NKα: Non-uniform epistatic interactions in an extended NK model

Tom Hebbron[1], Seth Bullock[1] and Dave Cliff[2]

[1] School of Electronics and Computer Science
University of Southampton
Southampton, UK
[2] Department of Computer Science
University of Bristol
Bristol, UK
tomhebbron@zepler.net

Abstract

Kauffman's seminal NK model was introduced to relate the properties of fitness landscapes to the extent and nature of epistasis between genes. The original model considered genomes in which the fitness contribution of each of N genes was influenced by the value of K other genes located either at random or from the immediately neighbouring loci on the genome. Both schemes ensure that (on average) every gene is as influential as any other. More recently, the epistatic connectivity between genes in natural genomes has begun to be mapped. The topologies of these genetic networks are neither random nor regular, but exhibit interesting structural properties. The model presented here extends the NK model to consider epistatic network topologies derived from a preferential attachment scheme which tends to ensure that some genes are more influential than others. We explore the consequences of this topology for the properties of the associated fitness landscapes.

Introduction

Recent advances in our understanding of natural genomes are beginning to reveal patterns in genomic organisation (Jeong et al., 2000; Barabási and Oltvai, 2004; Segrè et al., 2004). In particular, the epistatic networks that describe the manner in which genetically specified proteins interact with each other during cell metabolism have been shown to exhibit topologies that are scale-free in their degree distribution (Maslov and Sneppen, 2002; Fernández, 2007). In such networks, while the vast majority of proteins are involved in only a small number of protein-protein interactions, a few proteins are highly influential (Barabási et al., 1999).

Here, we explore the influence of this type of epistatic network topology on the structure of associated fitness landscapes using an extension of the NK model originally proposed by Kauffman (1989). In the canonical form of this model, the fitness associated with a particular genotype (i.e., the height associated with a particular point on the fitness landscape) is assessed by combining the fitness contributions of the binary alleles at each of its N loci. The fitness contribution of a locus, i, is determined by the allele at i and the alleles present at K additional loci. For each unique combination of $K+1$ alleles, a unique, but randomly determined fitness contribution is assigned. By considering the statistical properties of ensembles of NK landscapes, the generic influence of epistasis can be assessed.

Kauffman was able to demonstrate that the 'ruggedness' of a landscape increases with increasing K. For $K = 0$ landscapes, each locus contributes to fitness independently. The landscape is smooth, with the fitness of adjacent genotypes being highly correlated as a consequence of sharing $N - 1$ fitness components. An adaptive walk originating at any point on such a landscape will reach a single, unique optimum. Every step on such a walk will reduce the distance to the optimum by one as the allele at one locus mutates to a fitter variant. The mean length of such a walk is therefore $\frac{N}{2}$. By contrast, for landscapes where $K = N - 1$ a mutation at any locus has the side effect of changing the fitness contribution of the alleles at all other loci. Consequently, there is no correlation between the fitness of neighbouring genotypes, and the landscape is maximally rugged. A large proportion ($\frac{1}{N+1}$) of genotypes are now local optima, and adaptive walks tend to stall after $ln(N - 1)$ steps. Intermediate values of K give rise to intermediate levels of ruggedness, altering the average distance between local optima, the correlation amongst locally optimal genotypes, and the fitness distribution of local optima. For more details, see Altenberg (1997).

In the two most frequently explored forms of the model, the K loci that epistatically influence a particular locus, i, may either be randomly located on the genome, or may be the K nearest neighbour loci of i. In both cases, every gene influences the fitness contribution of (on average) the same number of other genes, ensuring that genes are equipotent in their contribution to genotypic epistasis. Many variants of this model have been considered, and its behaviour has been explored in various ways (Altenberg, 1997; Barnett, 1998; Geard et al., 2002; Gao and Culberson, 2002; Campos et al., 2002; Rivkin and Siggelkow, 2002; Verel et al., 2003; Skellett et al., 2005; Kaul and Jacobson, 2007).

Kauffman himself mentions briefly a variant of the model in which some genes are more influential than others (Kauffman, 1989, pp78). Here, we develop this idea and explore the implications of systematically manipulating the extent to which there is a particular scale-free non-uniformity in the degree of influence exerted by each gene on the fitness contribution of the remainder of the genome.

Scale-free degree distributions have been discovered to characterise connectivity in a wide variety of systems, from gene regulatory networks to scientific citation networks (Barabási et al., 1999; Rzhetsky and Gomez, 2001; Gisiger, 2001; Wolf et al., 2002; Barabási et al., 2002; Barabási, 2003). In each case, the frequency with which network nodes exhibit degree k is proportional to $k^{-\gamma}$, where $\gamma > 1$ (Barabási and Crandall, 2003). Scale-free networks of this kind may be grown via a process of 'preferential attachment' (Barabási et al., 1999; Newman, 2001; Caldarelli et al., 2002; Eisenberg and Levanon, 2003). Under such a scheme, nodes are added sequentially to an initial small graph. Upon being added to the graph, each node is allocated a number of edges linking it to existing nodes, where the probability of adding an edge to an existing node of degree k is proportional to k^{α}. Here, α is a model parameter governing the strength of preferential attachment.

