

University of Massachusetts Medical School

eScholarship@UMMS

University of Massachusetts Medical School Faculty Publications

2019-11-05

The adaptive potential of the M-domain of yeast Hsp90 [preprint]

Pamela A. Cote-Hammarlof

University of Massachusetts Medical School

Et al.

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/faculty_pubs



Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Biochemistry, Biophysics, and Structural Biology Commons](#), and the [Ecology and Evolutionary Biology Commons](#)

Repository Citation

Cote-Hammarlof PA, Fragata I, Flynn JM, Zeldovich KB, Bank C, Bolon DN. (2019). The adaptive potential of the M-domain of yeast Hsp90 [preprint]. University of Massachusetts Medical School Faculty Publications. <https://doi.org/10.1101/832022>. Retrieved from https://escholarship.umassmed.edu/faculty_pubs/1648

Creative Commons License



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 4.0 License](#). This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in University of Massachusetts Medical School Faculty Publications by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

The adaptive potential of the M-domain of yeast Hsp90

Pamela A. Cote-Hammarlof^{1*}, Inês Fragata^{2*§}, Julia Flynn¹, Konstantin B. Zeldovich¹,
Claudia Bank^{2#}, and Daniel N.A. Bolon^{1#}

*Co-first authors

#Co-corresponding authors

¹ University of Massachusetts Medical School, Worcester, USA

² Instituto Gulbenkian de Ciência, Oeiras, Portugal

§ Current address: Centre for Ecology and Evolution, Lisbon, Portugal

Corresponding authors:

Daniel N.A. Bolon, Dan.Bolon@umassmed.edu

Claudia Bank, evoldynamics@gmail.com

Abstract

Comparing the distribution of fitness effects (DFE) of new mutations across different environments quantifies the potential for adaptation in a given environment and its cost in other environments. So far, results regarding the cost of adaptation across environments have been mixed, and there were no sufficiently large data sets to map its variation along the genome. Here, we study the DFEs of ≈ 2500 amino-acid changing mutations obtained from deep mutational scanning of the 118 amino-acid-long middle domain of the heat-shock protein Hsp90 in five environments and at two expression levels. This region is known to be important for client binding, stabilization of the Hsp90 dimer, stabilization of the N-M and M-C interdomains and regulation of ATPase-chaperone activity. Despite the diverse and stressful environments, we find that fitness correlates well across environments, with the exception of one environment, diamide. Consistent with these results, we find very little cost of adaptation; on average only one in seven beneficial mutations is deleterious in another environment. We identify a hotspot of beneficial mutations in a region of the protein that is located within an allosteric center. The identified protein regions that are enriched in beneficial, deleterious, and costly mutations coincide with residues that are involved in the stabilization of Hsp90 interdomains and stabilization of client binding interfaces or residues that are involved in ATPase chaperone activity of Hsp90. Thus, our study yields information regarding the role and adaptive potential of a protein sequence that complements and extends known structural information.

Introduction

The distribution of fitness effects (DFE) describes the proportions of new mutations that are beneficial, deleterious or neutral (Eyre-Walker & Keightley 2007, Loewe 2010, Bataillon & Bailey 2014). This provides a snapshot of the robustness of the genome to changes in the

DNA and on the expected amount of genetic diversity within populations. Moreover, the beneficial part of the DFE informs on the adaptive potential of populations when introduced in a new environment (Sniegowski & Gerrish 2010, Bataillon & Bailey 2014). While maintaining the same shape of the DFE, mutations in one environment can be deleterious in another (Bataillon et al 2011). So far, it has been difficult to address the prevalence of such costs of adaptation (also termed antagonistic pleiotropy), because it was difficult to measure the fitness of the same mutations across various environments (but see, e.g., Hietpas et al 2013, Sane et al 2018).

Around a decade ago, inference of the complete DFE of different regions in several model species started to become feasible, using a combination of directed site mutagenesis and deep sequencing (Fowler et al. 2010, Hietpas et al 2012, Boucher et al. 2014). These studies usually reported a bimodal DFE with two peaks that represent neutral (or wild-type like) mutations and another that represents strongly deleterious mutations [Acevedo et al 2014, Bank et al 2014, Boucher et al 2014]. With improvements of the experimental and sequencing technologies, studies of hundreds of thousands of mutant effects have become feasible (Melamed et al 2013, Boucher et al 2014, Doud et al 2016, Sarkisyan et al 2016,). However, few studies have quantified the impact of environmental challenges on the shape of these DFEs (but see Hietpas et al 2013). That is partly because inferring the adaptive potential across environments and the cost of beneficial mutations in alternative environments requires a high experimental resolution of selection coefficient estimates.

Experimentally quantifying the DFE across environments is important because natural environments can change due to the alterations of abiotic factors including, temperature, salinity, and pH, which can have diverse evolutionary consequences that are directly related to the shape of the DFE (Dhar et al. 2011; Arribas et al. 2014; Mumby and van Woesik 2014; Brennan et al. 2017). Because environments are in general likely to change faster than genetic adaptation can proceed, the long-term fitness of individuals depends on their ability to deal with a wide range of environmental stresses.

Environmental challenges manifest at many biological levels, including the protein level. For example, increased temperature can cause protein unfolding and aggregation, which can disrupt function, ultimately affecting organismal fitness and survival (Richter et al. 2010). Chaperones, such as the heat shock protein Hsp90, help the cellular machinery survive stress conditions (Chen et al. 2018). They are an important line of defense for new and recurring environmental challenges and can also function to increase standing genetic variation that can facilitate adaptation to new stress conditions (Barrett and Schluter 2008; Jarosz and Lindquist 2010; Fitzgerald and Rosenberg 2019). Therefore, it is important to understand how chaperones themselves evolve, and how the selection pressure on the chaperone itself changes upon exposure to different environmental challenges.

Hsp90 is a chaperone that plays an essential role in protecting cells from environmental stress and is required at elevated levels for yeast growth at high temperature (Borkovich et al. 1989). Under optimal conditions Hsp90 can suppress the phenotypic consequences of mutations in signal transduction pathways by correctly folding mutant clients (Sangster et al. 2004; Jarosz and Lindquist 2010; Zabinsky et al. 2019). However, in stressful conditions there is an increase in the concentration of stress-induced unfolded and damaged proteins, which compete for Hsp90 chaperone function (Sangster et al. 2004; Taipale et al. 2010). Thus, mutations in protein clients and in Hsp90 itself may have phenotypic and ultimately fitness consequences.

Recent studies using systematic mutagenesis have begun to address how mutations to a highly conserved client binding site of Hsp90 can impact evolutionary adaptation in yeast (Hietpas et al. 2013). Multiple mutations in a nine amino acid client-binding site of yeast Hsp90 provided a growth advantage under elevated salinity conditions (Hietpas et al. 2013). Another recent study found that at low Hsp90 expression levels, changes in environment greatly changed the shape of the DFE, and that some environments showed a higher prevalence of both beneficial and deleterious mutations (Flynn et al 2019). However, another previous study proposed that at low expression the fitness effects of mutations, especially deleterious ones, should be larger (Jiang et al 2013). Thus, it was not possible to separate how much of the observed effect in this new study were due to expression level and which was due to environmental changes.

Here we examine a 119 amino acid region (encompassing positions 291-409, Fig 1) of the yeast middle domain Hsp90, known to bind to many client proteins (Nathan and Lindquist 1995; Nathan et al. 1997; Meyer et al. 2003; Hawle et al. 2006; Hagn et al. 2011). Several studies have demonstrated that the middle domain of Hsp90 plays a prominent role in client binding, and suggested that mutations in this region may impact the relative affinity or priority of different clients and physiological pathways with the potential to provide an adaptive benefit (Hagn et al. 2011, Zhang et al. 2005, Sato et al. 2000, Karagoz et al. 2014, Verba et al. 2016; Czemerer et al. 2017).

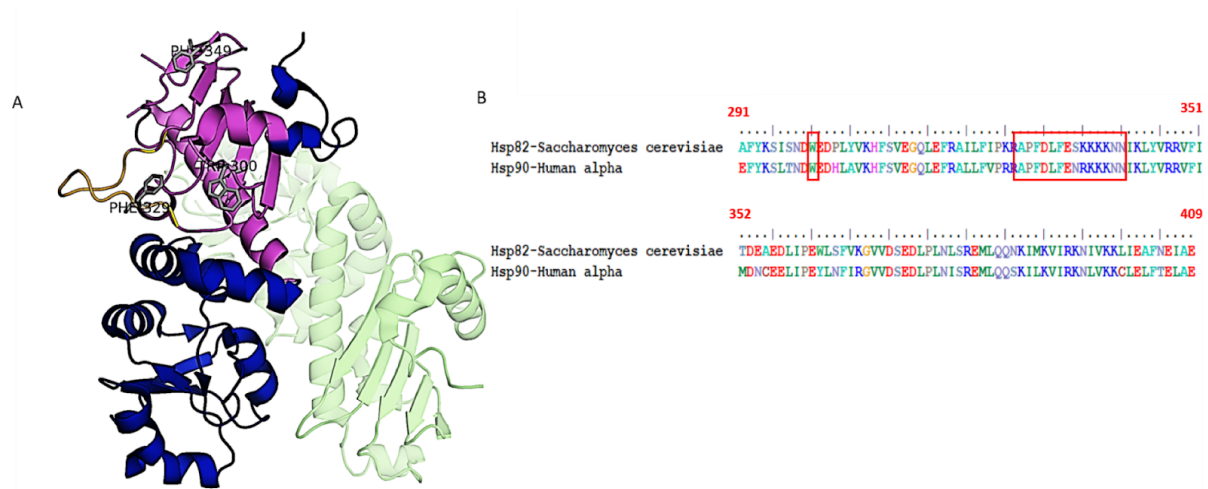


Figure 1: Middle domain of Hsp90

(A) In this image of the structure of Hsp90, the middle domain (amino acids 291-409) is highlighted in purple. A solvent exposed amphipathic loop implicated in client binding (amino acids 327-341) is highlighted in yellow and residues implicated in client binding (amino acids W300, F329, F349) are shown in grey. This representation is a modification of a monomer based on a crystal structure of the M domain of WT yeast Hsp90 (PDB 1HK7). (B) Amino acid sequence alignment of the M-domain region of study in yeast Hsp82 and human Hsp90. Boxed in red are the amino acid residues that form the solvent exposed amphipathic loop implicated in client binding (amino acids 300, 327-341).

We here used the EMPIRIC approach of deep mutational scanning (Hietpas et al 2012) to estimate the selection coefficient of all amino-acid changing mutations in the M-domain of Hsp90 across five environments. Whereas this study was performed at normal expression levels, we included low-expression data from Flynn et al. 2019 in diamide, salt and standard conditions in our analysis to test the impact of protein expression on the fitness effect of mutations. Using the inferred DFEs, we identified regions of the M-domain that stand out with respect to their potential for adaptation upon environmental change. To this end we: a) quantified the impact of environmental changes on the overall shape of the DFE, b) identified regions that show the largest proportions of beneficial or deleterious mutations, c) quantified hotspots of costs of adaptation, and d) compared the identified regions with known client binding or other structurally important sites to connect the phenotypic and fitness effects of mutations.

Results and Discussion

DMS across environments and expression levels reveals distributions of fitness effects.

To investigate the adaptive potential of the M domain region of yeast Hsp90 (amino acids 291-409) we used systematic site-directed mutagenesis in this region. This resulted in ≈ 2500 amino-acid changing mutations, whose selection coefficients were estimated from 2-3 replicates of bulk competitions that were performed for five environments. We focused our analysis on the standard lab conditions and four environmental stresses that affect growth rate in yeast (Supplementary figure 9) and the adaptive gene expression response (Gasch et al. 2000) , two that cause osmotic stress (0.5M salinity or 0.6M sorbitol) and two that cause oxidative stress (0.6 mM H₂O₂ or 0.85mM diamide). These conditions were chosen because of Hsp90's role in yeast response to these environmental stressors. In yeast, responses to environmental stress requires changes in signaling pathways that dramatically alters gene expression programs (de Nadal et al. 2011). For example, during osmotic stress, Hog1, the yeast homolog of the mammalian p38 mitogen activated protein kinase becomes activated (Hawle et al., 2007) and results in the induction of ~ 600 genes required for yeast response to osmotic stress (Westfall et al. 2004). Hsp90 function is required for adaptation to osmotic stress (Yang et al. 2006) and Hsp90 is necessary for the activation of the HOG pathway (Hawle et al. 2007; Yang et al. 2007). Although Hsp90 function is required for the osmotic stress response, increased expression of Hsp90 does not occur in this condition (Gasch et al. 2000). In contrast, the oxidative stresses of diamide and H₂O₂ cause an increase in the expression of chaperones, including Hsp90 (Gasch et al. 2000). Moreover, while each of these conditions causes changes in the expression of ~ 900 genes, diamide and H₂O₂ stress promote similar gene expression responses to those observed during thermal stress in yeast (Gasch et al. 2000). To allow for comparisons of mutational fitness across expression levels (Dekel and Alon 2005; Manceau et al. 2011), we included three data sets that were obtained in previous work (Flynn et al, Biorxiv), which evaluated the same mutations at constitutively low expression levels in the standard lab environment, the high-salinity environment, and the diamide environment.

DFEs across all environments show many wild-type like mutations. The shape of the DFE indicates the relative importance of purifying or directional selection as compared with neutral evolution. In apparent contrast with the strong conservation of the Hsp90 M-domain in natural populations, we observed DFEs with mostly wild-type like mutations across all environments (Supplementary Figure 1). Here, we categorized mutants as wild-type like if the 95% confidence interval of the estimated selection coefficient overlapped with 0 (see also Supplementary Table 1, Materials & Methods). According to this criterion, at high expression levels, between 50 % (in H₂O₂ environment) and 65% (in standard environment) of mutations showed a fitness effect that is indistinguishable from the reference type. At low expression levels, where the experimental resolution was lower, between 63% and 73% of mutations (in standard and diamide environments, respectively) showed a wild-type like effect. Large numbers of wild-type like mutations have been observed previously in DMS studies of DFEs (Soskine and Tawfik 2010, Hietpas et al 2013, Melamed et al 2013, Bank et al 2014, Hom et al 2019). Both biological and technical factors can be invoked as an explanation for the large number of wild-type like mutations. Firstly, the resolution of the experiment is likely much lower than the resolution at which natural selection may occur in large yeast populations (Ohta 1992). Secondly, selection pressures in the laboratory might differ greatly from those in nature (Reznick & Ghalambor 2005, Kvitek and Sherlock 2013). Finally and relatedly, selection pressures in nature might be fundamentally different and rapidly fluctuating (Mustonen and Lässig 2009). For example, a recent study of the DFE of the full Hsp90 sequence found that those mutations which were most robust across a set of diverse environments were those most likely observed in natural sequences (Flynn et al., 2019).

Correlations across conditions reveal that diamide stands out with respect to mutational effects. We next quantified the correlations of fitness effects across replicates, environments and expression levels (Figure 2, Supplementary Table 2). Consistent with the high accuracy of the experiment, we observed strong correlations between replicates (mean Pearson correlation $r=0.91$). Across environments and expression levels, we observed a large variation of correlations ranging from $r=0.36$ between diamide and low-expression salt to $r=0.98$ between H₂O₂ and the standard environment. In this analysis, diamide clearly stood out by showing consistently lower correlations with other environments and expression levels than all others (mean correlation including diamide $r=0.64$, mean correlation of all others $r=0.83$). This could be do to the fact that diamide, unlike any of the other conditions examined, has multiple negative effects on the cell, including oxidative stress on proteins, protein unfolding, and cell wall damage (Gasch et al. 2000). These detrimental effects place extra pressure on the chaperone system and thus require the increased expression of chaperone proteins including Hsp90 (Gasch et al. 2000). The observed differences between the two oxidative stress environments diamide and H₂O₂ can be explained by their different effects on the nuclear localization of Yap1 transcription factor (Gulshan et al. 2011). Yap1 is a transcription factor that is required to mount an adaptive response to diamide but is crucial for an adaptive response to H₂O₂ (Gulshan et al. 2011). Thus, the observed differences between fitness effects observed under diamide conditions and other conditions and expression levels are likely to arise from a combination of the global stress diamide exerts on the organism, and by specific differences between diamide and other oxidative stress conditions.

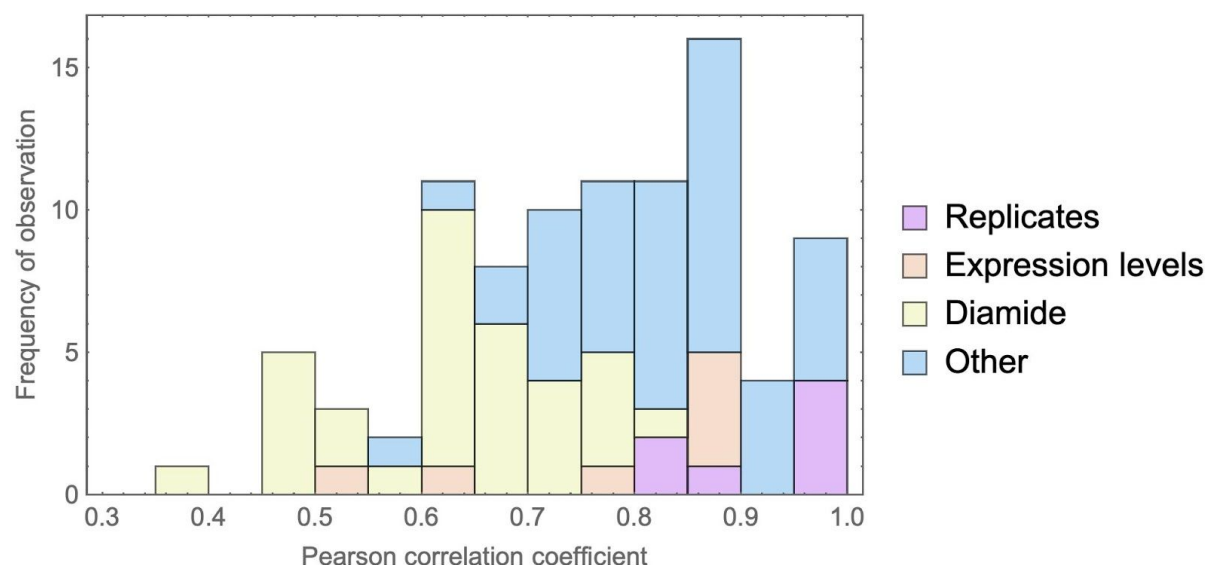


Figure 2: Histogram of correlations of fitness effects among replicates, expression levels, and environments reveals that diamide stands out as different. The stacked histogram is split into comparisons of replicates, expression levels, comparisons that include diamide, and all other comparisons.

We also observed a tendency to larger correlations of selection coefficients from the same batch of experiments (Supplementary Table 2). We tried to mitigate this effect by performing some of the analyses with the average of the estimated fitness effects between replicates and re-categorizing mutations based on the mutant category of both replicates, as indicated.

Finally, we compared the correlations of fitness effects across environments with those of the same environment, but obtained at a different (constitutive) expression level. A general expectation is that the Pearson correlation of DFEs between different expression levels should on average be larger than between environmental conditions. This would be consistent with the results and the verbal model of Jiang et al. 2013, who showed that the variance of the DFE increases at lower expression level. That is expected because at lower expression, small effects should become more important for protein production and thus become visible. However, we do not see any difference in the mean correlations within environments at the same expression level compared with correlations between expression levels (mean Pearson correlation across expression levels: $r=0.76$, vs. $r=78$ for correlations of environments with the same expression level). Again, we find that diamide stands out with a particularly low correlation between low and high expression levels. This makes sense with respect to the above-mentioned arguments, because an increase in Hsp90 expression levels is required to maintain protein and cellular homeostasis under diamide conditions (Gasch et al. 2000), which is compromised when Hsp90 is expressed at lower-than-normal expression levels.

The large correlations observed between most environments suggest that potential costs of adaptation (discussed in detail further below) should be the exception rather than the rule. This is at odds with previous results from a smaller region of Hsp90 (Hietpas et al 2013), where most beneficial mutations detected in high salinity were deleterious in other environments. A potential explanation for this discrepancy lies in the choice of the region. Whereas the previous study focused on only nine amino acids and thus chose a set of positions that are likely to show functionally interesting patterns, we here had the experimental power to scan a larger region of the protein. Along a larger, less specifically chosen, stretch of the protein, it is likely that many positions and mutations play the same role across most environments. Consequently, only a small number of mutations in this greater set may have a functionally different role across environments, leading to costs of adaptation. In such case, costs of adaptation of beneficial mutations could still be large, but due to their low proportion they barely affect the correlation of selection coefficients across environments.

Beneficial mutations appear across all environments, and are enriched in a region of Hsp90 that is implicated in stabilization of interdomains and client binding interfaces.

Next, we identified the proportion and identity of putatively beneficial mutations across all treatments (Supplementary Table 1, Supplementary Figure 2, 3). We categorized mutants as putatively beneficial if the lower limit of the 95% confidence interval of the selection coefficient was larger than 0, and we considered as strongest candidates those that overlapped between replicates from the same environment. As reported above, wild-type like mutations are the most abundant category in all environments, with a low proportion but considerable number of beneficial mutations in almost all environments. We find the lowest number of beneficial mutations in H₂O₂ and salt environments ($n_{\text{H}_2\text{O}_2}$ =65 and n_{Salt} =64, around 2.5% of the mutations) and the highest in diamide (n_{Diamide} =307, 12.3%). Interestingly, in diamide beneficial mutations were dispersed across the whole M-domain, whereas most of the other environments showed a clear enrichment of beneficial mutations in the region of positions 356-366 (Figure 3). This hotspot of beneficial mutations was particularly strong in sorbitol (Supplementary Figure 2). Out of the 31 beneficial mutations that show a beneficial effect in at least two environments, 18 are located in this hotspot.

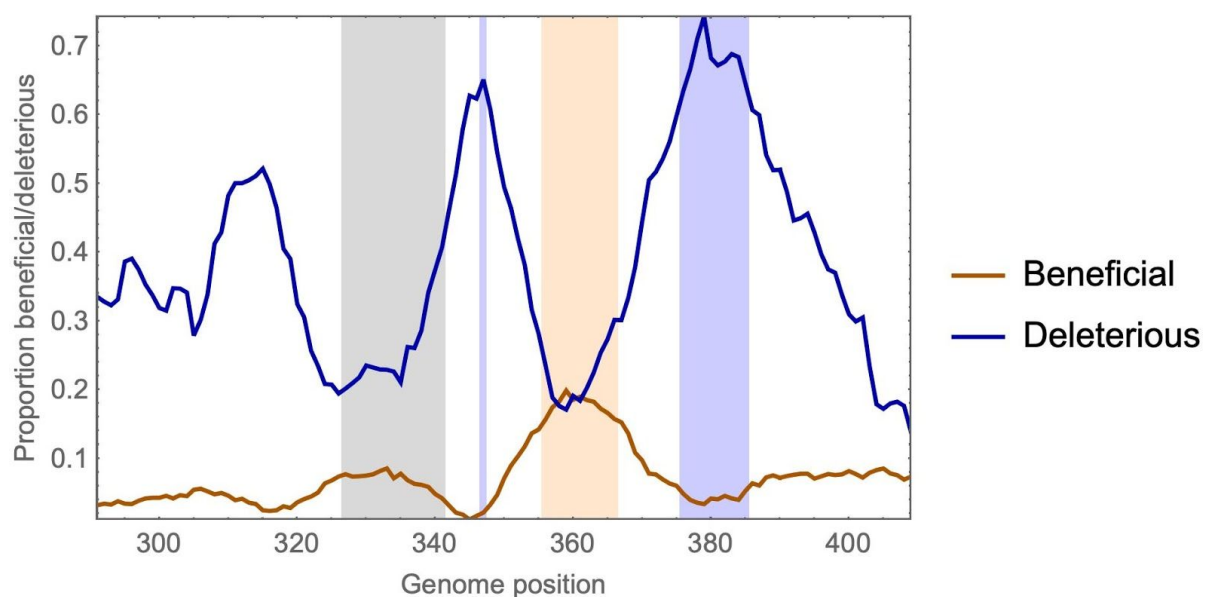


Figure 3: Hotspots of beneficial and deleterious mutations along the protein region do not coincide with known client binding loop. Average proportion of beneficial and deleterious mutations along the genome across all environments, illustrated using a 10-amino-acid sliding window. The region with the 10% largest proportions of beneficials is highlighted in light orange (position 356-366), the region with the 10% largest proportions of deleterious mutations is highlighted in light gray (positions 347 and 376-385). A known client binding loop (positions 327-341) is highlighted in light gray.

Interestingly the beneficial hotspot contains residues that are part of an allosteric center that mediates distinct conformations in structures with and without client (Czemerer et al. 2017). Despite its different structural arrangements, the hydrophobic amino acids in this region are mostly buried from solvent, indicating that the primary role of this region is mediating the stability and rearrangement of different conformations. Of note, the position with the greatest number of beneficial mutations is F364. The burial of the large hydrophobic side chain of F364 should provide local stability. The beneficial mutations at position 364 and in this region in general suggest that disruption of local stability and conformational dynamics may alter Hsp90 function. Based on the M-domains prominent role in binding to clients, we hypothesize that Hsp90 function may be altered in a manner that alters the relative affinity and thereby priority of different clients - a property that could provide benefits in specific conditions. The premise being that natural selection will balance client priorities integrated over multiple conditions, providing opportunities for Hsp90 mutations to improve priorities for individual conditions.

The proportion of deleterious, wild-type like, and beneficial mutations varies greatly along the protein sequence. We showed above that the correlations of fitness effects across environments are generally large, which indicates similar effects of the same mutations across environments. In contrast, the proportions of beneficial, wild-type like, and deleterious mutations vary greatly along the protein sequence, and whereas the overall pattern is similar between environments, the relative proportions differ between environments (Supplementary Figure 2, 3). For example, at low expression there is a decrease in the number of beneficial mutations, but the regions that show deleterious mutations are similar between expression levels. The finding that the same positions are enriched for deleterious and beneficial mutations across environments suggests that the structural properties of the M domain in those regions rather than its binding partners might be the most important factor in predicting the fitness effects of mutations.

One position that particularly stands out with respect to the proportions of beneficial, wild-type like and deleterious mutations is 364. Position 364 lies at the center of the cluster of the 10% largest average proportions of beneficial (Figure 3), and thus shows a clear deviation from the usually observed DFE with an enrichment of beneficials (Supplementary Figure 4). Specifically, we identified 6 mutations at this position that are beneficial in at least two environments. As mentioned above this may be due to a disruption on the stability of the protein that may alter the relative affinity of Hsp90 to its clients, which may be beneficial in specific environments.

Interestingly, the regions with the largest proportions of deleterious mutations are located near the ends of the beneficial hotspot region (position 347 and 376-385, respectively). At the

latter deleterious hotspot, around 70% of all mutations are deleterious in a sliding window of 10 amino-acid positions. Such a shift of the DFE towards deleterious mutations suggests that this protein region is under strong purifying selection. Structurally, the region 376-385 is part of a catalytic loop required for ATP hydrolysis which is required for the activation of all known clients (Wolmarans et al. 2016). R380 in the catalytic loop binds to and stabilizes the leaving phosphate of ATP and mutations at this position are lethal. The efficiency of catalysts depend strongly on geometry, and the precise location of R380 relative to ATP is likely tightly linked to Hsp90 function. The regions adjacent to R380 appear to be important for positioning the catalytic arginine, providing a rationale for the strong purifying selection that we observe.

Few mutations show costs of adaptation. Among the identified beneficial mutations, we were specifically interested in those that are deleterious in other environments, which results in a so called "cost of adaptation". This same phenomenon has also been termed antagonistic pleiotropy. In previous work we reported a large prevalence of costs of adaptation in a 9-amino-acid region of Hsp90 (Hietpas et al 2013). In this study, we found that costs of adaptation are not pervasive (Figure 4, Supplementary Table 3). Consistent with the comparatively low correlations of fitness effects and the special role of the diamide stress discussed above, we see the largest proportion of mutations that show costs of adaptation between diamide and other environments. Whereas the mean proportion of mutations displaying costs of adaptation across all comparisons of environments was 14.3%, comparisons that included diamide showed on average 22.7% mutations with costs of adaptation. The largest proportions of mutations with costs of adaptation (>40%) were seen in comparisons of low-expression diamide and high-expression H2O2.

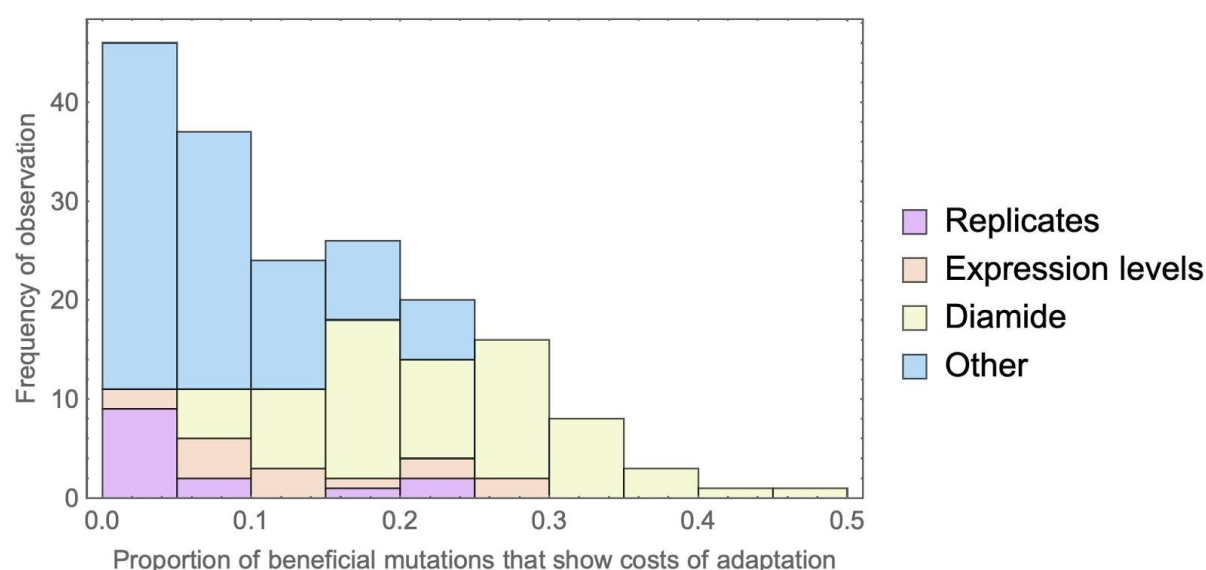


Figure 4: Histogram of proportions of mutations that display costs of adaptation illustrates that costs of adaptation are generally rare, and most likely if comparisons involve diamide environments. The stacked histogram is split into comparisons of

replicates, expression levels, the subset of other comparisons that include diamide, and all other comparisons.

Mapping the proportion of beneficials that are deleterious in other environments along the protein sequence, we observed that there is a hotspot for costs of adaptation at amino-acid positions 381-391 (Figure 5). Interestingly, this region, which partly overlaps with the hotspot of deleterious mutations (positions 376-385) showed costs of adaptation across environments, which indicates that each environment has specific beneficial mutations, which are deleterious in other environments. Structurally, this region also belongs to the catalytic loop involved in ATP hydrolysis discussed above.

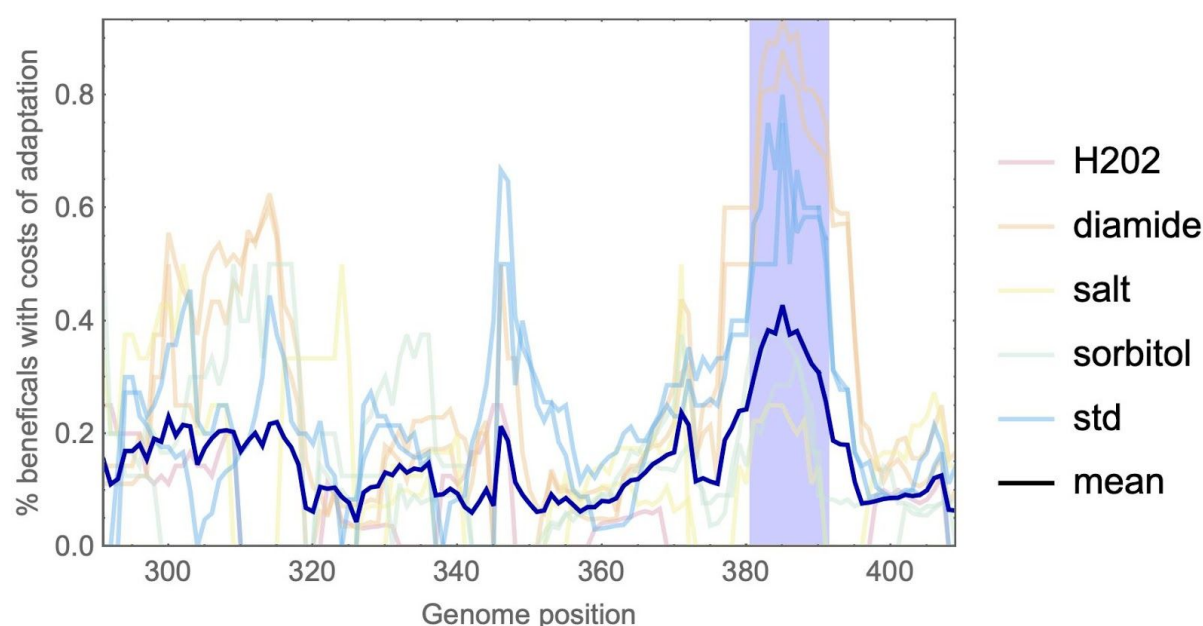


Figure 5: A hotspot of costs of adaptation is located at positions 381-391. The 10% positions with the largest mean proportion of mutations that display costs of adaptation (dark blue line) are highlighted in gray. The lighter curves indicate the mean proportion of beneficial mutations in a focal environment that is deleterious in another environments. A 10-amino-acid sliding window was used to locate region-specific effects. To avoid biases, the mean was computed across the subset of two replicates of all normal-expression environments.

We next computed whether there was a correlation between the effect of beneficial mutations in one environment with their effect in the other environments, which would indicate that larger-effect beneficial mutations are more likely to have stronger deleterious effects in other environments. Indeed, the effects of beneficial mutation in diamide are in general negatively correlated with the same mutation's effect in all other environments (Figure 5, Supplementary Figure 5) except H2O2. Thus, the larger the effect of a beneficial mutation in diamide, the lower it tends to be in salt, standard and sorbitol. Again, this suggests that the beneficial

mutations we identified in diamide may be involved in the response to a very specific type of stress. Across all other environments, we do not see any evidence that stronger beneficial mutations tend to have a more deleterious effect in other environments.

When comparing low and normal expression levels for diamide and salt (Supplementary Table 3), we observed that the percentage of beneficials at normal levels of expression that were deleterious at low expression is small, but that there is a larger percentage of beneficials at low expression that are deleterious at high expression. This pattern is inverted for the standard environment, which is consistent with the results from Jiang et al 2013. The asymmetric cost between expression levels in salt and diamide environments suggests that mutations that are beneficial in low expression levels and deleterious in normal levels, may be involved in changes of protein synthesis/increase in protein production. Conversely, the low cost of beneficials in normal expression, may point to mutations that are involved in the general response to the diamide stress.

Conclusion

Recent advances in experimental and technological approaches have led to the feasibility of large-scale screens of the DFE of new mutations, which, from an evolutionary point of view, is a key entity to determine the potential for adaptation. Here, we took a new route by mapping the proportions of beneficial, wild-type like, and deleterious mutations along a 119-amino-acid region of the Hsp90 protein in yeast, a protein that is heavily involved in the response to environmental stressors. By comparing the DFE along the protein sequence and between environments, we identified hotspots of beneficial and deleterious mutations that are common across environments, and a region in which beneficial mutations in one environment tend to be deleterious in other environments. Interestingly, neither of these regions coincided with the best described client binding loop in the studied region, which we had *a priori* considered the most likely candidate to display patterns different from the rest of the region. Moreover, our analyses suggested that mutational effects change little across environments except in diamide, which stands out both with respect to the number and also the distribution of beneficial mutations along the studied protein region. Altogether, our study highlights how the study of the DFE across environments sheds light on the evolutionary role of specific protein regions from a new perspective.

Materials and Methods

Generating point mutations. To accurately measure the fitness effects of all possible point mutations in a large portion of the middle (client-binding) domain of yeast Hsp90, we used saturation mutagenesis at positions 291-409 (see Figure 6). We used a cassette ligation strategy to introduce mutations as previously described (Hietpas et al. 2012). This strategy reduces the likelihood of secondary mutations because it avoids the potential for errors during PCR steps. As a control for the mutational procedure, twelve positions were randomized in isolation and Sanger sequenced to assess the level of incorporation of all four nucleotides at

each position in the target codon. At all randomized nucleotide positions within these twelve samples, we observed a similar magnitude of signal for each of the four nucleotides. For larger scale production, libraries were generated at 10 positions at a time as previously described (Hietpas et al. 2012). As additional controls, we generated a sample containing individual stop codons as well as the parental wildtype Hsp90 sequence. All variants of Hsp90 were generated in a plasmid (pRS414GPD) previously shown to produce endogenous levels of Hsp90 protein in yeast (Chang and Lindquist 1994).

Constructing barcoded libraries. To improve the efficiency and accuracy of fitness estimates, we added barcodes to a non-functional region of the plasmid, ~200 NTs downstream from the 3' untranslated region of Hsp90. Barcodes were introduced using a cassette ligation strategy (Supplementary Figure 5). Plasmid libraries were treated with NotI and AscI to generate directional sticky ends. NotI and AscI recognize 8-base cut sites that are unlikely to cut any of the Hsp90 variants in the library. We designed and annealed barcode forward and barcode reverse oligos together such that the resulting duplex product included a central N18 region bracketed by constant regions that facilitate annealing and overhangs that direct directional ligation into NotI and AscI overhangs. One of the constant ends in the designed oligo cassette contains an annealing region for an Illumina sequencing primer. Barcoded libraries were transformed into *E. coli* and pooled into a bulk culture that contained about 10-fold more transformants than Hsp90 variants in the library. We purified barcoded plasmids from this bulk culture. This procedure resulted in approximately 10 barcodes for each Hsp90 codon variant in the library (Supplementary figure 6). The potential diversity in the N18 barcode that we used ($4^{18} \sim 10^{11}$) far exceeds the number of barcodes that we utilize ($\sim 64 \times 119 \times 10 \sim 10^5$), which makes it likely that each Hsp90 variant will have a barcode that differs from all other barcodes at multiple nucleotides. With this setup, errors in sequencing of barcodes can be detected and eliminated from further analysis, which reduces the impact of sequencing misreads on estimates of variant frequency and fitness. Additional controls consisting of individual stop codons and wildtype Hsp90 were barcoded separately. For these controls, we isolated barcoded plasmid DNA from individual bacterial colonies and determined the barcodes by Sanger sequencing (Supplementary figure 5).

Associating barcodes to mutants. To identify the barcodes associated with each Hsp90 variant in our libraries, we used a paired-end sequencing approach essentially as previously described (Hiatt et al. 2010). Using paired-end sequencing on an Illumina MiSeq Instrument barcodes were associated with variant genotypes via reading from the Hsp90 gene with a 250 base-pair read and the associated N18 barcode with a 50 base-pair read. To reduce the size of the DNA fragments for efficient Illumina sequencing, we removed a portion of the Hsp90 gene such that the randomized regions were closer to the N18 barcode. This was done to increase the density of DNA on the sequencer by reducing the radius of clonal clusters during sequencing. Plasmid DNA was linearized with StuI and NotI endonucleases that removed bases of DNA. The linearized products were circularized by blunt ending followed by

ligation at DNA concentrations that predominantly lead to unimolecular ligations (REF). The resulting DNA products were amplified using a PE2_F primer and the standard Illumina PE1 primer that anneals next to the N18 barcode. Two PE2_F primers were designed in order to read across the region of Hsp90 that we randomized. PCR products were gel purified and submitted for paired end sequencing. We obtained sufficient paired-end reads such that the average barcode was read more than 10 times. The paired-end sequencing data was subjected to a custom data analysis pipeline to associate Hsp90 variants with barcodes. First, very low-quality reads with any Phred score less than 10 in reads 1 or 2 were discarded. Next, the data were organized by the barcode sequence. For barcodes with three or more reads, we constructed a consensus of the Hsp90 sequence read. We compared the consensus sequence to the parental Hsp90 sequence in order to determine mutations. Consensus sequences containing more than one protein mutation were discarded. The pipeline output generated is a file organized by point mutation that lists all barcodes associated with that mutation (Supplementary excel file-Hsp90_291_409_BarcodeAssociationTable). This output is what is used for calling a variant.

Yeast competitions. As in previous work (Hietpas et al. 2012; Hietpas et al. 2013), we performed Hsp90 competitions using a shutoff strain of *S. cerevisiae* (DBY288) (Figure 6). The sole copy of Hsp90 in DBY288 is driven by a galactose-inducible promoter, such that the strain requires galactose for viability and cannot grow on glucose. The introduction of a functional Hsp90 variant driven by a constitutive promoter rescues the growth of DBY288 in glucose media. We introduced the library of middle domain variants of Hsp90 driven by a constitutive promoter into DBY288 cells using the lithium acetate method (Gietz and Schiestl 2007). The efficiency of transformation was more than 10-fold higher than the number of barcodes, such that the average barcode was transformed into more than 10 individual cells. To enable the analyses of full biological replicates, transformations were performed in replicate such that a separate population of yeast were transformed for each biological replicate. Transformed yeast were initially selected in RGal-W (synthetic media that lacked tryptophan and contained 1% raffinose and 1% galactose supplemented with 100 µg/mL ampicillin to hinder bacterial contamination) to select for the plasmid, but not function of the plasmid encoded Hsp90 variant. This enabled us to generate a yeast library containing Hsp90 variants that could support a full range of fitness from null to adaptive. Cells were grown for 48 hours at 30 °C in liquid SRGal-W until the culture was visibly opaque compared to a control transformation sample that lacked plasmid DNA but was otherwise identical to the library sample.

To initiate the competition, cells were transferred to shut-off conditions and different environmental stresses. In order to deplete the pool of wild type Hsp90 protein, the library sample was diluted into SD-W (synthetic media lacking tryptophan with 2% glucose and 100 µg/mL of ampicillin) and grown at 30 °C for 10 hours. After depletion of wild type Hsp90 protein, the cells were split into different stress conditions in SD-W media at 30 °C: salt stress (0.5M NaCl), osmotic stress (0.6M sorbitol), oxidative stress (0.6 mM hydrogen peroxide or 0.85 mM diamide), as well as a control non-stress condition. We used a custom

built turbidostat (Figure 6) to provide constant growth conditions during this competition phase of the experiment. Cells were grown under consistent density and population size. 10^9 cells were maintained in a 50 mL volume over the course of 40 hours for each condition. Rapid magnetic stirring was used to provide aeration to the media and cells. Samples of 10^8 cells were collected after 0, 4, 8, 12, 24, 32, 40, and 48 hours of competition in the different conditions. At the time of collection, samples were centrifuged and the pellets immediately frozen at -80°C .

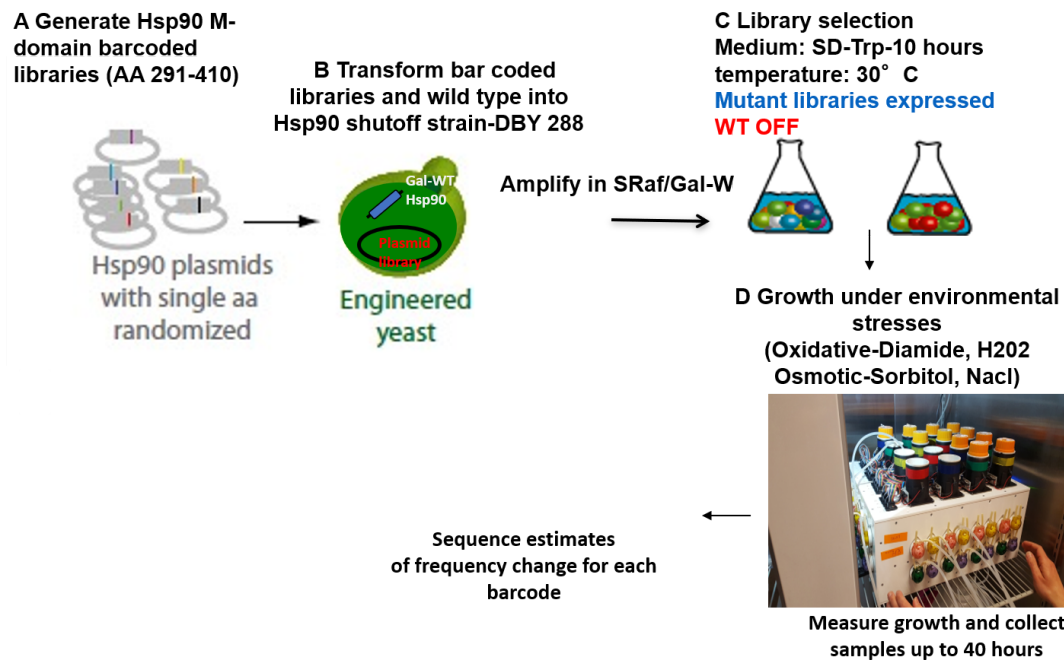


Figure 6: Systematic approach to measure the adaptive potential of the M domain of yeast Hsp90 (amino acids 291-409).

Sequencing of time samples. The fitness effects of Hsp90 variants were estimated based on frequencies observed in next-generation sequencing analyses essentially as previously described (Hietpas et al. 2012). DNA was isolated from each timepoint sample as described (Hietpas et al. 2013) and the barcodes were amplified and sequenced. Barcodes were amplified using the standard Illumina PE1 primer that anneals next to the N18 barcode and custom designed barcode_forward primers. A set of barcode_forward primers were designed with identifier sequences that could be read during sequencing and used to distinguish each timepoint sample. Each identifier sequence was eight nucleotides in length and differed by at least two bases from all other identifier sequences. Twenty cycles of PCR were sufficient to generate clear products from all samples. These PCR products were purified on a silica column (Zymo Research) and sequenced for 100 base single reads on an Illumina NextSeq instrument. The barcode corresponded to the first 18 bases and the identifier at positions 91-98. The resulting fastq files were processed and analyzed using customized software tools. First, poor quality reads containing any positions with a Phred score less than 20 were discarded. Reads were tabulated if the barcode matched a barcode associated with a point

mutation and if the identifier matched with a timepoint. The analyses scripts output a file with the number of reads for each amino acid point mutant at each timepoint in each condition (Supplementary excel file-Hsp90_291_409_BarcodeAssociationTable). Barcodes with 0 reads at the first time point, or with a total of 1 read along the whole trajectory were removed from the analysis.

Estimation of selection coefficients. Inference of selection coefficients was performed via log-linear regression as described in Matuszewski et al 2016. To improve estimation accuracy by incorporating information from each individual barcode, the linear model for each amino-acid changing mutation included barcode identities as nominal variables. For each amino-acid changing mutation we obtained a selection coefficient and 95% confidence interval (CI) of the estimates, representing the variation within amino acid due to differences in codon*Barcode tag throughout time. This allowed to test for consistency of barcode trajectory through time, reducing the impact of potential outliers and also to account for repeatability of synonymous codons within amino acid. Mutants with 50 or less reads at the first time point were removed from the data set. Finally, we normalized all selection coefficients, by subtracting the median of all mutations that were synonymous to the wild type. This ensures that the average of the selection coefficients of wild type synonyms represents a selection coefficient of 0. To categorize mutations as beneficial, deleterious or neutral we tested the overlap of the CI with 0. Namely, if the lower CI limit was larger than 0, a mutation was considered beneficial, and if the upper CI limit was smaller than 0, a mutation was considered deleterious. All other mutations were considered wild-type like.

Costs of adaptation. To compute the costs of adaptation for Figure 4 and 5, we first selected all mutations in a focal environment that were categorized as beneficial. We then computed the proportion of this subset of mutations that was deleterious either in one other environment (for Fig. 4) or across all other environments (excluding replicates of the focal environments and low-expression environments; for Fig. 5). Calculating the proportion of mutations with costs of adaptation compensates for variable numbers of beneficial mutations across environments, which occur due to both biological and experimental reasons.

Acknowledgements

This work was supported in part by grant R01GM112844 from the National Institutes of Health (to DNAB). IF was supported by a postdoctoral fellowship from the FCT (Fundação para a Ciência e a Tecnologia) within the project JPIAMR/0001/2016. CB is grateful for support by EMBO Installation Grant IG4152 and by ERC Starting Grant 804569 - FIT2GO.

References

- Acevedo, A., Brodsky, L., & Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature*, 505(7485), 686–690.
- Arribas, M., K. Kubota, L. Cabanillas, and E. Lazaro. 2014. Adaptation to fluctuating temperatures in an RNA virus is driven by the most stringent selective pressure. *PLoS One* 9:e100940.
- Bank, C., Hietpas, R.T., Wong, A., Bolon, D.N. and Jensen, J.D., 2014 A Bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: uncovering the potential for adaptive walks in challenging environments, *Genetics* 196(3): 841–852.
- Barrett, R. D. and D. Schluter. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38-44.
- Bataillon, T., Zhang, T., & Kassen, R. (2011). Cost of adaptation and fitness effects of beneficial mutations in *Pseudomonas fluorescens*. *Genetics*, 189(3), 939–949.
- Bataillon T., Bailey S.F. 2014 Effects of new mutations on fitness: insights from models and data. *Ann N Y Acad Sci.* 1320:76-92.
- Borkovich, K. A., F. W. Farrelly, D. B. Finkelstein, J. Taulien, and S. Lindquist. 1989. hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Mol. Cell. Biol.* 9:3919-3930.
- Boucher, J. I., Cote, P., Flynn, J., Jiang, L., Laban, A., Mishra, P., ... Bolon, D. N. (2014). Viewing protein fitness landscapes through a next-gen lens. *Genetics*, 198(2), 461–471.
- Brennan, G. L., N. Colegrave, and S. Collins. 2017. Evolutionary consequences of multidriver environmental change in an aquatic primary producer. *Proc. Natl. Acad. Sci. U. S. A.* 114:9930-9935.
- Chang, H. C. and S. Lindquist. 1994. Conservation of Hsp90 macromolecular complexes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 269:24983-24988.
- Chen, B., M. E. Feder, and L. Kang. 2018. Evolution of heat-shock protein expression underlying adaptive responses to environmental stress. *Mol. Ecol.* 27:3040-3054.
- Cunningham, C. N., K. A. Krukenberg, and D. A. Agard. 2008. Intra- and intermonomer interactions are required to synergistically facilitate ATP hydrolysis in Hsp90. *J. Biol. Chem.* 283:21170-21178.
- Czemes, J., K. Buse, and G. M. Verkhivker. 2017. Atomistic simulations and network-based modeling of the Hsp90-Cdc37 chaperone binding with Cdk4 client protein: A mechanism of chaperoning kinase clients by exploiting weak spots of intrinsically dynamic kinase domains. *PLoS One* 12:e0190267.
- de Nadal, E., G. Ammerer, and F. Posas. 2011. Controlling gene expression in response to stress. *Nature reviews. Genetics* 12:833-845.

- Dekel, E. and U. Alon. 2005. Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436:588-592.
- Dhar, R., R. Sagesser, C. Weikert, J. Yuan, and A. Wagner. 2011. Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *J. Evol. Biol.* 24:1135-1153.
- Doud, M. and Bloom, J., 2016 Accurate Measurement of the Effects of All Amino-Acid Mutations on Influenza Hemagglutinin, *Viruses* 8(6), 155.
- Eyre-Walker, A. and Keightley, P.D. 2007 The distribution of fitness effects of new mutations. *Nature Review Genetics* 8(8): 610.
- Firnberg E, Labonte JW, Gray JJ, Ostermeier M. 2014 A Comprehensive, High-Resolution Map of a Gene's Fitness Landscape. *Mol Biol Evol.*
- Fitzgerald, D. M. and S. M. Rosenberg. 2019. What is mutation? A chapter in the series: How microbes "jeopardize" the modern synthesis. *PLoS Genet.* 15:e1007995.
- Flynn, J. M., Rossouw, A., Cote-Hammarlof, P. A., Fragata, I., Mavor, D., Hollins, C., Bank, C. & Bolon, D. N. (2019). Investigating the influence of environment on the evolution of Hsp90 using comprehensive fitness maps. *bioRxiv*, 823468.
- Fowler, D. M., Araya, C. L., Fleishman, S. J., Kellogg, E. H., Stephany, J. J., Baker, D., & Fields, S. (2010). High-resolution mapping of protein sequence-function relationships. *Nature methods*, 7(9), 741–746.
- Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein, and P. O. Brown. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11:4241-4257.
- Genest, O., M. Reidy, T. O. Street, J. R. Hoskins, J. L. Camberg, D. A. Agard, D. C. Masison, and S. Wickner. 2013. Uncovering a region of heat shock protein 90 important for client binding in *E. coli* and chaperone function in yeast. *Mol. Cell* 49:464-473.
- Gietz, R. D. and R. H. Schiestl. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* 2:31-34.
- Gopinath, R. K., S. T. You, K. Y. Chien, K. B. Swamy, J. S. Yu, S. C. Schuyler, and J. Y. Leu. 2014. The Hsp90-dependent proteome is conserved and enriched for hub proteins with high levels of protein-protein connectivity. *Genome Biol. Evol.* 6:2851-2865.
- Hagn, F., S. Lagleder, M. Retzlaff, J. Rohrberg, O. Demmer, K. Richter, J. Buchner, and H. Kessler. 2011. Structural analysis of the interaction between Hsp90 and the tumor suppressor protein p53. *Nat. Struct. Mol. Biol.* 18:1086-1093.
- Hawle, P., M. Siepmann, A. Harst, M. Siderius, H. P. Reusch, and W. M. Obermann. 2006. The middle domain of Hsp90 acts as a discriminator between different types of client proteins. *Mol. Cell. Biol.* 26:8385-8395.
- Hiatt, J. B., R. P. Patwardhan, E. H. Turner, C. Lee, and J. Shendure. 2010. Parallel, tag-directed assembly of locally derived short sequence reads. *Nature methods* 7:119-122.
- Hietpas, R., B. Roscoe, L. Jiang, and D. N. Bolon. 2012. Fitness analyses of all possible point mutations for regions of genes in yeast. *Nat. Protoc.* 7:1382-1396.
- Hietpas, R. T., C. Bank, J. D. Jensen, and D. N. A. Bolon. 2013. Shifting fitness landscapes in response to altered environments. *Evolution* 67:3512-3522.

- Hom N., Gentles L., Bloom J.D., and Lee K.K. 2019 Deep Mutational Scan of the Highly Conserved Influenza A Virus M1 Matrix Protein Reveals Substantial Intrinsic Mutational Tolerance. *Journal of Virology* 93(13):e00161-19.
- Jarosz, D. F. and S. Lindquist. 2010. Hsp90 and environmental stress transform the adaptive value of natural genetic variation. *Science* 330:1820-1824.
- Jiang, L., Mishra, P., Hietpas, R. T., Zeldovich, K. B., & Bolon, D. N. (2013). Latent effects of Hsp90 mutants revealed at reduced expression levels. *PLoS genetics*, 9(6), e1003600.
- Kaplan, K. B. and R. Li. 2012. A prescription for 'stress'--the role of Hsp90 in genome stability and cellular adaptation. *Trends Cell Biol.* 22:576-583.
- Karagoz, G. E., A. M. Duarte, E. Akoury, H. Ippel, J. Biernat, T. Moran Luengo, M. Radli, T. Didenko, B. A. Nordhues, D. B. Veprintsev, C. A. Dickey, E. Mandelkow, M. Zweckstetter, R. Boelens, T. Madl, and S. G. Rudiger. 2014. Hsp90-Tau complex reveals molecular basis for specificity in chaperone action. *Cell* 156:963-974.
- Karagoz, G. E. and S. G. Rudiger. 2015. Hsp90 interaction with clients. *Trends Biochem. Sci.* 40:117-125.
- Kvitek DJ, Sherlock G 2013 Whole Genome, Whole Population Sequencing Reveals That Loss of Signaling Networks Is the Major Adaptive Strategy in a Constant Environment. *PLOS Genetics* 9(11): e1003972
- Lindquist, S. 2009. Protein folding sculpting evolutionary change. *Cold Spring Harb. Symp. Quant. Biol.* 74:103-108.
- Loewe, L. 2010 The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 1153-1167.
- Manceau, M., V. S. Domingues, R. Mallarino, and H. E. Hoekstra. 2011. The developmental role of Agouti in color pattern evolution. *Science* 331:1062-1065.
- Matuszewski, S., Hildebrandt, M. E., Ghenu, A.-H., Jensen, J. D., and Bank, C. 2016 A Statistical Guide to the Design of Deep Mutational Scanning Experiments, *Genetics* 204(1), 77–87.
- McClellan, A. J., Y. Xia, A. M. Deutschbauer, R. W. Davis, M. Gerstein, and J. Frydman. 2007. Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches. *Cell* 131:121-135.
- Melamed D., Young D. L., Gamble C. E., Miller C. R., Fields S., 2013. Deep mutational scanning of an RRM domain of the *Saccharomyces cerevisiae* poly(A)-binding protein. *RNA* 19: 1537–1551.
- Meyer, P., C. Prodromou, B. Hu, C. Vaughan, S. M. Roe, B. Panaretou, P. W. Piper, and L. H. Pearl. 2003. Structural and functional analysis of the middle segment of hsp90: implications for ATP hydrolysis and client protein and cochaperone interactions. *Mol. Cell* 11:647-658.
- Mumby, P. J. and R. van Woesik. 2014. Consequences of ecological, evolutionary and biogeochemical uncertainty for coral reef responses to climatic stress. *Curr. Biol.* 24:R413-423.
- Mustonen V and Lässig M. 2009 From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. *Trends Genet.* 25(3):111-9.

- Nathan, D. F. and S. Lindquist. 1995. Mutational analysis of Hsp90 function: interactions with a steroid receptor and a protein kinase. *Mol. Cell. Biol.* 15:3917-3925.
- Nathan, D. F., M. H. Vos, and S. Lindquist. 1997. In vivo functions of the *Saccharomyces cerevisiae* Hsp90 chaperone. *Proc. Natl. Acad. Sci. U. S. A.* 94:12949-12956.
- Ohta T. 1992 The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics* 23(1):263-286.
- Park, S. J., B. N. Borin, M. A. Martinez-Yamout, and H. J. Dyson. 2011a. The client protein p53 adopts a molten globule-like state in the presence of Hsp90. *Nat. Struct. Mol. Biol.* 18:537-541.
- Park, S. J., M. Kostic, and H. J. