals with the genotypes ab, Ab and aB, and the bias in the population is g_2/g_1 as the system evolves according to:

$$g_{t+1} = S \cdot M \cdot g_t \tag{2}$$

In the equation above, M represents the redistribution of individuals within g through mutation and S represents their increase in numbers through reproduction and selection. M and S are described below.

$$S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+s_1 & 0 \\ 0 & 0 & 1+s_2 \end{pmatrix} \qquad M = \begin{pmatrix} 1 & 0 & 0 \\ \mu_1 & 1 & 0 \\ \mu_2 & 0 & 1 \end{pmatrix} \tag{3}$$

Instead of iterating over many generations, one can look at a time-continuous version of the model: $\dot{(g)} = S \cdot M \cdot g$. The matrix SM can be diagonalized and the system becomes analytically solvable, with the following result: for an initial condition of $g(t=0) = [1\,0\,0\,0]^T$, that is, only ab individuals, $g(t)$ is:

$$g(t) = \begin{pmatrix} e^t \\ a_1[\, e^{(1+s1)t} - e^t \,] \\ a_2[\, e^{(1+s2)t} - e^t \,] \end{pmatrix}, \quad a_1 = \frac{\mu_1\,(1+s1)}{\mu_1+\mu_2+s_1}, \quad a_2 = \frac{\mu_2\,(1+s2)}{\mu_1+\mu_2+s_2} \tag{4}$$

where the bias $b(t)$ is

$$b(t) = \frac{a_2[e^{(1+s_2)t} - e^t\,]}{a_1[\,e^{(1+s_1)t} - e^t\,]} = \frac{\mu_2(1+s_2)(\mu_1+\mu_2+s_1)[\,e^{(1+s_2)t} - e^t\,]}{\mu_1(1+s_1)(\mu_1+\mu_2+s_2)[\,e^{(1+s_1)t} - e^t\,]} \tag{5}$$

or, asymptotically,

$$log(b) \approx log(\mu_2/\mu_1) + (s_2 - s_1)\,t \tag{6}$$

This, as shown in Fig. 4, matches the results in the Yampolsky-Stoltzfus paper: the s_1/s_2 ratio does not matter as much as the μ_1/μ_2 ratio. It is interesting to notice that, even though the difference in fitnesses has an exponential effect in the bias between genotypes, it is modulated by a ratio of the rates of introduction of those genotypes. Also, given enough time for the population to grow, the outcome is determined by selection, but if it does not reach the stage where the second term in the sum exceeds the first one, mutation will have a stronger impact.

One can also look at this model including the mutation rates from B to b and from A to a, but although tractable, the analytical solution of this model is not practical, due to the very nonlinear dependancy of the eigenvalues of SM on all parameters. In that case, solving the model numerically is probably the most practical choice.

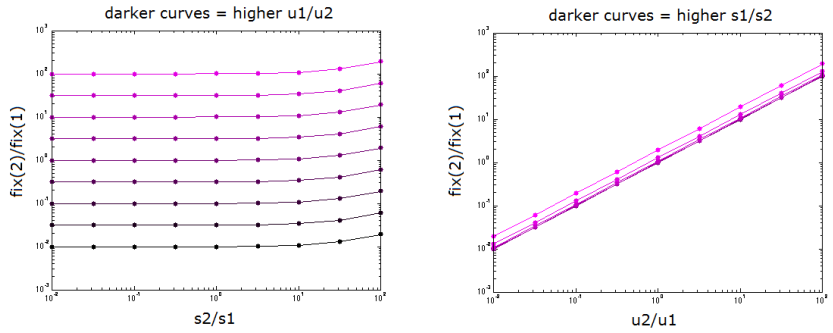

Figure 4: The plots above show the Markov chain implementation of the Yampolsky-Stoltzfus model. Both show that the μ_1/μ_2 ratio is the main parameter in determining the bias b between aB and Ab populations.