Networks with a scale-free topology have some distinct properties.

Self similarity at different scales: properties of local areas of the network are echoed in the whole.

The small-world phenomenon: shortest paths between any pair of nodes are remarkably short (Watts and Strogatz, 1998; Albert et al., 1999; Lazer and Friedman, 2005; Giacobini et al., 2006).

Robust to random failure: removal of nodes at random has little effect on network structure. However they are vulnerable to attacks that target the highly connected hubs (Albert et al., 2000; Barabási, 2003; Barabási and Crandall, 2003).

This paper first specifies an extended NK model, NKα, and describes the metrics that will be used to characterise its fitness landscapes. Results from the novel model are then compared with those of the canonical NK models, and their implications discussed before, finally, some future work is suggested.

Methods

An extensible NK model was implemented using a variation on the hashing method described by Altenberg (Altenberg, 1994, 1997), and using an efficient hashing algorithm proven against funnelling effects (Jenkins, 1997). The model was validated against published data from several sources for the Kauffman local and random variants (Kauffman, 1989; Weinberger, 1991; Kauffman, 1993, 1995; Altenberg, 1997). The random number generator and hashing functions were tested using the NIST validation suite (Rukhin et al., 2001).

The network of epistatic interactions between loci was represented as an $N \times N$ Boolean matrix, A, with $A_{ij} = 1$ iff locus i influences the fitness contribution of locus j. Since each locus always contributes to its own fitness contribution, $A_{ii} = 1 \; \forall i$. Furthermore, $\sum_i A_{ij} = K + 1 \; \forall j$, since each row of A contains K entries in addition to the self-connection, corresponding to j's incoming edges. By contrast, the sum of each column of A corresponds to the out-degree of each locus, which, in general, may be free to vary such that $1 \leq \sum_j A_{ij} \leq N$. Under all schemes considered here $\sum_{i,j} A_{ij} = N(K + 1)$, i.e., the total number of edges in the network is conserved.

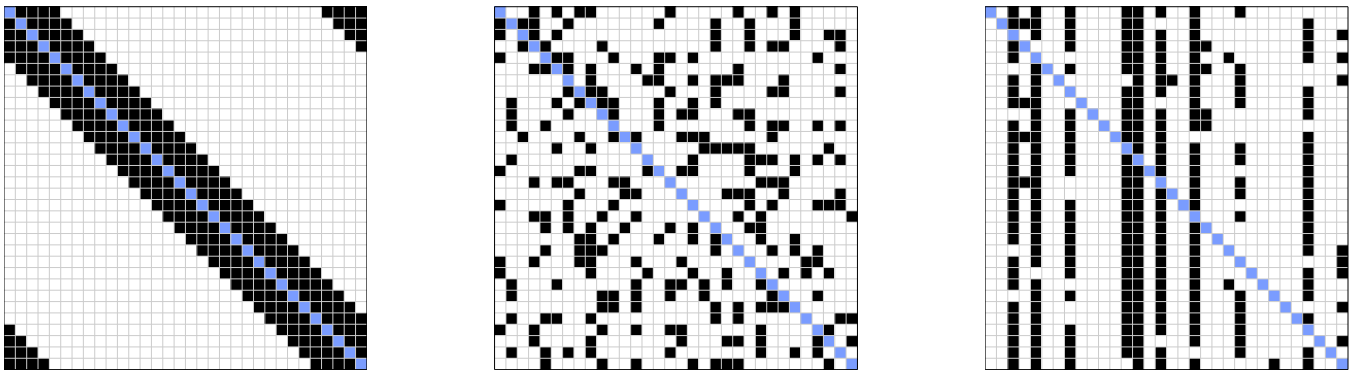
Kauffman's original NK model employed two schemes for allocating the epistatic links: local or random. In the former, each locus is influenced by its K nearest neighbours, giving rise to an epistatic network with a ring-lattice topology (see fig. 1a). In the latter, for each locus, K unique influential loci are chosen at random, giving rise to a random graph topology (see fig. 1b). Under the local scheme both in- and out-degree are uniform, whereas under the random scheme in-degree is uniform, but the out-degree is a Poisson distribution with a mean of K Newman et al. (2001).

Here, we introduce NKα, a variant of the NK model that employs a scale-free epistatic topology, parametrised by a single exponent, α. As before, the network contains $N(K + 1)$ edges, N of which are self-connections, and each locus has the same in-degree $(K + 1)$. However, the out-degree distribution approximates a power-law as a consequence of the following preferential attachment growth process.

Initially each locus is connected only to itself, giving a degree of 1. Subsequently, we perform K passes through the list of loci. Each pass visits each locus once in random order. On each visit, the visited locus is assigned one incoming edge from a random locus, i, chosen with probability $\propto (k_i)^{\alpha}$, where k_i is the out-degree of locus i and is updated after each visit, and the magnitude of α determines the strength of preferential attachment. This process assigns a total of $N(K+1)$ edges with a power-law like degree distribution, save that a ceiling threshold exists: no locus can have more than N connections (including its self-connection). With sufficiently high K or α, some loci will attract the maximum N connections deforming the power-law curve. When $\alpha = 0$, the resulting epistatic matrix is equivalent to the random map explored in the original model. Where $\alpha > 0$, increasingly skewed degree-distributions are generated, conferring increasing influence on a minority of loci (see Fig. 1c).

Measuring landscape properties

For the model introduced here, a triple (N, K, α) specifies an ensemble of landscapes that we sample and evaluate



(a) Local (b) Random (c) Scale-free: $\alpha = 1.8$

Figure 1: Epistatic maps for $N = 32$ with $K = 8$ using (a) local connectivity, (b) random connectivity, and (c) the NKα variant with $\alpha = 1.8$.

below. In addition to sampling the fitness distribution over each landscape as a whole by sampling 10,000 genomes at random on each landscape, we perform a number of walks across landscapes. Walks are of two types.

Adaptive walks were carried out by simple hill-climbers. At each step, a hill-climber calculates the fitness of all N single bit mutation neighbours of the current genotype, and selects one of the fitter neighbours at random to move to. If no fitter neighbour exists, then the hill-climber has reached a local optimum, and terminates. By undertaking multiple independent walks on the same landscape, an assay of available local optima can be compiled. Additionally, the length of adaptive walks is an indicator of a landscape's 'ruggedness'.

Random walks start from a random position in the landscape, and proceed by a series of random single-point mutations. Here, random walks were terminated after 2048 steps, as described by Weinberger (Kauffman, 1993). Such walks allow an assay of fitness distributions, and the correlation between the fitness of points separated by intervening genotypes.

Results

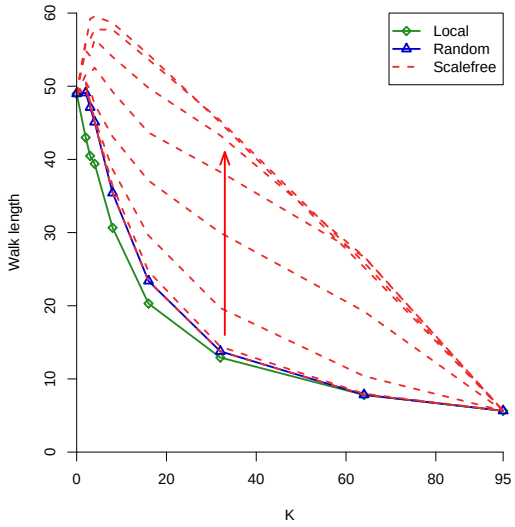
Unless otherwise stated, genotype length is held constant with $N = 96$, K ranges over $\{0, 1, 2, 3, 4, 8, 16, 32, 64, 81, 95\}$, and α ranges over $\{0.0, 0.5, 1.0, 1.5, 1.8, 2.0, 2.5\}$. By 'a full range of landscapes' we will mean all combinations of N and K for the local and random variants of the original NK model, and all combinations of N, K and α for the NKα variant. For the majority of results presented, the data is an aggregation of 10 repetitions of all combinations, each of these repetitions having a different seed and consequently different epistatic matrix.

Figure 2a shows the manner in which the average length of an adaptive walk varies with K for the landscapes considered here. While, in general, walk length increases with K, it is also apparent that walks tend to be longer for landscapes with higher values of α. Moreover, for low K and high α walks tend to involve a number of steps that exceeds $\frac{N}{2}$, the maximum average walk length observed for the local and random variants.

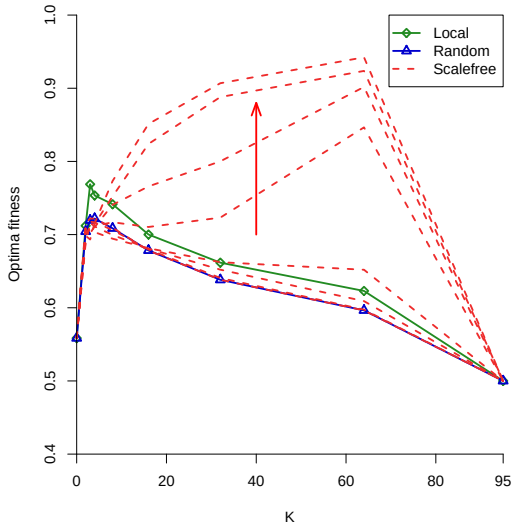
Figure 2b demonstrates that increasing α has a dramatic effect on the way in which K influences the mean fitness of landscape optima. For intermediate values of K, increasing α is associated with increasingly fit optima. In both figures we see that as α reaches high values, its influence asymptotes. This results from the ceiling effect mentioned above, which restricts the distribution of epistatic influences such that a few loci have the maximum influence, while the remainder have little or no influence at all.

For both the local and random variants of the NK model the correlation between the fitness at a local optimum and the fitness of its neighbours decreases with increasing K. Here, Figure 3 compares the distribution of fitness values of genotypes adjacent to a local optimum in a 'local' landscape with the distribution of fitness values adjacent to a local optimum on a landscape with relatively high α. The comparison is made for $N = 96$ and $K = 64$, but the qualitative results are characteristic of the comparison in general. For the local variant, fitness values adjacent to a local optimum are relatively tightly distributed around a value somewhat lower than the fitness at the local optimum. For the NKα variant, however, the distribution of adjacent fitnesses is much broader with many values close to the fitness of the local optimum, and many values far from it.

Figure 4a demonstrates that, for low K, the random NK model variant exhibits optima with a range of basin sizes and that there is a weak correlation between the fitness of an optimum and the size of its basin of attraction. Increasing K destroys this correlation, rendering every optima essentially equally attainable regardless of fitness (Figure 4b). However, the NKα variant gives rise to optima with a variety of



(a) Adaptive walk length

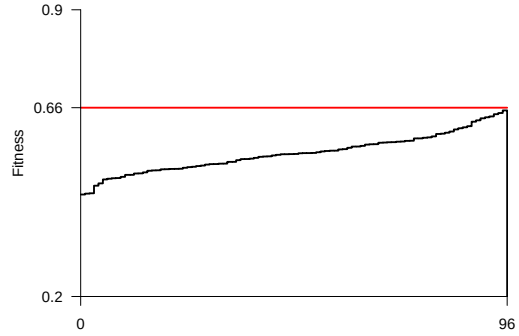


(b) Optima fitness

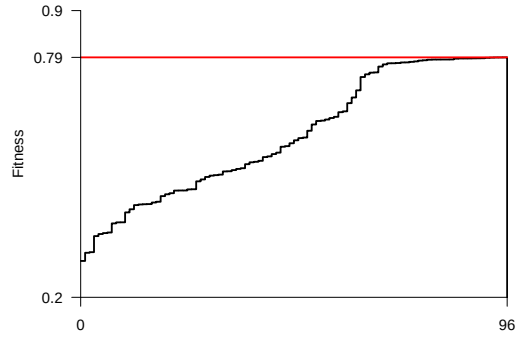
Figure 2: The mean length and mean local optimum fitness found for 100,000 adaptive walks for each (K, α) combination ($N = 96$). The arrow indicates increasing values of $0 \leq \alpha \leq 2.5$ for NKα variant.

basin sizes for high K landscapes, and here, optima fitness is strongly correlated with basin size. This accounts both for the fact that adaptive walks are taken on NKα landscapes tend to be longer than those carried out for equivalent landscapes from the local or random model variants, and that they tend to terminate at optima of higher fitness.

Kauffman used the term 'Massif Central' to describe a global structure he discovered in landscapes with small K. He used this term to refer to the tendency for high-fitness local optima to be located in the vicinity of the global optimum, rather than being randomly distributed as is the case for high K. The inverse correlation between fitness and distance in Figure 5a and its gradual erosion in Figures 5b and 5c reflect this observation. However, when we consider the NKα variant, we find a similar but stronger relationship with many fit optima close to the global optimum. Unlike for the original model variants, for the NKα variant this relationship between optima fitness and hamming distance is *strengthened* by increasing K.



