nature ecology & evolution

ARTICLES

PUBLISHED: 23 JANUARY 2017 | VOLUME: 1 | ARTICLE NUMBER: 0045

A thousand empirical adaptive landscapes and their navigability

José Aguilar-Rodríguez[1,2†], Joshua L. Payne[1,2†] and Andreas Wagner[1,2,3*]

The adaptive landscape is an iconic metaphor that pervades evolutionary biology. It was mostly applied in theoretical models until recent years, when empirical data began to allow partial landscape reconstructions. Here, we exhaustively analyse 1,137 complete landscapes from 129 eukaryotic species, each describing the binding affinity of a transcription factor to all possible short DNA sequences. We find that the navigability of these landscapes through single mutations is intermediate to that of additive and shuffled null models, suggesting that binding affinity—and thereby gene expression—is readily fine-tuned via mutations in transcription factor binding sites. The landscapes have few peaks that vary in their accessibility and in the number of sequences they contain. Binding sites in the mouse genome are enriched in sequences found in the peaks of especially navigable landscapes and the genetic diversity of binding sites in yeast increases with the number of sequences in a peak. Our findings suggest that landscape navigability may have contributed to the enormous success of transcriptional regulation as a source of evolutionary adaptations and innovations.

An adaptive landscape is a mapping from a high-dimensional space of genotypes onto fitness or some other related quantitative phenotype, which defines the 'elevation' of each coordinate in genotype space[1]. Evolution can be viewed as a hill-climbing process in an adaptive landscape, where populations tend to move towards peaks as a consequence of natural selection. The ruggedness of an adaptive landscape has important evolutionary consequences, particularly for the evolution of sex, reproductive isolation and mutational robustness, and for the predictability of evolution[2]. An adaptive landscape that is smooth and single peaked does not pose any obstacle to evolutionary exploration. It is therefore highly navigable, in that it is possible to reach the global peak via positive selection through a series of small mutations that only move 'uphill'. In contrast, a rugged landscape can block the approach to the highest peak by entrapping populations on local suboptimal peaks[3].

We know very little about the navigability of empirical adaptive landscapes, largely due to the incompleteness of the landscapes that have been constructed to date. With few exceptions[4,5], these landscapes were built by assaying the phenotypes of only a small number of mutations in all possible combinations within a single wild-type background[2]. These studies have helped form our intuition about the structure and navigability of empirical adaptive landscapes, but their conclusions are limited by the fact that they describe only a minute fraction of any complete landscape. An additional caveat of earlier studies is their focus on just one or a few landscapes, which limits the generality of their findings.

To study the navigability of a large number of complete, empirical adaptive landscapes, we consider data that describe the binding affinity of a transcription factor (TF)—a sequence-specific DNA-binding protein that helps regulate gene expression—to all possible DNA sequences (TF binding sites) of eight nucleotides in length. TFs are fundamental mediators of gene expression and are involved in numerous evolutionary innovations[6]. Their regulatory effect can be modulated via mutations in TF binding sites, which may alter a TF's affinity for a site and thereby affect gene expression[7–9]. We describe the mapping of DNA sequence to binding affinity as an adaptive landscape, where we can study selection for TF binding. This is a common approach for exploring the evolution of TF binding sites[10–14], other protein–DNA interactions[4,15,16] and protein–RNA interactions[17]. In this context, adaptive evolution is an exploration of sequence space that attempts to optimize the capacity of a sequence to bind a particular TF.

Results

Adaptive landscapes of TF binding affinity. We obtained protein-binding microarray (PBM) data for 1,137 TFs from the UniPROBE[18] and CIS-BP[19] databases. These TFs represent 129 eukaryotic species and 62 different DNA-binding domain structural classes (Supplementary Table 1). For each TF, we construct an adaptive landscape from the enrichment score (*E*-score)—a proxy for relative binding affinity[20–22]—of each of the 32,896 possible sites that bind the TF ('Methods'). These landscapes are complete because they describe the affinity with which a TF binds all possible sites, in the absence of confounding factors such as epigenetic marks, chromatin context, local sequence context or interactions with protein partners. We consider a sequence as 'bound' if its *E*-score exceeded 0.35[20,22–24] ('Methods'). We use this binding affinity threshold (τ) to differentiate sequences that are specifically bound by a TF via hydrogen-bond donors and acceptors from those that are unspecifically bound by a TF, for example, via its affinity for the DNA backbone. To facilitate the analysis of these landscapes, we represent each of them as a genotype network[25], in which vertices represent bound DNA sequences and edges connect sequences that differ by a single point mutation or a short insertion/deletion[22] (Fig. 1a; 'Methods'). These networks sometimes comprise multiple disconnected components (Supplementary Section 3.1 and Supplementary Figs 1–3); when this occurs, we consider only the largest component, which we refer to as the dominant genotype network. Each dominant genotype network forms the basis of an adaptive landscape, in which binding affinity defines the 'elevation' of each coordinate (TF binding site) in genotype space.

[1]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. [2]Swiss Institute of Bioinformatics, Quartier Sorge - Bâtiment Génopode, 1015 Lausanne, Switzerland. [3]The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA. †These authors contributed equally to this work. *e-mail: andreas.wagner@ieu.uzh.ch

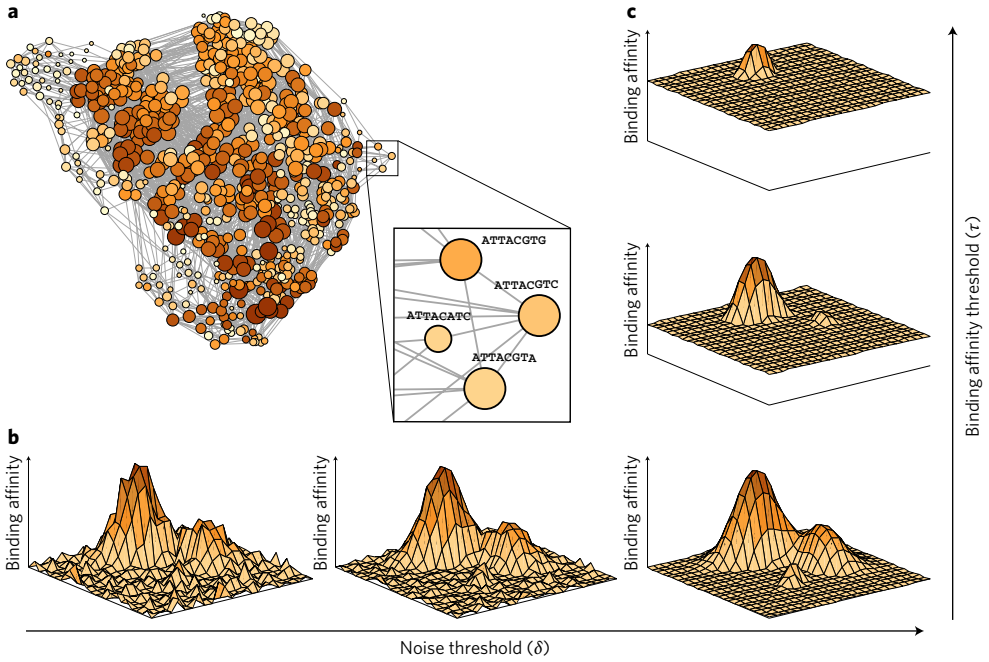


Figure 1 | Adaptive landscapes of TF binding affinity. a, The largest connected component of the genotype network for the yeast TF Gcn4, visualized using a force-directed algorithm. Each vertex corresponds to a DNA sequence bound by Gcn4 (*E*-score > 0.35). The colour of a vertex indicates its binding affinity (darker = higher), that is, the 'elevation' of the landscape, whereas vertex size corresponds to the number of neighbouring sequences (bigger = more). The inset shows that two vertices are connected by an edge if their corresponding sequences are separated by a single small mutation ('Methods'). **b**, To determine whether two sequences differ in their binding affinity, that is, if they differ in their 'elevation', we used a noise threshold, δ, which accounts for experimental noise in the PBM data. Increasing δ increases the 'smoothness' of the landscape, as shown in the three hypothetical landscapes at the bottom of the figure, where the *x*–*y* plane represents genotype space, and the *z* axis represents binding affinity. **c**, To delineate bound from unbound sequences we used a binding affinity threshold, τ. As τ increases, the number of bound sequences decreases, thus reducing the size of the landscapes and pruning local peaks, as shown in the three hypothetical landscapes at the right of the figure. In Supplementary Sections 3.6.1 and 3.6.2, we assess the sensitivity of our main results to broadly varying δ and τ values.