2.2 Allele fixation and Kimura equations

Although the previous model is able to correctly reproduce the mentioned results, its populations grow indefinitely. A more realistic approach would be to consider a limited number of individuals in the population, allowing some drift to interfere with the fixation of one or another genotype. Hence I developed a second implementation of the model, now considering a limited number of L sites filled initially by ab individuals, eventually mutating to Ab or aB. At each timestep, after mutations occur, one random individual is removed, and replaced it by a member of the

populations present in that time (i.e. allowing them to reproduce) with probability p_i proportional to their frequency and fitness:

$$p_i = \frac{w_i x_i}{\sum_i w_i x_i} \tag{7}$$

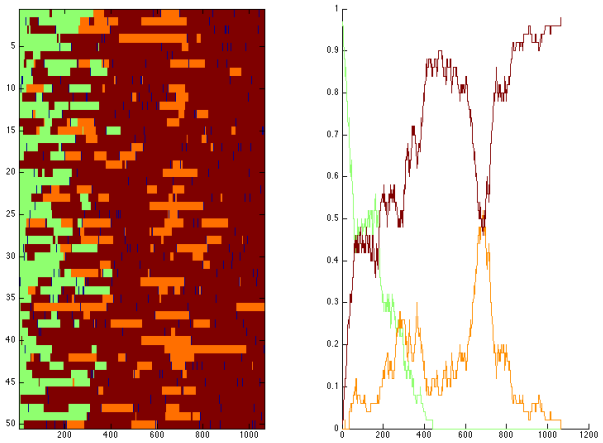

Figure 5: One run of the gene fixation implementation of the Y-S problem. Here, green individuals (ab) mutates to orange (Ab) and brown (aB) individuals, and they all suffer selection and drift. In this case, the $b \to B$ rate is higher, and the aB genotype is fixed.

The results obtained in the simulation of gene fixation (Fig. 5) are the same as with the unbounded populations, which suggests that other formalisms such as the diffusion equations employed by Motoo Kimura [18], more suited to gene fixation problems, should also apply here. These models describe the fixation of an allele as a Fokker-Planck equation on a simplex defined in this case by $x_0 + x_1 + x_2 = 1$, $x_i \in [0, 1]$, $i = \{0, 1, 2\}$. In this case, the resulting object is a triangle, and the case of only two mutually exclusive alleles, this object would be the interval $[0, 1]$. In this formalism, genetic drift is described as diffusion, and both selection and mutation come as advection terms. The final formulation of the problem is the following:

$$\frac{\partial P}{\partial t} = \frac{\kappa}{2} \sum_{ij} \partial_{ij}^2 [D_{ij}\, P] - \alpha \sum_i \partial_i [\Omega_i\, P] - \alpha \sum_i \partial_i [A_i\, P] \tag{8}$$

$$D_{ij} = x_i \left(\delta_{ij} - x_j\right)$$

$$A_i = \sum_j M_{ji} x_j - M_{ij} x_i$$

$$\Omega_i = \sum_j D_{ij} \psi^{(j)}(x)$$

In the equations above [18][4][3], $P(\mathbf{x}) = (x_0, x_1, x_2)$ is the p.d.f. for a population with allele frequencies x_0, x_1 and x_2, κ and D_{ij} describe the contribution from genetic drift, $\Omega_i(x)$ is the contribution of the fitness of each allele to the flux in the direction of x_i, and A_i is likewise for mutations. Each vertex of the triangle represents when a certain gene is fixed: $\mathbf{x} = \mathbf{e}_i$, that is, $x_j = \delta_{ij}$.

The strength this model has from being so general comes with a drawback: its analytical solutions are hardly practical. In his original text, Kimura [18] solves the equation for two alleles with diffusion only, obtaining a hypergeometrical function - that should be enough warning even for the bravest reader. Still, numerical simulations can provide good insight: since this model is written in the form of a conservation equation, it should be possible to discretize it in a way that it becomes a Markov chain with a stochastic matrix - that is, A such that all entries $A_{ij} \geq 0$, and $\sum_i A_{ij} = 1$.

A stochastic matrix has all of its eigenvalues on or in the unitary circle ($|\lambda| \leq 1$). Since a population concentrated at $\mathbf{x} = \mathbf{e}_i$, $i = \{1, 2\}$ should remain there, both $\delta(\mathbf{x} - \mathbf{e}_1)$ and $\delta(\mathbf{x} - \mathbf{e}_2)$ - or their equivalents, in the discretisation - should be eigenvectors with $\lambda = 1$. Hence, the projection of the initial condition $P(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{e}_0)$ on those eigenvectors should represent how much each one (x_1 and x_2) is going to grow. Since both eigenvectors have the same eigenvalue, the ratio between such projections should be the bias discussed previously.

A final approach to the same problem would be the one in Wang et al [30]: when genetic drift is weak, the problem is mostly deterministic, and a vector field can be drawn to describe the flow of the population towards the vertices of the simplex. Every one of the models described yields the same results, but coming from different simplifications. When the population is large enough (or the selection and mutation rates large enough) so that drift is negligible, a simple Markov chain as described in section 2.1 can account for the desired behaviour, while the more intractable models can provide more insight into the coupling of those different biological phenomena.

It is important to notice how many questions arise just from a bias in the introduction of variation, without any GP maps, gene networks, or interaction between populations. Still, one should be able to couple all of these in a single description.

3 Bias in models for the cell cycle

3.1 The Budding Yeast Cell Cycle Model

Mathematical descriptions of biological phenomena are proven to be a useful tool in their understanting - in particular, in molecular biology, they can suggest experiments or provide explanations for the ones already realised, just as they can present the scientist with new directions to follow. One celebrated model in this sense is the budding yeast cell cycle model: a collection of approximately 50 equations and 150 parameters that were all carefully parametrised to describe the cell cycle of *Saccharomyces cerevisae*, resulting in a robust model which relates faithfully to experiment [5]. A picture of the gene regulatory network described by the model is shown on Fig. 7.

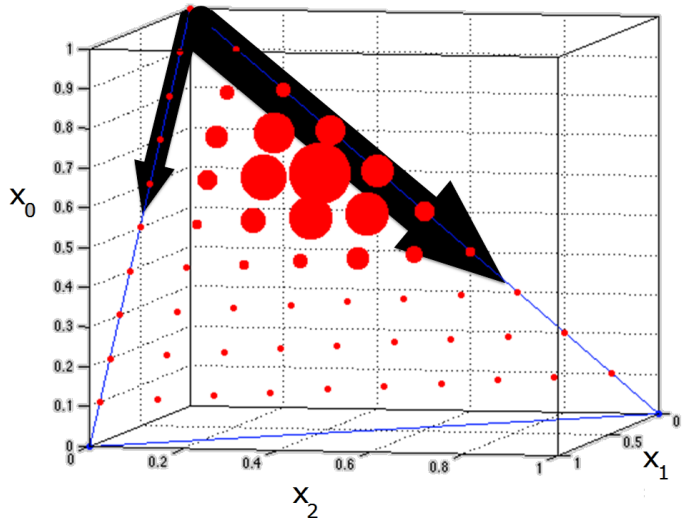

Figure 6: One run of the gene fixation Y-S problem implemented as the Kimura equation on the $x_0 + x_1 + x_2 = 1$ triangle. Larger circles represent higher population density, while the dark arrows indicate mutation, the $x_0 \to x_2$ mutation rate μ_2 being larger than the $x_0 \to x_1$ mutation rate μ_1.

3.2 The Generic Cell Cycle Model

Another model studied here is the generic cell cycle model [8]. Similar to the one described before, this model proposes to explain that differences between the cell cycle regulation networks from dfferent organisms are due only to the different parameters of those reactions, such as chemical reaction rates or saturation constants. Once those differences are removed, the cell cycle would exhibit the same behaviour both in the budding yeast and in the fission yeast, but also in mammalian cells. Built in a similar way to the budding yeast cell cycle model, it has a set of equations and parameters related to a wiring diagram (displayed in Fig. 8), and also presents robust phenotypes with precise descriptions of cellular physiology.

In order to study phenotypic biases, these models can be seen as maps from a set of parameters to a set of curves (the gene expression patterns). A coarse discretisation of these curves via the 'up-down' method [13] can turn the curves into binary strings, which a 1 in position k when the slope of the curve in time t_k is positive, and an 0 otherwise. Those strings can be then clustered, so one can then observe how many sets of parameters map to a specific string - that is, how many genotypes map to a phenotype. However, generating different genotypes in a way that the whole parameter space is covered is no easy task, specially if taken into consideration that it is not at all clear how mutations in the DNA would map to parameters in the model (which, by itself, is a very difficult task). The approach here was then similar to the one used to study phenotypic robustness in [5]: all the parameters were multiplied by a factor of $\sqrt{2^\gamma}$, where $\gamma \in \{0, \pm 1, \pm 2, \pm 3, \pm 4\}$.