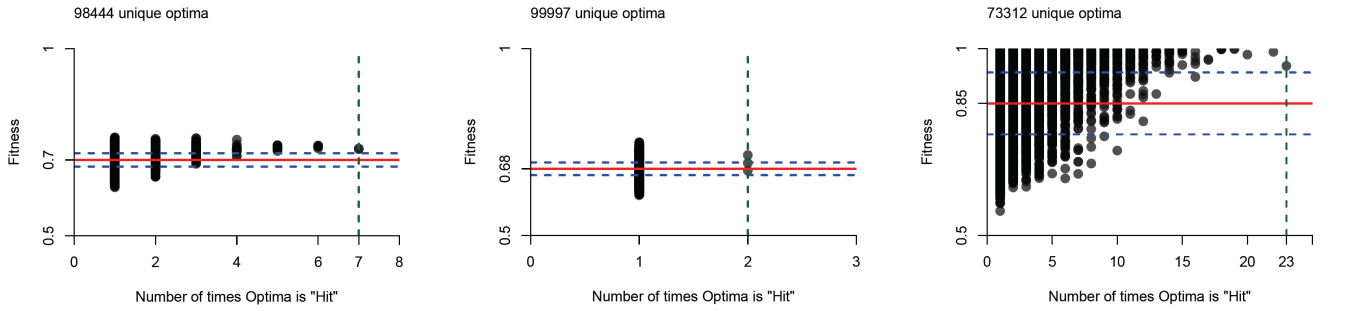
(a) Local



(b) Scale-free $\alpha = 1.8$

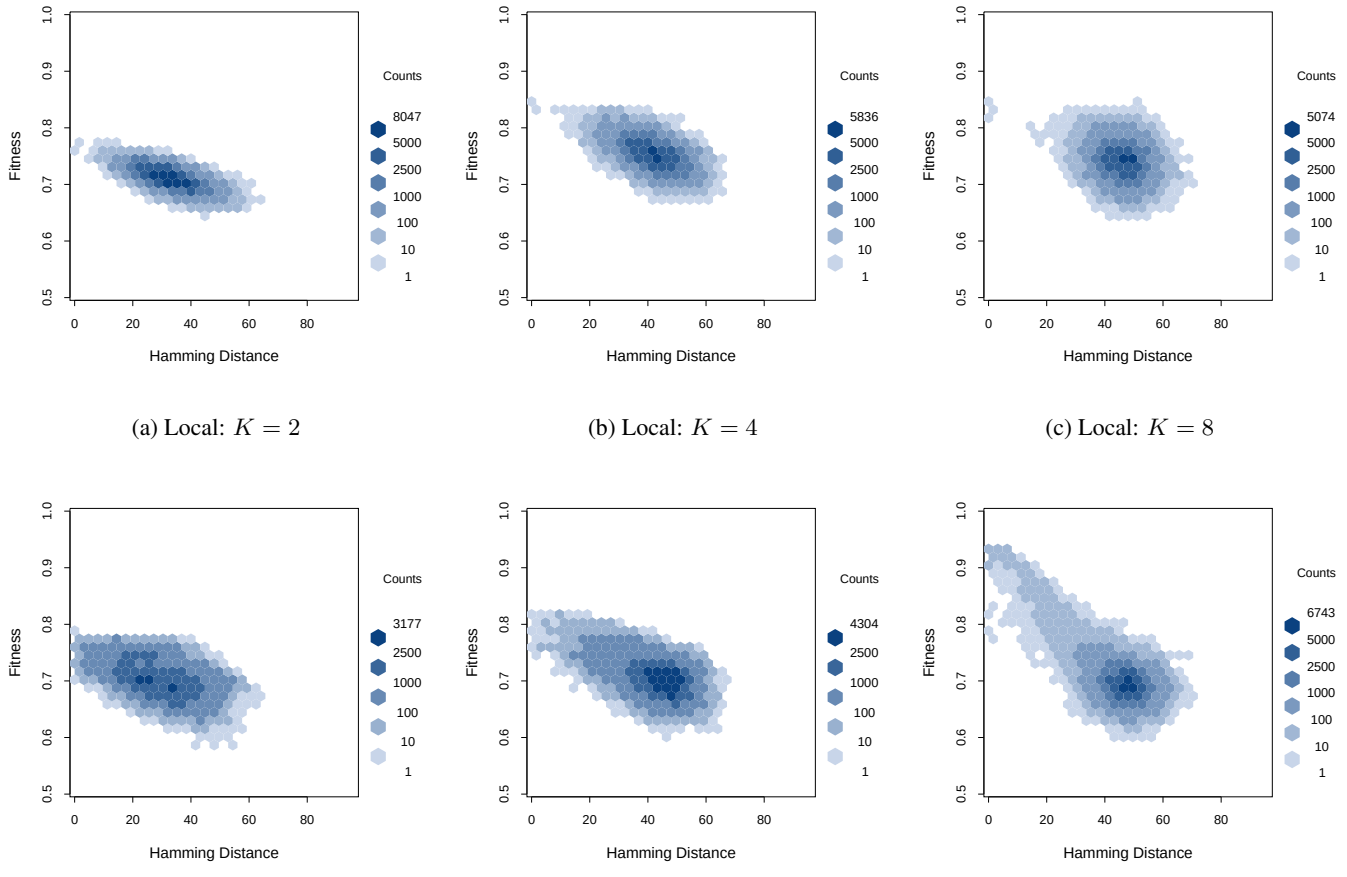
Figure 3: The fitness of neighbours of a local optimum are plotted for two landscapes ($N = 96$, $K = 64$), ordered by fitness. Solid horizontal lines indicate the local optimum fitness.