We define landscape navigability as the ability to access a global peak via an evolutionary exploration involving random mutation and natural selection. Landscape navigability is highest when all mutational paths to the global peak exhibit a monotonic increase in binding affinity, which implies a landscape that is smooth and single peaked. Landscape navigability is lowest when no mutational paths to the global peak exhibit a monotonic increase in binding affinity. This implies a rugged landscape with many peaks. Our measures of landscape navigability depend on two parameters, the noise threshold (δ) and τ, which we use for noise filtering and for delineating bound from unbound sequences, respectively (Fig. 1b,c; 'Methods'). We compare the navigability of the 1,137 empirical landscapes to landscapes generated via two different null models that provide lower and upper bounds on navigability ('Methods'). In the first null model (the additive model), we deterministically assign a binding affinity to each of the TF binding sites in the genotype network using the position weight matrix (PWM) of the TF, which assumes additive interactions between nucleotides and therefore produces smooth and highly navigable landscapes. In the second null model (the shuffled model), we randomly permute the affinities of the TF binding sites in the genotype network, which yields a rugged landscape that hinders navigability. Due to the stochastic nature of the shuffled model, we generate 1,000 shuffled landscapes for each TF. These null models capture two opposing extremes of landscape navigability while maintaining the structure of the underlying genotype network and therefore provide two points of reference for the empirical landscapes.

Landscape navigability: the number of peaks. The number of peaks in an adaptive landscape is an important indicator of its navigability. The more peaks a landscape has, the less navigable it becomes, if the peaks are of unequal height. We find that 42% of the empirically derived landscapes (478 of 1,137) have multiple peaks of unequal height (Supplementary Figs 4 and 5), with peak numbers ranging from 2 to 36 (Fig. 2a). In comparison, only 4.5% of the additive landscapes (51 of 1,137) have multiple peaks, whereas 99% of the shuffled landscapes (1,125,800 of 1,137,000) have multiple peaks (Fig. 2a). One might think that larger landscapes with more binding sites also contain more peaks and earlier theoretical work hints at that possibility[3]. However, the experimental data we analyse show no such scaling relationship (Supplementary Fig. 6a). Thus, by the criterion of peak number, smaller landscapes are not necessarily more navigable than larger landscapes. Also, while one usually thinks of a peak as a single sequence, the data do not support this notion. Except for 36 landscapes, the global peaks—those containing the highest-affinity site—are plateaus containing between 2 and 121 sequences (Supplementary Section 3.2 and Supplementary Figs 5 and 7).

Landscape navigability: epistasis. Epistasis[26]—non-additive effects of different mutations on a quantitative phenotype (or fitness)—can increase the ruggedness of an adaptive landscape. In the absence of epistasis, adaptive landscapes are smooth and single peaked, and thus do not hinder evolutionary exploration. Epistasis can be partitioned into three different classes—magnitude, simple sign and reciprocal sign—with increasingly detrimental effects on landscape navigability (Box 1)[27]. To study epistasis, we first identify all squares in a genotype network—a binding site, two of its one-mutant neighbours and the double mutant that can be formed from the single mutants—and classify their affinity relationships according to the

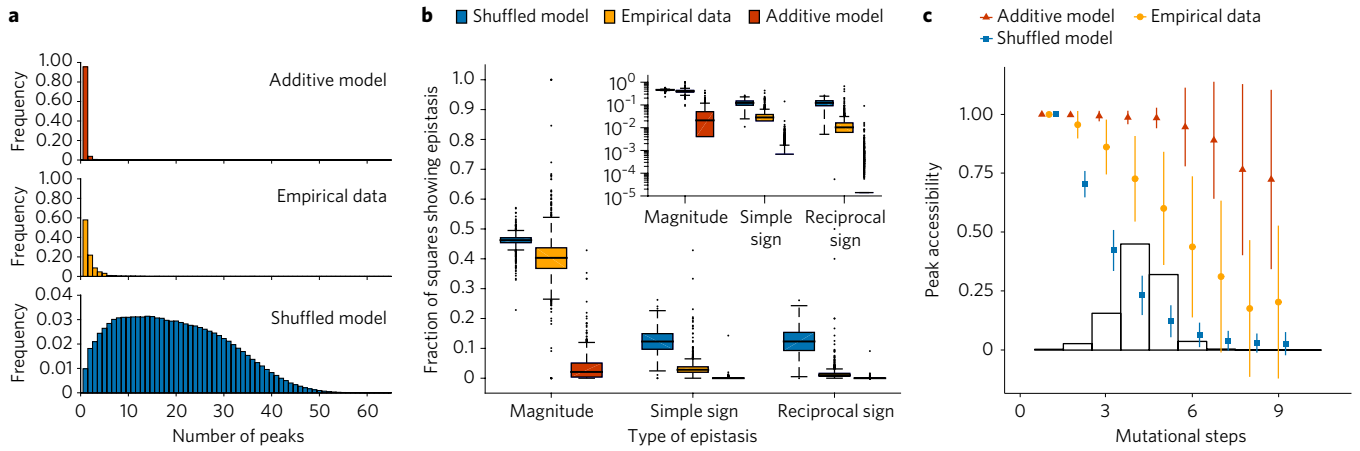


Figure 2 | The navigability of adaptive landscapes of TF binding affinity. a, The distribution of the number of peaks for the 1,137 empirical adaptive landscapes (yellow) and the additive (red) and shuffled (blue) null models. **b**, Boxplots of the fraction of squares showing magnitude, simple sign and reciprocal sign epistasis for the 1,137 adaptive landscapes. The thick horizontal line in the middle of each box represents the median of the data, while the bottom and top of each box represent the 25th and 75th percentiles, respectively. For the shuffled model, the boxplot summarizes 1,137 data points, each of which is an average of 1,000 shuffled landscapes. All pairwise differences are significant (paired *t*-tests: *P* value $< 2.2 \times 10^{-16}$). The inset shows the same data, but with a logarithmically scaled *y* axis, which obscures the following numbers of data points with zero epistasis: magnitude epistasis (empirical, 9; additive, 184), simple sign epistasis (shuffled, 2; empirical, 53; additive, 681) and reciprocal sign epistasis (shuffled, 1; empirical, 98; additive, 853). The absence of the thick horizontal line indicates that the median of the distribution is below the lowest value of the logarithmically scaled *y* axis. **c**, For each of the 1,137 adaptive landscapes we show the mean (symbols) and standard deviation (error bars) of the fraction of accessible paths to the highest-affinity site in the landscape (that is, the peak accessibility). The histogram shows the distribution of mutational steps to the highest-affinity site for all sequences in all landscapes. For the shuffled model, each symbol represents the mean of 1,137 data points, each of which is an average of 1,000 landscapes.

scheme in Box 1 ('Methods'). In the additive landscapes, on average, no more than 0.1% of squares show either of the two kinds of epistasis that impede landscape navigability most severely (simple or reciprocal, collectively referred to as sign epistasis) (Supplementary Section 3.3 and Supplementary Fig. 8). In the experimental data, sign epistasis is more frequent and affects 4.7% of squares, on average (Fig. 2b). However, its incidence is still five times lower than in the shuffled landscapes, where it affects 24.5% of the squares. In the experimental data only magnitude epistasis, which does not affect landscape navigability, approaches the levels observed in the shuffled landscape. In addition, sign epistasis preferentially occurs among nearby nucleotides in a binding site[28], whereas magnitude epistasis shows no such preference (Supplementary Section 3.4 and Supplementary Fig. 9).