The complexity of each binary coarse-grained phenotype then was measured by using a standard Lempel-Ziv compressing algorithm to measure their Kolmogorov

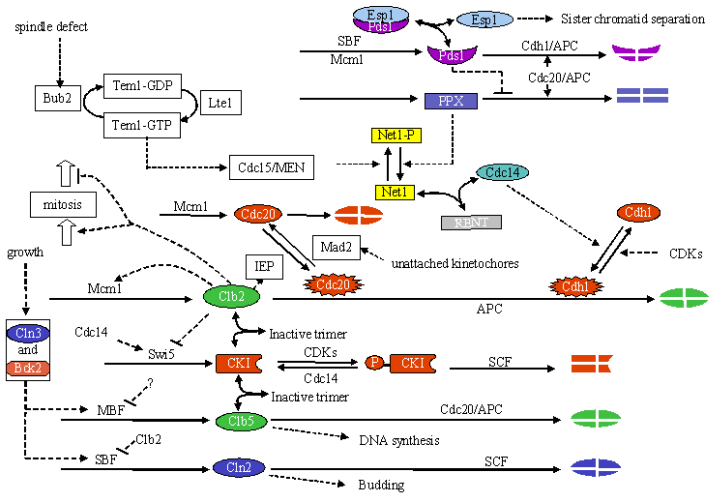

Figure 7: Wiring diagram of the budding yeast cell cycle model, from [5]. Its ~ 150 parameters can be seen as a genotype, where as the expression patterns of its ~ 50 variables can be seen as the phenotype. This mapping produces less complex outputs with a higher probability, as seen in Fig. 9.

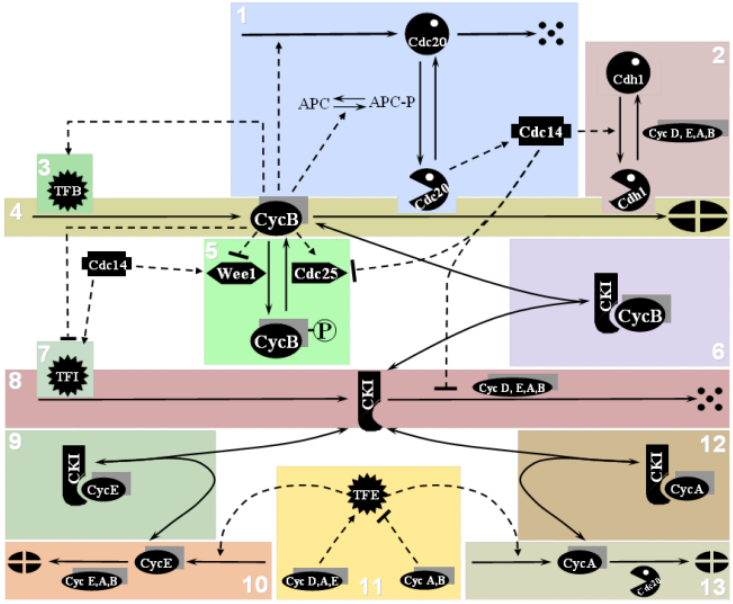

Figure 8: Wiring diagram of the generic yeast cell cycle model, from [8]. Just as in the budding yeast cell cycle model, it maps parameters in chemical reactions to protein expression patterns and other chemical variables, exhibiting bias towards less complex outputs, as in Fig. 10.

complexity [19]. As in Fig. 9, one can definitely see a trend: as the complexity of a phenotype increases, its genotype set size (GSS) decreases. It is worthy to note

the log-scale on the vertical axis, indicating a reproduction of the law derived from the $P(x) \propto 2^{-K(x)}$ law described above.