For the local and random variants of the NK model, loci are (roughly) epistatically equipotent. However, in the NKα model, some loci are more influential than others. How does this affect the rate at which different loci are mutated during an adaptive walk? At each step of an adaptive walk, plotting the out-degree of the mutated gene (how many loci it influences epistatically) reveals that for $\alpha > 0$, the most influential loci become fixed early in the adaptive walk. This effect increases with increasing α. For fixed α and N, increasing K increases the number of influential loci, and (if $\alpha > 0$) decreases the number of weakly connected loci. This lengthens the phase during which influential nodes are 'locked in.' By contrast, in the random variant (a), the variation in out-degree is much reduced, walks tend to be shorter, and there is no relationship between the out-degree of a locus and its tendency to be mutated early or late in an adaptive walk.



(a) Random: $K = 2$ (b) Random: $K = 16$ (c) Scale-free $K = 16, \alpha = 2.5$

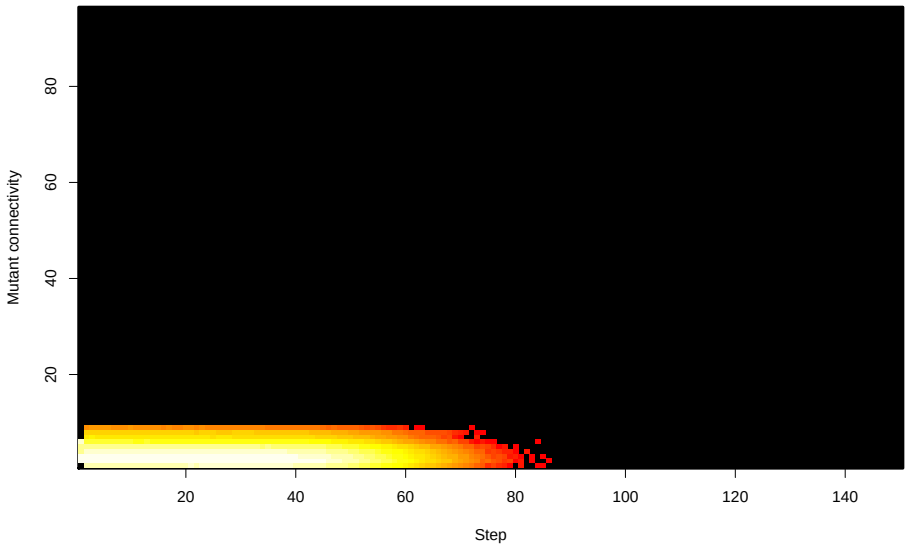
Figure 4: The accessibility of local optima discovered by 10 repetitions of 10,000 independent adaptive walks on three classes of landscape. Solid horizontal lines indicates mean fitness with standard deviation indicated by dashed horizontal lines. Vertical dashed lines indicate the most frequently reached optima.



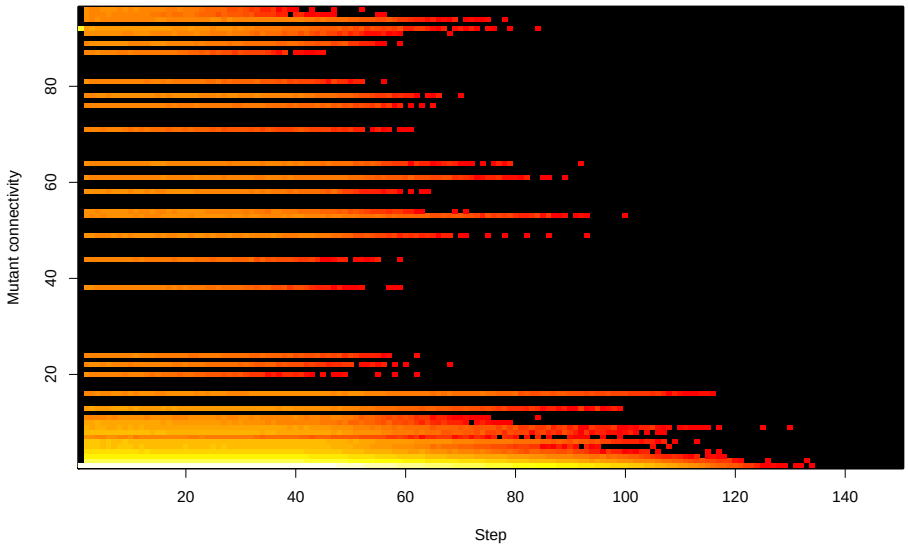
(a) Local: $K = 2$ (b) Local: $K = 4$ (c) Local: $K = 8$

(d) Scale-free: $K = 2, \alpha = 1.5$ (e) Scale-free: $K = 4, \alpha = 1.5$ (f) Scale-free: $K = 8, \alpha = 1.5$

Figure 5: Each plot shows the hamming distance from the fittest optima found to the other optima found on 100,000 adaptive walks with $N = 96$ and $K = 2, 4, 8$. The data has been binned, and plotted with boundaries {0,1,100,1000,2500,5000,10000} to give a sense of the density of clustered optima. Darker tone indicates higher density of optima. The key on each plot shows only those bins used, with an upper limit of the most dense point in the plot. For the local variant (a-c) increasing K gradually erodes an inverse correlation between hamming distance and optima fitness. For the NKα variant, (d-f) increasing K strengthens a similar relationship.



(a) Random



(b) Scale-free $\alpha = 2.5$

Figure 6: When are influential loci mutated on an adaptive walk? For a sample of 100,000 adaptive walks, the heat map depicts the degree of mutated loci at each time step for (a) the random model variant, (b) the NKα variant. Here, $N = 96, K = 2$. The heat colour scheme is used: black indicates no mutation events, increasing to white indicating the most frequent mutation events.

Discussion

In general, imposing an increasingly scale-free structure on the network of epistatic interactions brings about a number of significant changes to the behaviour of adaptive walks on the associated fitness landscape: longer adaptive walks, higher fitness optima, more clustering of optima in the landscape and increased correlation between their fitness and the distance between them.

When the K epistatic influences of a locus are uniformly distributed, the resultant landscape is essentially isotropic. Statistical properties in one part of the landscape are largely predictive of the whole. Consequently, the effect of increasing K is to impose ruggedness globally. Conversely, when the same number of epistatic interactions are allocated non-uniformly, the genome is structured such that there exist a few influential loci and a majority of loci with little or no influence. This structure gives rise to a radically anisotropic landscape. Portions of the landscape exhibit properties that are very different from one another. More specifically, fixing alleles at influential loci confines an adaptive walk to a relatively correlated sub-landscape, while fixing the same number of low-influence loci confines an adaptive walk to a much less correlated landscape. Adaptive walks on such landscapes tend to initially spend time fixing influential loci, since mutating these alleles can bring about significant fitness changes. Once a satisfactory configuration of highly influential loci is discovered, low influence loci can be fixed relatively easily, since each is essentially independent from the others.

As yet it is unclear the extent to which one might describe NKα landscapes as modular. Are there multiple Massif Centrals on these landscapes, each characterised by a cluster of local optima of similar fitness? Or is there a more gross organisation of optima across the landscape as a whole? Relatedly, we have not considered assortativity in the network of epistatic interactions. While it has been known for some time that, for instance, the network of protein-protein interactions for yeast exhibits a scale-free degree distribution, recent work has shown that although the network for ancestral yeast has high degree proteins tending to interact directly with one another, the network for contemporary yeast is less assortative, with what has been interpreted as a more modular structure. For instance, precursors to modern yeast feature an epistatic network with a single hub related to the ribosome, whereas the modern yeast network exhibits two hubs, one ribosomal and the other related to signalling. These hubs are connected, but only via other poorly connected proteins, making the whole network appear modular (Fernández, 2007). Scale-free network topologies tend to be robust to failure unless the hubs are targetted (Albert et al., 2000; Barabási, 2003; Barabási and Crandall, 2003; Jeong et al., 2001), and a modular topology has the advantage of preventing the failure of one hub triggering the failure of another.

The preferential attachment algorithm used here defines an epistatic network topology with a scale-free out degree, which has significant effects on the resulting fitness landscape. However, the class of networks with scale-free degree distribution encompasses a range of topologies. In future work, we will extend the current NKα variant to consider the influence of assortative epistatic network topologies; the fitness landscapes and evolutionary dynamics to which they give rise.

Conclusions

The human genome project revealed a far lower number of genes than anticipated, increasing the significance of the study of their interactions. By extending an existing model, the paper demonstrates how a scale-free epistatic network topology alters the properties of a fitness landscape in a way that makes adaptive dynamics on it much more liable to discover high-fitness optima despite strong epistasis. To the best of our knowledge, and also to our surprise, this is the first systematic study of how the standard NK results vary when a preferential attachment scheme is used for determining the epistatic linkages between loci.

References

Albert, R., Jeong, H., and Barabási, A. (1999). The diameter of the world wide web. *Arxiv preprint cond-mat/9907038*.

Albert, R., Jeong, H., and Barabási, A. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

Altenberg, L. (1994). Evolving better representations through selective genome growth. In *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, pages 182–187. IEEE Press.

Altenberg, L. (1997). NK fitness landscapes. In Bäck, T., Fogel, D. B., and Michalewicz, Z., editors, *Handbook of Evolutionary Computation*, pages B2.7.1–B2.7.10. Institute of Physics Press and Oxford University Press, New York.

Barabási, A. (2003). *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Plume.

Barabási, A., Albert, R., and Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A*, 272(1):173–187.

Barabási, A. and Crandall, R. (2003). Linked: The new science of networks. *American Journal of Physics*, 71:409.

Barabási, A., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., and Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A*, 311(3-4):590–614.

Barabási, A. and Oltvai, Z. (2004). Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2):101–113.

Barnett, L. (1998). Ruggedness and neutrality: The NKp family of fitness landscapes. In Adami, C., Belew, R., Kitano, H., and Taylor, C., editors, *Artificial Life VI: Proceedings of the*

Sixth International Conference on Artificial Life, pages 18–27. MIT Press, Cambridge, MA.

Caldarelli, G., Capocci, A., De Los Rios, P., and Munoz, M. (2002). Scale-free networks without growth or preferential attachment: Good get richer. *Arxiv preprint cond-mat/0207366*.

Campos, P., Adami, C., and Wilke, C. (2002). Optimal adaptive performance and delocalization in NK fitness landscapes. *Physica A*, 304(3-4):495–506.

Eisenberg, E. and Levanon, E. (2003). Preferential attachment in the protein network evolution. *Physical Review Letters*, 91(13):138701.

Fernández, A. (2007). Molecular basis for evolving modularity in the yeast protein interaction network. *PLoS Comput Biol*, 3(11):e226.

Gao, Y. and Culberson, J. C. (2002). An analysis of phase transition in NK landscapes. *Journal of Artificial Intelligence Research*, 17:309–332.

Geard, N., Wiles, J., Hallinan, J., Tonkes, B., and Skellett, B. (2002). A comparison of neutral landscapes—NK, NKp and NKq. In Fogel, D. B., El-Sharkawi, M. A., Yao, X., Greenwood, G., Iba, H., Marrow, P., and Shackleton, M., editors, *Proceedings of the 2002 Congress on Evolutionary Computation*, pages 205–210. IEEE Press.

Giacobini, M., Preuss, M., and Tomassini, M. (2006). Effects of scale-free and small-world topologies on binary coded self-adaptive CEA. In *Evolutionary Computation in Combinatorial Optimization*, pages 86–98. Springer, Berlin/Heidelberg.

Gisiger, T. (2001). Scale invariance in biology: Coincidence or footprint of a universal mechanism? *Biological Reviews*, 76(02):161–209.

Jenkins, B. (1997). Hash functions. *Dr. Dobbs Journal*, 9709.

Jeong, H., Mason, S., Barabási, A., Oltvai, Z., et al. (2001). Lethality and centrality in protein networks. *Nature*, 411(6833):41–42.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z., Barabási, A., et al. (2000). The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654.

Kauffman, S. (1989). Adaptation on rugged fitness landscapes. In Stein, D., editor, *Lectures in the Sciences of Complexity*, volume 1, pages 527–618. Addison-Wesley, Redwood City, CA.

Kauffman, S. A. (1993). *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press, Oxford.

Kauffman, S. A. (1995). *At Home in the Universe: The Search for Laws of Self-organization and Complexity*. Oxford University Press, Oxford.

Kaul, H. and Jacobson, S. (2007). New global optima results for the Kauffman NK model: Handling dependency. *Mathematical Programming*, 108(2):475–494.

Lazer, D. and Friedman, A. (2005). The parable of the hare and the tortoise: Small worlds, diversity, and system performance. KSG Working Paper No. RWP05-058, John F. Kennedy School of Government, Harvard University.

Maslov, S. and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science*, 296(5569):910.

Newman, M. (2001). Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2):25102.

Newman, M. E. J., Strogatz, S. H., and Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64(2):026118.

Rivkin, J. and Siggelkow, N. (2002). Organizational sticking points on NK landscapes. *Complexity*, 7(5):31–43.

Rukhin, A., Soto, J., Nechvatal, J., Smid, M., Barker, E., Leigh, S., Levenson, M., Vangel, M., Banks, D., Heckert, A., Dray, J., and Vo, S. (2001). A statistical test suite for the validation of random number generators and pseudo random number generators for cryptographic applications. Special Publication 800-22, NIST.

Rzhetsky, A. and Gomez, S. (2001). Birth of scale-free molecular networks and the number of distinct DNA and protein domains per genome. *Bioinformatics*, 17(10):988–996.

Segrè, D., DeLuna, A., Church, G., and Kishony, R. (2004). Modular epistasis in yeast metabolism. *Nature Genetics*, 37:77–83.

Skellett, B., Cairns, B., Geard, N., Tonkes, B., and Wiles, J. (2005). Maximally rugged NK landscapes contain the highest peaks. In Beyer, H. G., editor, *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 579–584. ACM Press, New York, NY.

Verel, S., Collard, P., and Clergue, M. (2003). Where are bottlenecks in NK fitness landscapes? In Sarker, R., Reynolds, R., Abbass, H., Tan, K. C., McKay, B., Essam, D., and Gedeon, T., editors, *Proceedings of the 2003 Congress on Evolutionary Computation*, pages 273–280. IEEE Press.

Watts, D. and Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):409–10.

Weinberger, E. (1991). Local properties of Kauffman's NK model: A tunably rugged energy landscape. *Physical Review A*, 44(10):6399–6413.

Wolf, Y. I., Karev, G., and Koonin, E. V. (2002). Scale-free networks in biology: New insights into the fundamentals of evolution? *Bioessays*, 24(2):105–9.