Landscape navigability: accessible mutational paths. Another important indicator of landscape navigability is the fraction of accessible mutational paths to a given genotype from all other genotypes in the landscape (Supplementary Fig. 10). Here, a mutational path is considered accessible if each mutation in the path increases binding affinity monotonically[29]. Figure 2c shows the fraction of accessible paths to the highest-affinity binding site in a global peak (that is, peak accessibility), in relation to the length of the mutational path, for the empirical data and both null models. In all three cases, the fraction of accessible mutational paths to the highest-affinity site decreases with the length of the path. However, the rate of decrease for the empirical data is intermediate to that of the additive and shuffled models, indicating that the empirical landscapes are always less navigable than the additive landscapes, but more navigable than the shuffled landscapes. Even for the longest mutational paths to the highest-affinity site, more than 20% of the paths are accessible in the empirical landscapes. Moreover, even when unbound sequences are included in the landscapes—a modification that decreases landscape navigability—the global peak remains accessible for all but the longest mutational paths (Supplementary Fig. 11 and Supplementary Section 3.5).

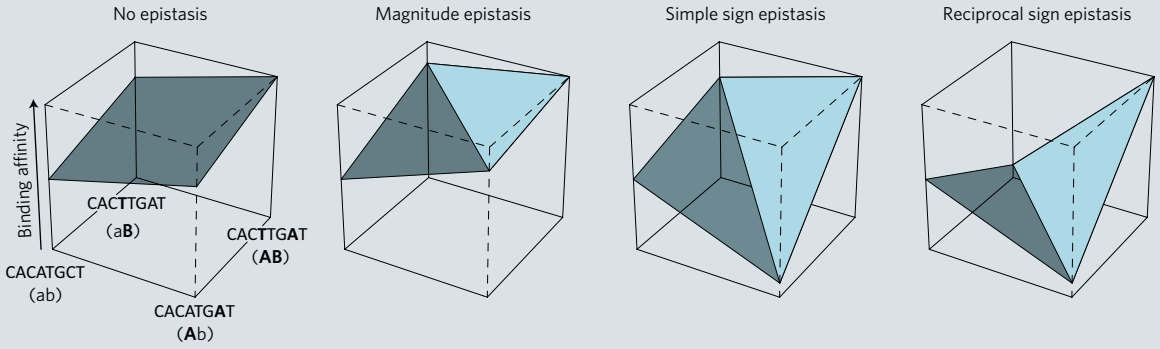
Taken together, these three measures of landscape navigability—number of peaks, epistasis and peak accessibility—indicate that transcription factor binding affinity landscapes are more navigable than shuffled landscapes, but less navigable than additive landscapes. This conclusion is robust to broadly varying parameter choices (δ and τ) and modelling assumptions (Supplementary Section 3.6 and Supplementary Figs 12–30). Although experiments are required to determine which parameters and assumptions best reflect the true binding affinity landscapes, the following *in vivo* analyses suggest that the baseline parameter combination studied here provides meaningful information about TF binding in both yeast and mouse, two highly diverged eukaryotic species.

Navigability influences the *in vivo* abundance of binding sites. Landscape navigability varies among TFs within the same species. This led us to reason that global peak sequences from more navigable landscapes might be more abundant in the regulatory regions of living organisms than global peak sequences from less navigable landscapes. The reason is that smooth landscapes pose fewer obstacles to the evolution of global peak sequences than rugged landscapes. To test this hypothesis, we consider two sources of *in vivo* data from 14 cell and tissue types in *Mus musculus*: RNA sequencing (RNA-seq) transcript abundance estimates[30] and maps of genome-wide DNase I footprints[31]. The RNA-seq data indicate which TFs are expressed in each of the 14 cell and tissue types, which is important because we only expect landscape navigability to impact the *in vivo* abundance of a binding site if its cognate TF is expressed in that cell or tissue type. The DNase I footprints demarcate DNA sequences in open chromatin that are bound by protein[32] ('Methods') and can therefore be used to predict TF binding sites.

For the 187 murine TFs in our dataset that are expressed in a given cell or tissue type ('Methods'; Supplementary Table 2), we determine the *in vivo* abundance of the TF's highest-affinity binding site by counting the number of times the site appears in DNase I footprints that are predicted to bind the TF. We determine the statistical significance of each count by comparing it with the number

Box 1 | Epistasis.

A square in a genotype network connects a 'wild type' sequence (ab) to a double mutant (AB) through two single mutants (Ab and aB). The left-most panel shows an example with binding sites of length eight where there is no epistasis: the binding affinity of the double mutant is simply the addition of the affinity contributions of the two single mutants. That is, the mutation a to A has the same effect on binding affinity in the two different genetic backgrounds (b and B). Magnitude epistasis occurs when the magnitude (but not the sign) of a mutation's effect on binding affinity depends on the genetic background. Simple sign epistasis occurs when one single mutant has a lower binding affinity than both the wild type and the double mutant, while the other single mutant has an affinity that is intermediate to the wild type and double mutant[60]. Reciprocal sign epistasis occurs when both mutations decrease affinity independently, but increase affinity in combination[27]. To account for the noise in our data, we only considered a single mutant's binding affinity to be lower than that of the wild type if the affinity difference exceeded a threshold value δ, which was derived from the empirical data and was specific to each TF ('Methods'). For all squares, we quantified epistasis along a single axis of the square by designating the highest-affinity sequence as the double mutant ('Methods'). Figure adapted from ref. [55], Macmillan Publishers Ltd.



of times the sequence is expected to appear in stretches of DNA that have the same length and the same mono- and di-nucleotide frequencies as the footprints ('Methods'). We find that the highest-affinity sites in landscapes with multiple peaks are less abundant in regulatory regions genome-wide than those from landscapes with a single peak across all of the 14 cell and tissue types (Supplementary Fig. 31 and Supplementary Table 3), as shown for heart tissue in Fig. 3a (Wilcoxon rank-sum test, P value $= 1.42 \times 10^{-6}$). We also observe that highly accessible global peak sequences are more abundant in protein-bound regions of the mouse genome than less

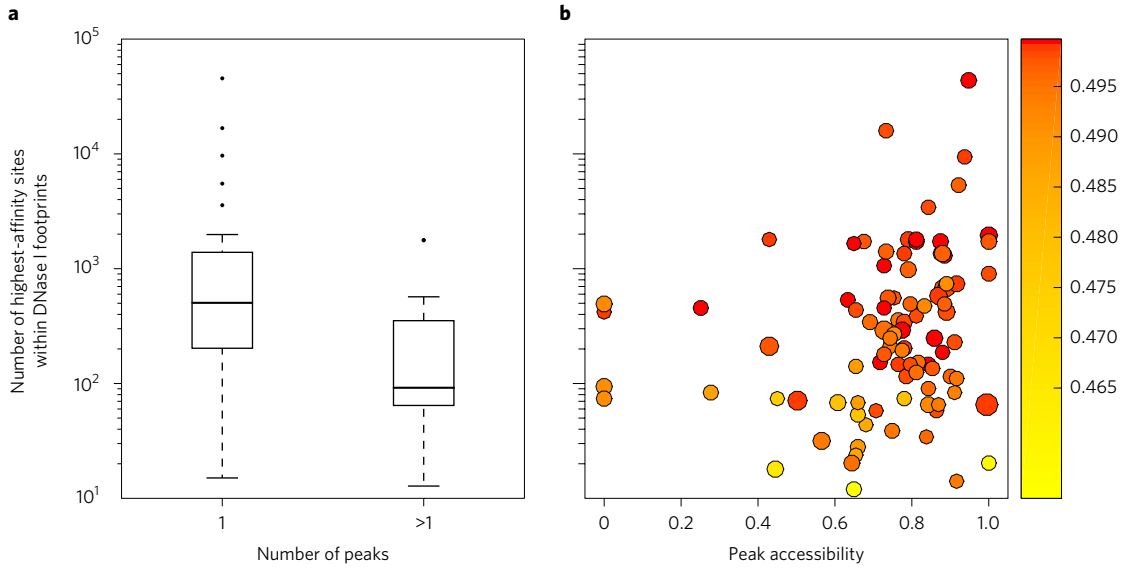


Figure 3 | ***In vivo*** **binding site abundance correlates with landscape navigability. a**,**b**, The vertical axis of each panel indicates the abundance of a TF's highest-affinity site in protein-bound regions of the *M. musculus* genome in heart tissue (according to DNase I footprint data; 'Methods'). In panel **a**, the horizontal axis indicates the number of peaks and classifies landscapes into single-peaked and multi-peaked categories. Global peak sequences from single-peaked landscapes are more abundant than those from multi-peaked landscapes (Wilcoxon rank-sum test, P value $= 1.42 \times 10^{-6}$). The thick horizontal line in the middle of each box represents the median of the data, while the bottom and top of each box represent the 25th and 75th percentiles, respectively. In panel **b**, the horizontal axis shows peak accessibility through mutational paths of length four (Spearman's rank correlation coefficient $= 0.27$, P value $= 9.2 \times 10^{-3}$), which are the most abundant paths in our dataset (Fig. 2c). Each circle corresponds to a single TF expressed in heart tissue. Circle colour indicates the binding affinity of the TF's highest-affinity site (darker = higher; colour bar). Circle size corresponds to the TF's expression level (larger = higher, 'Methods'). Note the logarithmic scale of the y axis for both panels.

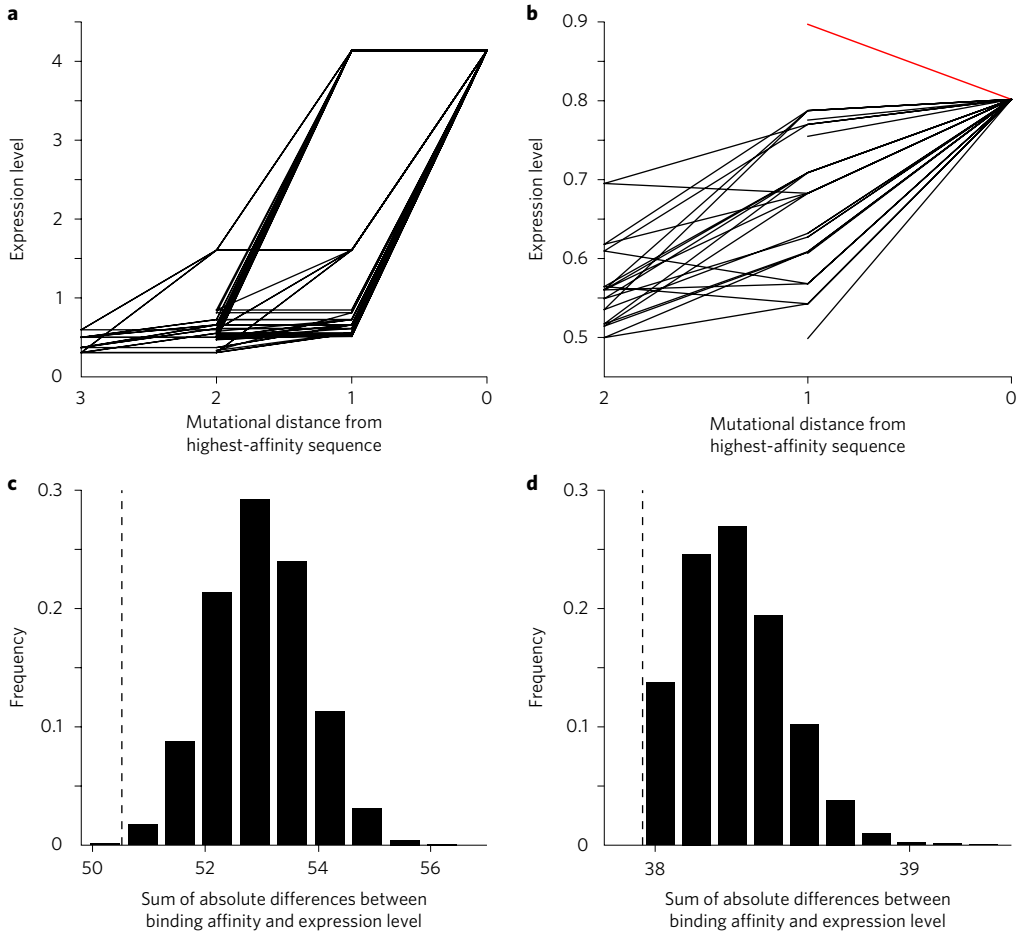


Figure 4 | Gene expression increases along accessible mutational paths and reflects landscape topography. a,**b**, Each line shows the expression levels of the sequences in an accessible mutational path to the highest-affinity sequence for the yeast TFs Gcn4 (**a**) and Fhl1 (**b**). Black lines denote accessible mutational paths in which gene expression increases monotonically. The red line denotes the single accessible mutational path in which gene expression does not increase monotonically. Note that although expression appears to decrease along some of the black lines in panel **b**, it does not decrease beyond the noise threshold (δ) necessary to classify the trend as a true expression decrease ($\delta = 0.7$ for Gcn4 and $\delta = 0.1$ for Fhl1; 'Methods'). Note also that the highest-affinity sequence does not necessarily correspond to the sequence with the highest level of expression (red line in panel **b**). **c**,**d**, Vertical dashed lines indicate the observed sum of the absolute differences between binding affinity and gene expression, and the black bars show the null distribution of this sum for 10^5 random permutations of the gene expression values for the TFs Gcn4 (**c**) and Fhl1 (**d**).

accessible global peak sequences across 10 of the 14 cell and tissue types (Fig. 3b, Supplementary Table 3 and Supplementary Fig. 32). As a negative control, we repeat the above analyses using the DNase I hypersensitive regions that flank the footprints, rather than the footprints themselves. For this control data, the abundance of the highest-affinity sites is not significantly associated with the number of peaks in any of the 14 cell and tissue types (Supplementary Table 4). Also, for the control data and in 13 of the 14 tissue types, the abundance of the highest-affinity sites is not significantly associated with peak accessibility; in the remaining tissue (heart), this association is only marginally significant (Spearman's rank correlation coefficient = 0.27, P value = 0.034; Supplementary Table 4). Importantly, the effects of landscape navigability on binding site abundance still hold after controlling for binding affinity (Supplementary Section 3.7 and Supplementary Tables 5 and 6), although statistical significance is lost in three tissues for peak accessibility. Moreover, our observations hold in all cell and tissue types after controlling for the information content of each TF's PWM (Supplementary Section 3.7 and Supplementary Tables 5 and 7). Taken together, these findings suggest that landscape navigability has influenced the evolution of TF binding sites in the mouse genome.

Gene expression reflects landscape topography. Gene expression levels can be fine-tuned via affinity-altering mutations in TF binding sites[7,9]. Models of regulatory evolution commonly assume a direct mapping between binding affinity and gene expression[33,34], such that monotonic changes in binding affinity lead to monotonic changes in gene expression. We test this assumption with *in vivo* gene expression data from a recent high-throughput promoter screen in *Saccharomyces cerevisiae*[9]. These data comprise replicated *in vivo* gene expression measurements for every single-base-pair and many double- and triple-base-pair mutants of a TF's consensus binding site ('Methods'). PBM data are available for two of these TFs (Gcn4 and Fhl1), facilitating the superposition of *in vivo* transcriptional output with *in vitro* binding affinity. For the subset of each TF's genotype network where *in vivo* expression data are available, we determine whether gene expression levels increase monotonically along accessible mutational paths to the site with the highest affinity. For Gcn4, all 71 accessible mutational paths exhibit monotonic increases in gene expression ($P = 0.01$, permutation test; Fig. 4a) and for Fhl1 all except one of the 37 accessible mutational paths exhibit monotonic increases in gene expression ($P = 0.001$, permutation test; Fig. 4b). These findings suggest that, at least for these two TFs, the navigability

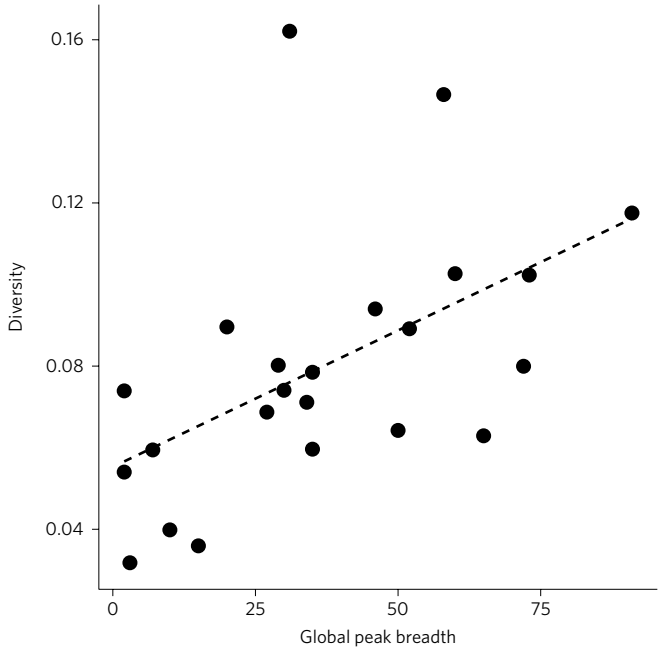


Figure 5 | Global peak breadth influences the diversity of TF binding sites in the yeast genome. Scatter-plot for 23 TFs showing the relationship between the number of binding sites in the global peak (that is, global peak breadth) of a particular TF and the average diversity of all its polymorphic binding sites in 19 strains of *S. cerevisiae* ('Methods'; Spearman's rank correlation coefficient = 0.61, *P* value = 0.002). The dashed line represents the best linear-regression fit to the data.

of the adaptive landscape of binding affinity facilitates the evolution of increased gene expression.

We also compare the binding affinity landscapes of Gcn4 and Fhl1 to incomplete landscapes constructed using gene expression data from the high-throughput promoter screens. To do this, we develop a measure of similarity between the complete *in vitro* landscapes and the incomplete *in vivo* landscapes. This measure is simply the sum of the absolute differences between binding affinity and gene expression for each of the binding sites. We reason that if the *in vitro* binding affinity landscapes are truly informative of *in vivo* gene expression, then this sum will be significantly smaller than expected, if instead the gene expression measurements are randomly permuted. To test this hypothesis, we compare the observed sum to a null distribution of sums obtained via 10^5 random permutations of the gene expression data. The fraction of permuted datasets in which the sum is smaller than that of the measured expression data yields an empirical *P* value for each TF. Based on this test, *in vitro* binding affinity is indeed informative of *in vivo* gene expression (Gcn4: *P* value = 0.0011; Fhl1: *P* value = 0.0125; Fig. 4c,d), which provides additional validation of the binding affinity landscapes studied here.

Global peak breadth affects the diversity of binding sites. We have shown that global peaks typically comprise many different binding sites of similar affinity. The broader a peak is, the more sequences it has that have mutant neighbours with approximately the same affinity. One would thus expect that sequences in broader peaks could evolve more freely by means of nucleotide changes and thus accumulate greater genetic diversity. In contrast, mutations in the binding sites of narrower peaks—those with fewer binding sites—will more often lead to a decrease in binding affinity and thus be eliminated. To find out if global peak breadth has any effect on binding site evolution, we analyse whether global peak breadth affects genetic diversity in binding sites. Using single nucleotide polymorphism data across 19 strains of *S. cerevisiae*[35], we calculate the diversity within binding sites[36] of 23 different TFs as the average Shannon diversity index per site ('Methods'). Indeed, the broader the global peak of a TF's affinity landscape, the greater the average diversity of polymorphic sites that bind the TF (Spearman's rank correlation coefficient = 0.61, *P* value = 0.002; Fig. 5). We emphasize that this trend is not driven by TF specificity, since the information content of the TFs' PWMs exhibit no correlation with binding site diversity (Spearman's rank correlation coefficient = −0.08, *P* value = 0.71). This finding provides further validation that the topography of a binding affinity landscape can impact the evolution of TF binding sites.

Discussion

We have used measurements from PBMs to construct and analyse more than 1,000 complete, empirical adaptive landscapes, each describing the binding affinity of a TF to all possible short DNA sequences. Such landscapes are important objects of study, because changes in the level, location or timing of gene expression commonly underlie evolutionary innovations[37,38] and gene expression patterns are readily fine-tuned via small changes in binding affinity[7,9]. Understanding how an evolutionary process might navigate a TF binding affinity landscape is therefore an important step towards understanding how gene regulatory programs evolve. Here, we have taken this step, demonstrating that the navigability of binding affinity landscapes is intermediate to that of additive and shuffled (rugged) landscapes, but closer to the additive expectation, in terms of the number of peaks and the incidence of three forms of epistasis.

Our measures of landscape navigability allow us to understand how individual binding sites can evolve towards a higher binding affinity in the absence of confounding factors (such as chromatin context[39]), a pursuit that is motivated by earlier theoretical work[10,13,33] and by several empirical observations, which indicate that high-affinity sites are used preferentially *in vivo*. For example, studies of blastoderm patterning in *Drosophila melanogaster* have shown that high-affinity sites typically reside near probable functional targets, whereas low-affinity sites are more often found near genes that are not transcribed in the early embryo[40] and that are less likely to drive expression in transgenic reporter assays[41]. Moreover, in both microbes and humans, affinity-decreasing mutations are predominately under negative selection, whereas affinity-increasing mutations are under positive selection[42,43], consistent with the inferred monotonic increases in organismal fitness that accompany increased binding affinity[13,14,44]. Nevertheless, low-affinity sites do sometimes play important regulatory roles[45–48] and it is therefore worth noting that, by symmetry, accessible paths to the highest-affinity site are also accessible in the opposite direction by permitting only monotonically decreasing changes in binding affinity. More generally, the landscapes constructed here can easily be transformed to study selection for low or intermediate binding affinity, by assuming that the fitness conferred by a binding site is a decreasing function of the difference between the site's affinity and an arbitrary optimal affinity[33]. Understanding how such transformations affect the navigability of TF binding affinity landscapes is an exciting direction for future work.

To summarize, while our analyses of *in vitro* and *in vivo* measurements of TF–DNA interactions have some caveats (Supplementary Section 4), they suggest that the navigability of TF binding affinity landscapes has left a trace in the portfolios of regulatory DNA within the mouse and yeast genomes. Landscape navigability may therefore have contributed to the enormous success of transcriptional regulation as an evolutionary mechanism for generating variation and innovation.

Methods

***In vitro* data.** The *in vitro* data we studied came from PBMs[21,49], which measure the binding affinity of a TF to all 32,896 possible eight-nucleotide, double-stranded DNA sequences. There were $(4^8 - 4^4) / 2 + 4^4 = 32{,}896$ sequences,

rather than $4^8 = 65{,}536$ sequences, because each sequence was merged with its reverse complement and because there were 4^4 sequences that were identical to their reverse complement and therefore could not be merged.

We had three criteria for including a TF in our dataset. First, it had to be analysed on two different PBM designs. Second, it had to bind at least one DNA sequence with an *E*-score above 0.45, as this indicates a high level of data quality[19]. Third, its genotype network (Fig. 1a) had to contain at least one square to permit the analysis of epistasis (Box 1). Based on these criteria, we obtained PBM data for 42 yeast[23] and 104 mouse TFs[20] from the UniPROBE database[18] and 991 TFs belonging to 129 different eukaryotic species (including 4 additional TFs for *S. cerevisiae* and 83 for *M. musculus*) from the CIS-BP database[19]. In total, our dataset comprised 1,137 TFs, representing 129 eukaryotic species and 62 DNA-binding domain structural classes (Supplementary Table 1).

For each TF, the PBM data included a non-parametric, rank-based *E*-score for each of the 32,896 DNA sequences. The *E*-score is a variant of the Wilcoxon–Mann–Whitney statistic[21] and ranges from −0.5 (most disfavoured site) to 0.5 (most favoured site). *E*-scores correlate with the relative dissociation constants of TFs[20,21] and can be used as a proxy for relative binding affinity. We therefore refer to this measure as binding affinity and used it to delineate bound from unbound sequences and as the quantitative phenotype that defines the surface of our adaptive landscapes. Except for the 42 yeast TFs from Zhu *et al.*[23], the data also included a median signal intensity *Z*-score for each site. We used this score as an alternative proxy for binding affinity in Supplementary Section 3.6.5. Following earlier work[22,23], we only considered a sequence as bound by a TF if its *E*-score exceeded a threshold of 0.35. The reason for this threshold was that it has precedent[20,24] and, more importantly, an analysis of the relationship between the *E*-score and false discovery rate (FDR) in 104 mouse TFs[20] revealed that all sequences with an *E*-score exceeding 0.35 had an FDR below 0.001. This threshold could therefore be used to delineate sequences that are specifically bound by a TF from those that are unspecifically bound. In addition to the threshold $\tau = 0.35$, we conducted a sensitivity analysis of our results by broadly varying τ (Supplementary Section 3.6.2).