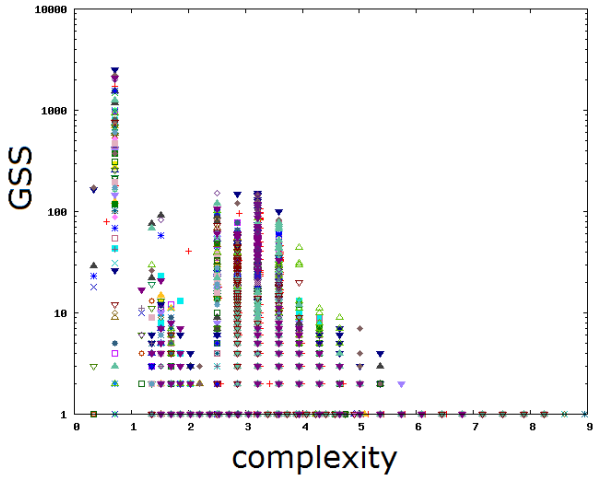

Figure 9: Scatter plot of GSS versus Kolmogorov complexity for the budding yeast cell cycle model. Different symbols represent the discretised expression patterns of all variables in the system. There is a clear trend for the upper bound of the GSS of every phenotype, following $GSS(x) \propto 2^{-K(x)}$.

When the same reasoning is applied to the generic cell cycle model, the same behaviour is observed: for different organisms, namely fission yeast, budding yeast and mammals, the least complex phenotypes show the highest GSS, whereas the more complex ones are harder to obtain. It is a small data set, even though 1 000 000 parameter sets were explored for each organism, but the results do point in the same direction as the budding yeast model: that complexity of a phenotype correlates inversely with the log of its GSS.

It is remarkable that the same trend is showed in maps governed by such different rules, such as physicochemical interactions for the RNA secondary structures, or mathematical rules and parameters as in the cell cycle models. These results support the idea that Levin's coding theorem can be adapted to maps other than universal Turing machines [10].

4 Evolution of gene networks

The phenotype robustness observed in the cell cycle models suggests that the parameter values might not be the most important factors in the dynamics of a gene network. In fact, many studies affirm that the connectivity pattern is much more important than actual rates and constants [24][12][1]. Given that, a genetic network can be reduced to a Boolean network model, as in [17]. These models assume that genes can be either "on" or "off", and therefore the output of the network at a given time t would be a string of ones and zeros. Likewise, interactions between genes can be either positive $(+1)$, negative (-1) or just non-existent (0).

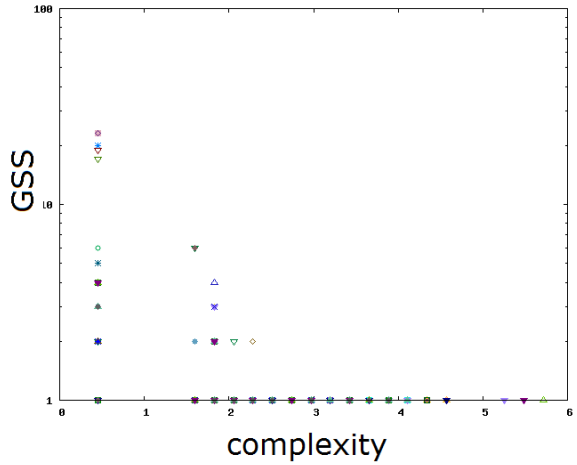

Figure 10: Scatter plot of GSS versus Kolmogorov complexity for the generic cell cycle model. Again, different symbols represent the different variables in the system, and the GSS of a phenotype follows $GSS(x) \propto 2^{-K(x)}$.

In spite of having so many simplifications, these models have successfully analysed several gene networks [11], including a model for the budding yeast [21] and another for the fission yeast [9] cell cycles. Fig. 11 shows those networks: in both yeast cell cycles, the natural phenotype visibly has a large basin of attraction, which can also be seen as a form of environmental robustness, where many outputs that do not correspond to states of the cell cycle eventually result on it, without any form of exterior regulation.

Having these results in mind, as well as the previous sections, the question of how a gene network works as a GP map might come to mind. The system does have the requirements: its genotype space quickly becomes hyperastronomical (the number of possible Boolean networks with 15 genes is $3^{15\times15} > 10^{100}$), while the number of possible outputs is significantly smaller - the output of such network can be described as a transition matrix, mapping one state to another. Given $2^{15} = 32768$ states, since each state only maps to another one, there are $32768^2 \approx 10^9$ possible outputs. Therefore, the mapping from gene network to transition matrix seems to be a good example of GP map.

And indeed, it does present features of such: Fig. 12 shows the size L of the largest cycles found in 10000 random Boolean networks with 10 nodes. It is possible to spot a certain trend for the probability of having size L for the largest limit cycle: $P(L) \propto 2^{-L}$. Taking Nochomovitz and Li's argument that the complexity of the phenotype is proportional to cycle size [23], this would mean $P(x) \propto 2^{-K(x)}$, pointing again to Levin's coding theorem, and to the general trend here observed in GP maps.

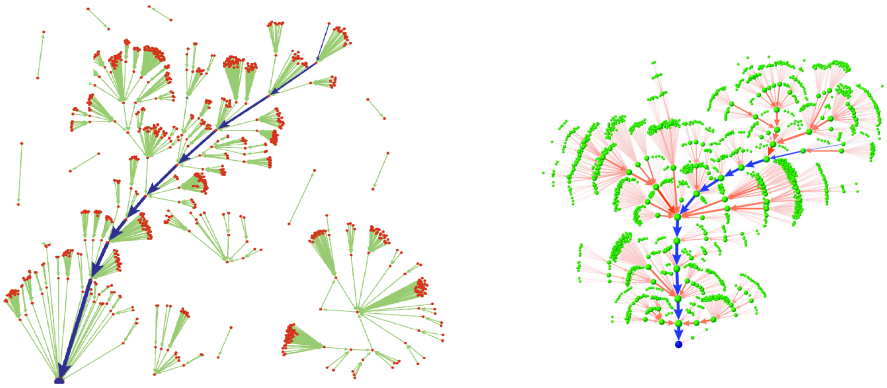

Figure 11: States in the boolean models for the budding (left) and fission (right) yeast cell cycle boolean models. In these networks, a state is a pattern of activation/inhibition of its components, represented by the nodes in the figure above. The arrows represent which state leads to which, the blue arrows being the chain of reactions of the cell cycle. This shows how in both cases the cell cycle is a strong attractor, given that so many states, after one or two steps, map into it.

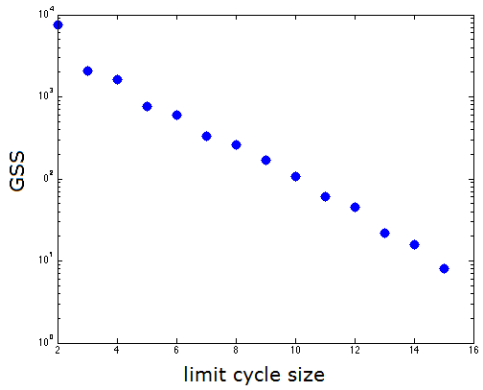

Figure 12: Random boolean networks of size 10 produce limit cycles with probability depending on cycle length, which can be seen as a proxy for the complexity of that output: $P(L) \propto 2^{-L}$.

5 Conclusions

This report has covered some examples of bias in nature - how cell cycles are robust and how mutation can beat selection. It is very clear that variation has a very important role in evolution - more specifically, that bias, either in rates or maps, can outweigh selection in determining how structures will evolve. The systems here studied indicate that the generalisation of the coding theorem for GP maps [10] is

true, but they also raise more questions than answers: is the distance between sets of genotypes reflected in the distance between phenotypes? How does this work for the boolean networks? How can one know that the bias of a given map is more than just random, since it is impossible to explore the whole parameter space? How do selection, mutation and drift operate on maps, how do they change over time? How important is contingency? How do sloppiness, robustness, and evolvability emerge?

Those among many others are questions that have yet to be answered: how phenotypic biases and multilevel mapping work together with selection, mutation and drift. At this point it is certain that no understanding of life will come without comprehensive knowledge of how life not only is shaped by, but also how it shapes, evolution.

References

[1] Réka Albert and Hans G Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *Journal of theoretical biology*, 223(1):1–18, 2003.

[2] Elhanan Borenstein and David C Krakauer. An end to endless forms: epistasis, phenotype distribution bias, and nonuniform evolution. *PLoS computational biology*, 4(10):e1000202, 2008.

[3] Fabio ACC Chalub and Max O Souza. The frequency-dependent wright–fisher model: diffusive and non-diffusive approximations. *Journal of mathematical biology*, 68(5):1089–1133, 2014.

[4] Fabio ACC Chalub, Max O Souza, et al. A non-standard evolution problem arising in population genetics. *Communications in Mathematical Sciences*, 7(2):489–502, 2009.

[5] Katherine C Chen, Laurence Calzone, Attila Csikasz-Nagy, Frederick R Cross, Bela Novak, and John J Tyson. Integrative analysis of cell cycle control in budding yeast. *Molecular biology of the cell*, 15(8):3841–3862, 2004.

[6] Matthew C Cowperthwaite, Evan P Economo, William R Harcombe, Eric L Miller, and Lauren Ancel Meyers. The ascent of the abundant: how mutational networks constrain evolution. *PLoS computational biology*, 4(7):e1000110, 2008.

[7] Francis HC Crick. The origin of the genetic code. *Journal of molecular biology*, 38(3):367–379, 1968.

[8] Attila Csikász-Nagy, Dorjsuren Battogtokh, Katherine C Chen, Béla Novák, and John J Tyson. Analysis of a generic model of eukaryotic cell-cycle regulation. *Biophysical journal*, 90(12):4361–4379, 2006.

[9] Maria I Davidich and Stefan Bornholdt. Boolean network model predicts cell cycle sequence of fission yeast. *PloS one*, 3(2):e1672, 2008.

[10] Kamaludin Dingle. *Probabilistic Bias in Genotype-Phenotype Maps*. PhD thesis, University of Oxford, July 2014.

[11] Kamaludin Dingle and Ard Louis. Unpublished.

[12] Carlos Espinosa-Soto, Pablo Padilla-Longoria, and Elena R Alvarez-Buylla. A gene regulatory network model for cell-fate determination during arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *The Plant Cell Online*, 16(11):2923–2939, 2004.

[13] Thomas MA Fink, Karen Willbrand, and Francis CS Brown. 1-d random landscapes and non-random data series. *EPL (Europhysics Letters)*, 79(3):38006, 2007.

[14] Dimitri Gilis, Serge Massar, Nicolas J Cerf, Marianne Rooman, et al. Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome Biol*, 2(11):49–1, 2001.

[15] Fred Hoyle and Chandra Wickramasinghe. Evolution from space. *Evolution from space., by Hoyle, F.; Wickramasinghe, C.. London (UK): JM Dent, 176 p.*, 1, 1981.

[16] Shalev Itzkovitz and Uri Alon. The genetic code is nearly optimal for allowing additional.

[17] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467, 1969.

[18] Motoo Kimura. Diffusion models in population genetics. *Journal of Applied Probability*, 1(2):177–232, 1964.

[19] Abraham Lempel and Jacob Ziv. On the complexity of finite sequences. *Information Theory, IEEE Transactions on*, 22(1):75–81, 1976.

[20] Leonid A Levin. Laws of information conservation (nongrowth) and aspects of the foundation of probability theory. *Problemy Peredachi Informatsii*, 10(3):30–35, 1974.

[21] Fangting Li, Tao Long, Ying Lu, Qi Ouyang, and Chao Tang. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14):4781–4786, 2004.

[22] AL Mackay. Optimization of the genetic code. *Nature*, 216:159–160, 1967.

[23] Yigal D Nochomovitz and Hao Li. Highly designable phenotypes and mutational buffers emerge from a systematic mapping between network topology and dynamic output. *Proceedings of the National Academy of Sciences of the United States of America*, 103(11):4180–4185, 2006.

[24] Isabelle S Peter, Emmanuel Faure, and Eric H Davidson. Predictive computation of genomic logic processing functions in embryonic development. *Proceedings of the National Academy of Sciences*, 109(41):16434–16442, 2012.

[25] Karthik Raman and Andreas Wagner. The evolvability of programmable hardware. *Journal of The Royal Society Interface*, page rsif20100212, 2010.

[26] Peter Schuster, Walter Fontana, Peter F Stadler, and Ivo L Hofacker. From sequences to shapes and back: a case study in rna secondary structures. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 255(1344):279–284, 1994.

[27] Andreas Wagner. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630):91–100, 2008.

[28] Sewall Wright. *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, volume 1. na, 1932.

[29] Lev Y Yampolsky and Arlin Stoltzfus. Bias in the introduction of variation as an orienting factor in evolution. *Evolution & development*, 3(2):73–83, 2001.

[30] Feng Zhang, Li Xu, Kun Zhang, Erkang Wang, and Jin Wang. The potential and flux landscape theory of evolution. *The Journal of chemical physics*, 137(6):065102, 2012.