***In vivo* data.** We collected DNase I footprints for 14 cell and tissue types in *M. musculus*[31]. These genome-wide data specify DNA sequences that are in open chromatin and bound by protein, at single nucleotide resolution. For each of the 187 murine TFs in our dataset, we used FIMO[50] and the TF's PWM (obtained from UniPROBE[18] and CIS-BP[19]) to scan these footprints for potential binding sites using $P < 1 \times 10^{-4}$ as a threshold. We then counted the number of times that each eight-nucleotide DNA sequence appeared in the predicted binding sites, considering both strands of the DNA. To determine the statistical significance of each count, we compared it with the number of times the same sequence is expected to appear in stretches of DNA that have the same length and mono- and di-nucleotide frequencies as the footprints of each cell and tissue type, thus controlling for the GC content of the footprints. Following van Helden *et al.*[51], the statistical significance of each observed count was determined using the binomial formula, with a conservative significance threshold of $P < 1 / 32{,}896$ (the inverse of the number of possible eight-nucloetide DNA sequences). Across the 14 cell and tissue types, the counts for the highest-affinity sites of each of the 187 mouse TFs in our dataset ranged from 0 to 73,174.

We also collected the DNase I hypersensitive regions that flank the footprints in each of the 14 cell and tissue types. We used these regions to perform a negative control, in which we correlated our landscape navigability measures with *in vivo* binding site abundance within regions of open chromatin that do not show evidence of protein binding. Specifically, we counted the number of times that each eight-nucleotide DNA sequence appeared in the DNase I hypersensitive regions, after having removed the footprints of the 14 cell and tissue types. To determine the statistical significance of each count, we compared it with the number of times the same sequence is expected to appear in stretches of DNA that have the same length and mono- and di-nucleotide frequencies as the DNase I hypersensitive regions of each cell and tissue type, again using the binomial formula and a stringent significance threshold ($P < 1 / 32{,}896$).

To determine which of the 187 TFs were expressed in each of the 14 cell and tissue types, we collected RNA-seq data for the same cell and tissue types[30] and used cufflinks[52] to calculate the FPKM (fragments per kilobase of transcript per million fragments mapped) for each TF. We considered a TF as expressed in a given cell or tissue type if its FPKM ≥ 1. Supplementary Table 2 shows that the number of TFs expressed in a given cell or tissue type ranged from 89 to 133. This table also provides the file names of the DNase I hypersensitive regions, DNase I footprints and the RNA-seq data that we analysed.

For two yeast TFs (Gcn4 and Fhl1), we also collected gene expression data from high-throughput promoter screens[9]. These data include gene expression measurements from a large library of engineered promoters, each 150 nucleotides long. Each of these promoters contains between zero and three point mutations to the TF's consensus sequence. For Gcn4, which binds a seven-nucleotide sequence, this library includes the consensus sequence (TGACTCA), all 21 single mutants, 42 double mutants and 10 triple mutants (73 sequences in total). For Fhl1, which binds an eight-nucleotide sequence, this library includes the consensus sequence (GACGCAAA), all 24 single mutants, 56 double mutants and 10 triple mutants (90 sequences in total). For each promoter in the library, the data include gene expression measurements from two biological replicates. We used the average gene expression measurement per promoter.

To map the gene expression measurements for the seven-nucleotide sequences bound by Gcn4 onto the eight-nucleotide sequences for which we had PBM data, we first located each of the 73 seven-nucleotide sequences within the 150-nucleotide promoters. We then padded each of the seven-nucleotide sequences with one nucleotide upstream and one nucleotide downstream from its respective promoter, forming two eight-nucleotide sequences. We assigned the gene expression measurement for the seven-nucleotide sequence to each of the eight-nucleotide sequences that were formed in this fashion. This procedure generated gene expression measurements for 132 eight-nucleotide sequences.

For the subset of the sequences for which we had measurements of both *in vitro* binding affinity and *in vivo* gene expression, we determined the number of accessible mutational paths to the sequence with the highest level of binding affinity. For Gcn4, there were 71 accessible mutational paths: 30 of length 3, 37 of length 2 and 14 of length 1. For Fhl1, there were 37 accessible mutational paths: 22 of length 2 and 15 of length 1. For each TF, we then determined the fraction of these accessible paths in which gene expression increased monotonically. To determine whether two gene expression levels truly differed from one another, we used a noise threshold that covered the same proportional range of expression levels as covered by the noise threshold δ in the range of affinity values. To calculate the statistical significance of the fraction of accessible paths in which gene expression increased monotonically, we performed a permutation test, in which we randomly permuted the gene expression measurements of the sequences, while preserving their binding affinities. We repeated this process 1,000 times for both Gcn4 and Fhl1.

We obtained the genomic coordinates of TF binding sites in *S. cerevisiae* from a map of conserved regulatory sites[36] and analysed binding sites detected with a stringent binding *P* value cut-off of 0.001, but no conservation cut-off. We obtained single nucleotide polymorphism (SNP) data for 19 different strains from the Saccharomyces Genome Resequencing Project (SGRP)[35]. The TF binding sites were based on the January 2006 Saccharomyces Genome Database (SGD) version of the reference strain *S. cerevisiae* S288c genome sequence, while the SNP data were based on the January 2010 version. We used liftOver to convert both sets of genomic coordinates to the coordinates in the February 2011 version. Then we used the *intersect* function from the BEDTools suite (version 2.25.0)[53] to determine the presence of SNPs within binding sites. Using these data, we calculated the genetic diversity of polymorphic TF binding sites across the 19 yeast strains. For each position j in a binding site we calculated the Shannon diversity index (H):

$$H_j = -\sum_i p_i \log_2(p_i) \quad (1)$$

where p_i is the frequency of allele i, which we computed as the fraction of strains with allele i. We computed the diversity (D) of a binding site as the average of H over all L positions:

$$D = \frac{1}{L}\sum_{j=1}^{L} H_j \quad (2)$$

Genotype networks. The procedure for constructing genotype networks of TF binding sites has been described elsewhere[22]. In brief, for each TF, we first determined the set of sequences that were bound by the TF (*E*-score > 0.35). We then used an alignment algorithm to calculate the mutational distance between all pairs of bound sequences. Finally, we used these mutational distances to define the edges of the genotype network, connecting two sequences if they differed by a single small mutation. The mutations we considered were point mutations and small indels that shift an entire contiguous binding site by a single base[22].

Quantitative measures of landscape navigability. We used several measures to quantify the navigability of an adaptive landscape. All of them were parameterized by δ, a threshold value that is used to determine whether two affinity values truly differ from one another. This parameter is necessary because PBM data are inherently noisy[49]. For each TF, we calculated δ as the residual standard error of a linear regression between the affinity values of all bound sequences from the two replicate PBMs. Thus, each TF had its own δ, which reflected the noise in the replicated PBM measurements for that particular TF (Supplementary Table 1). We considered that the binding affinity E_i of site i was greater than the binding affinity E_j of site j if $E_i > E_j + \delta$. Analogously, we considered that the binding affinity E_i of site i was less than the binding affinity E_j of site j if $E_i + \delta < E_j$. Otherwise, we considered that we could not differentiate between the two affinity values. The average empirical value of δ across the 1,137 TFs was 0.028, which covered 18.4% of the range of affinity values for bound sequences ($0.35 < E\text{-score} \leq 0.50$). Our criterion for considering two affinity values as different was therefore highly conservative. In addition to these empirical values of δ, we conducted a sensitivity analysis of our results by broadly varying δ (0, 0.001, 0.01, 0.03 and 0.05) above and below the average empirical value (Supplementary Section 3.6.1).

Our first measure of the navigability of an adaptive landscape was its number of peaks. To detect a peak, we followed a procedure similar to one previously described[54]. We categorized each sequence in a genotype network as belonging to a peak (either as an individual sequence or as a member of a plateau) or not. To do this, we selected sequences in decreasing order of binding affinity to seed a breadth-first search of the genotype network. Each iteration of the search considered sequences that were an additional mutational step away from the seed sequence. In the first iteration of the search, we determined whether the seed sequence was a peak or not; it was considered a peak if all of its neighbours in the genotype network had a lower binding affinity and was not a peak if at least one of its neighbours had higher binding affinity. If the seed sequence had at least one neighbour with an affinity that was neither greater nor lower, then these neighbours were retained as belonging to a plateau that may be a peak and the breadth-first search continued. If at any subsequent iteration of the search a sequence was found that neighboured any of those on the plateau and had an affinity that was higher than the seed sequence, then the seed sequence did not belong to a peak and the search was halted. If at any iteration of the search, all sequences that neighboured those on the plateau had a lower affinity than the seed sequence, then the seed sequence belonged to a plateau that was a peak.

Our second measure of landscape navigability was based on the concept of accessible mutational paths[29,55,56]. A mutational path is the shortest path that connects two bound sequences, *i* and *j*, on a genotype network, such that sequence *i* can be transformed into sequence *j* via a series of intermediates that are also on the genotype network. A mutational path from sequence *i* to sequence *j* is considered accessible if binding affinity increases monotonically along the path. We report the fraction of mutational paths that are accessible, starting from all sequences in the genotype network and ending at the highest-affinity site in the global peak. We refer to this fraction as peak accessibility. We note that, by symmetry, accessible paths to the highest-affinity site are also accessible in the opposite direction by permitting only monotonically decreasing changes in binding affinity. Accessibility of low-affinity binding sites is important, as such sites are known to play crucial roles in the regulation of certain genes[46,48,57].

Our final measures of landscape navigability pertained to epistasis, motivated by a recent debate over the significance of non-additive interactions between the individual bases of TF binding sites in their contribution to binding affinity[58,59]. Epistasis can have detrimental effects on landscape navigability. We quantified the incidence of three classes of epistasis for each of the 1,137 adaptive landscapes[27]. First, we detected all squares in each TF's genotype network. A square is a quadruplet of sequences that contain a binding site, two of its one-mutant neighbours, and the double mutant that can be formed from the single mutants (Box 1). Second, we designated the highest-affinity sequence as the 'double mutant', thus forcing the labelling of the other three sites as the wild type or single mutants. Third, we calculated the magnitude of epistasis as:

$$\varepsilon = E_{\mathrm{AB}} + E_{\mathrm{ab}} - E_{\mathrm{Ab}} - E_{\mathrm{aB}} \tag{3}$$

where E_{AB} is the binding affinity of the double mutant, E_{ab} is the binding affinity of the wild type and E_{Ab} and E_{aB} are the binding affinities of the single mutants. We only considered a mutational pair to be epistatic if $|\varepsilon|$ was greater than or equal to the noise threshold δ. If this condition was met, we classified the epistatic interaction as magnitude epistasis, simple sign epistasis or reciprocal sign epistasis[27]. The interaction was classified as magnitude epistasis when the following relation held:

$$|\Delta E_{\mathrm{ab}\rightarrow\mathrm{Ab}} + \Delta E_{\mathrm{aB}\rightarrow\mathrm{AB}}| = |\Delta E_{\mathrm{ab}\rightarrow\mathrm{Ab}}| + |\Delta E_{\mathrm{aB}\rightarrow\mathrm{AB}}| \tag{4}$$

where ΔE is the 'mutational effect', that is, the change in binding affinity caused by a mutation (for example, ab → Ab). The interaction was classified as simple sign epistasis when the following relation held:

$$|\Delta E_{\mathrm{ab}\rightarrow\mathrm{Ab}} + \Delta E_{\mathrm{aB}\rightarrow\mathrm{AB}}| < |\Delta E_{\mathrm{ab}\rightarrow\mathrm{Ab}}| + |\Delta E_{\mathrm{aB}\rightarrow\mathrm{AB}}| \tag{5}$$

The interaction was classified as reciprocal sign epistasis when both equation (3) and the following relation held:

$$|\Delta E_{\mathrm{ab}\rightarrow\mathrm{aB}} + \Delta E_{\mathrm{Ab}\rightarrow\mathrm{AB}}| < |\Delta E_{\mathrm{ab}\rightarrow\mathrm{aB}}| + |\Delta E_{\mathrm{Ab}\rightarrow\mathrm{AB}}| \tag{6}$$

If a mutational effect was smaller than δ, we assigned it a value of zero. If all mutational effects were smaller than δ, even despite $|\varepsilon| \geq \delta$ being true, then we classified the interaction as non-epistatic. Taken together with the fact that we excluded unbound sequences from these calculations, our measures of epistasis were conservative (Supplementary Section 3.8 and Supplementary Fig. 33).

Null models. For each of the 1,137 TFs, we considered two null models. Both changed the topography of the adaptive landscape, but maintained the structure of the underlying genotype network. The additive model was based on the PWM of a TF. A PWM represents the binding preferences of a TF as a $4 \times L$ matrix, where each row corresponds to one of the four bases and each column corresponds to one of the L positions in the binding site. Each matrix entry, $f_{i,b}$, is the frequency of base b at position i. We obtained the PWMs of the 1,137 TFs from the UniPROBE[18] and CIS-BP[19] databases.

Since the width L of a TF's PWM may not equal eight, we used a sliding-window to assign a score (S_{pwm}) to each of the eight-nucleotide sequences in each genotype network. Specifically, we slid each sequence in a genotype network through the corresponding TF's PWM from left to right, assigning a score to each of the subsequences. The sliding-window procedure was carried out in such a way that the first scored subsequence was the single right-most position of the sliding sequence and occupied the left-most column of the PWM. The procedure ended with a subsequence that corresponded to the single left-most position in the sequence and occupied the right-most column of the PWM. We then repeated this process with the sequence's reverse complement. We took the maximum of these scores as the score for the sequence. As an example, if L was equal to eight, then we took the maximum of 30 separate scores: 15 for the sequence and another 15 for its reverse complement. Each of these scores was calculated as

$$S_{\mathrm{pwm}} = \sum_{i=1}^{l} f_{i,\,b} I(i) \tag{7}$$

where l is the window length of the sliding sequence and $I(i)$ is the information content at position i:

$$I(i) = 2 + \sum_{b} f_{i,b} \log_2 f_{i,b} \tag{8}$$

Base matches at positions with high information content thus contribute more to a sequence's overall score than base matches at positions with low information content. Importantly, this scoring technique was purely additive, that is, the contribution of each binding site position to the overall score was independent of the other positions in the binding site. To facilitate comparison among TFs, we normalized the scores by the maximum score in each genotype network. To analyse the navigability of landscapes constructed with this model, we used a noise threshold δ that covers the same range of scores as the δ used for the empirical data.

In the shuffled model, we randomly permuted the binding affinities of the sequences in the genotype network, yielding a rugged landscape topography. Since this process was stochastic, we repeated it 1,000 times per TF. In these landscapes, we used the same δ as that used for the empirical landscapes, because, unlike the additive model, the shuffled null model does not change the range of affinity values.

Data availability. All data analysed during this study are available in public repositories; accession information is provided in Supplementary Tables 1 and 2. All data generated during this study are included in this published article (and its supplementary information files).

Received 6 June 2016; accepted 5 December 2016; published 23 January 2017

References

1. Wright, S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. in *Proc. Sixth Int. Congress Genetics Vol. 1* (ed. Jones, D. F.) 356–366 (The Genetics Society of America, 1932).
2. Szendro, I. G., Schenk, M. F., Franke, J., Krug, J. & de Visser, J. A. G. M. Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech-Theory E.* **2013,** P01005 (2013).
3. Kauffman, S. & Levin, S. Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.* **128,** 11–45 (1987).
4. Rowe, W. *et al.* Analysis of a complete DNA-protein affinity landscape. *J. R. Soc. Interface* **7,** 397–408 (2010).
5. Jiménez, J. I., Xulvi-Brunet, R., Campbell, G. W., Turk-MacLeod, R. & Chen, I. A. Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc. Natl Acad. Sci. USA* **110,** 14984–14989 (2013).
6. Wray, G. A. The evolutionary significance of *cis*-regulatory mutations. *Nat. Rev. Genet.* **8,** 206–216 (2007).
7. Gertz, J., Siggia, E. D. & Cohen, B. A. Analysis of combinatorial *cis*-regulation in synthetic and genomic promoters. *Nature* **457,** 215–218 (2009).
8. Shultzaberger, R. K., Malashock, D. S., Kirsch, J. F. & Eisen, M. B. The fitness landscapes of *cis*-acting binding sites in different promoter and environmental contexts. *PLoS Genet.* **6,** e1001042 (2010).
9. Sharon, E. *et al.* Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* **30,** 521–530 (2012).
10. Gerland, U. & Hwa, T. On the selection and evolution of regulatory DNA motifs. *J. Mol. Evol.* **55,** 386–400 (2002).
11. Berg, J., Willmann, S. & Lässig, M. Adaptive evolution of transcription factor binding sites. *BMC Evol. Biol.* **4,** 42 (2004).
12. Maerkl, S. J. & Quake, S. R. A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315,** 233–237 (2007).

13. Mustonen, V., Kinney, J., Callan, C. G. & Lässig, M. Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. *Proc. Natl Acad. Sci. USA* **105,** 12376–12381 (2008).
14. Haldane, A., Manhart, M. & Morozov, A. V. Biophysical fitness landscapes for transcription factor binding sites. *PLoS Comput. Biol.* **10,** e1003683 (2014).
15. Carlson, C. D. *et al.* Specificity landscapes of DNA binding molecules elucidate biological function. *Proc. Natl Acad. Sci. USA* **107,** 4544–4549 (2010).
16. Weghorn, D. & Lässig, M. Fitness landscape for nucleosome positioning. *Proc. Natl Acad. Sci. USA* **110,** 10988–10993 (2013).
17. Buenrostro, J. D. *et al.* Quantitative analysis of RNA–protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat. Biotechnol.* **32,** 562–568 (2014).
18. Newburger, D. E. & Bulyk, M. L. UniPROBE: an online database of protein binding microarray data on protein–DNA interactions. *Nucleic Acids Res.* **37,** D77–D82 (2009).
19. Weirauch, M. T. *et al.* Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158,** 1431–1443 (2014).
20. Badis, G. *et al.* Diversity and complexity in DNA recognition by transcription factors. *Science* **324,** 1720–1723 (2009).
21. Berger, M. F. *et al.* Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat. Biotechnol.* **24,** 1429–1435 (2006).
22. Payne, J. L. & Wagner, A. The robustness and evolvability of transcription factor binding sites. *Science* **343,** 875–877 (2014).
23. Zhu, C. *et al.* High-resolution DNA-binding specificity analysis of yeast transcription factors. *Genome Res.* **19,** 556–566 (2009).
24. Nakagawa, S., Gisselbrecht, S. S., Rogers, J. M., Hartl, D. L. & Bulyk, M. L. DNA-binding specificity changes in the evolution of forkhead transcription factors. *Proc. Natl Acad. Sci. USA* **110,** 12349–12354 (2013).
25. Maynard Smith, J. Natural selection and the concept of a protein space. *Nature* **225,** 563–564 (1970).
26. Lehner, B. Molecular mechanisms of epistasis within and between genes. *Trends Genet.* **27,** 323–331 (2011).
27. Poelwijk, F. J., Tănase-Nicola, S., Kiviet, D. J. & Tans, S. J. Reciprocal sign epistasis is a necessary condition for multi-peaked fitness landscapes. *J. Theor. Biol.* **272,** 141–144 (2011).
28. Jolma, A. *et al.* DNA-binding specificities of human transcription factors. *Cell* **152,** 327–339 (2013).
29. Weinreich, D. M., Delaney, N. F., Depristo, M. A. & Hartl, D. L. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312,** 111–114 (2006).
30. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515,** 355–364 (2014).
31. Stergachis, A. B. *et al.* Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* **515,** 365–370 (2014).
32. Hesselberth, J. R. *et al.* Global mapping of protein–DNA interactions *in vivo* by digital genomic footprinting. *Nat. Methods* **6,** 283–289 (2009).
33. Lynch, M. & Hagner, K. Evolutionary meandering of intermolecular interactions along the drift barrier. *Proc. Natl Acad. Sci. USA* **112,** E30–E38 (2015).
34. MacArthur, S. & Brookfield, J. F. Y. Expected rates and modes of evolution of enhancer sequences. *Mol. Biol. Evol.* **21,** 1064–1073 (2004).
35. Bergström, A. *et al.* A high-definition view of functional genetic variation from natural yeast genomes. *Mol. Biol. Evol.* **31,** 872–888 (2014).
36. MacIsaac, K. D. *et al.* An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* **7,** 113 (2006).
37. Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. & Carroll, S. B. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433,** 481–487 (2005).
38. Rister, J. *et al.* Single-base pair differences in a shared motif determine differential *Rhodopsin* expression. *Science* **350,** 1258–1261 (2015).
39. Siggers, T. & Gordân, R. Protein–DNA binding: complexities and multi-protein codes. *Nucleic Acids Res.* **42,** 2099–2111 (2014).
40. Li, X. Y. *et al.* Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6,** 0365–0388 (2008).
41. Fisher, W. W. *et al.* DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in *Drosophila*. *Proc. Natl Acad. Sci. USA* **109,** 21330–21335 (2012).
42. Mustonen, V. & Lässig, M. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* **25,** 111–119 (2009).
43. Arbiza, L. *et al.* Genome-wide inference of natural selection on human transcription factor binding sites. *Nat. Genet.* **45,** 723–729 (2013).
44. Mustonen, V. & Lässig, M. Evolutionary population genetics of promoters: predicting binding sites and functional phylogenies. *Proc. Natl Acad. Sci. USA* **102,** 15936–15941 (2005).
45. Swanson, C. I., Schwimmer, D. B. & Barolo, S. Rapid evolutionary rewiring of a structurally constrained eye enhancer. *Curr. Biol.* **21,** 1186–1196 (2011).
46. Grönlund, A., Lötstedt, P. & Elf, J. Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nat. Commun.* **4,** 1864 (2013).
47. Ramos, A. I. & Barolo, S. Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution. *Phil. Trans. R. Soc. B.* **368,** 20130018 (2013).
48. Crocker, J. *et al.* Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* **160,** 191–203 (2015).
49. Berger, M. F. & Bulyk, M. L. Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat. Protoc.* **4,** 393–411 (2009).
50. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27,** 1017–1018 (2011).
51. van Helden, J., André, B. & Collado-Vides, J. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *J. Mol. Biol.* **281,** 827–842 (1998).
52. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515 (2010).
53. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).
54. Dawid, A., Kiviet, D. J., Kogenaru, M., de Vos, M. & Tans, S. J. Multiple peaks and reciprocal sign epistasis in an empirically determined genotype-phenotype landscape. *Chaos* **20,** 26105 (2010).
55. Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M. & Tans, S. J. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445,** 383–386 (2007).
56. Franke, J., Klözer, A., de Visser, J. A. G. M. & Krug, J. Evolutionary accessibility of mutational pathways. *PLoS Comput. Biol.* **7,** e1002134 (2011).
57. Parker, D. S., White, M. A., Ramos, A. I., Cohen, B. A. & Barolo, S. The *cis*-regulatory logic of Hedgehog gradient responses: key roles for gli binding affinity, competition, and cooperativity. *Sci. Signal.* **4,** ra38 (2011).
58. Zhao, Y. & Stormo, G. D. Quantitative analysis demonstrates most transcription factors require only simple models of specificity. *Nat. Biotechnol.* **29,** 480–483 (2011).
59. Morris, Q., Bulyk, M. L. & Hughes, T. R. Jury remains out on simple models of transcription factor specificity. *Nat. Biotechnol.* **29,** 483–484 (2011).
60. Weinreich, D. M., Watson, R. A. & Chao, L. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59,** 1165–1174 (2005).

Acknowledgements

J.A.-R. and J.L.P. acknowledge support through the Forschungskredit program of the University of Zurich (grant numbers FK-14-076 and K-74301-04-01). J.L.P. acknowledges additional support through the Ambizione program of the Swiss National Science Foundation. A.W. acknowledges support through the Swiss National Science Foundation (grant 31003A_146137) and the University Priority Research Program in Evolutionary Biology at the University of Zurich. We thank S. Bratulic, F. Khalid, A. Moya, Y. Schaerli and M. Toll-Riera for discussions and helpful comments on this manuscript.

Author contributions

J.A.-R., J.L.P. and A.W. designed the research. J.A.-R. and J.L.P. performed the research. J.A.-R., J.L.P. and A.W. analysed the data and wrote the paper.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.W.

How to cite this article: Aguilar-Rodríguez, J., Payne, J. L. & Wagner, A. A thousand empirical adaptive landscapes and their navigability. *Nat. Ecol. Evol.* **1,** 0045 (2017).

Competing interests

The authors declare no competing financial interests.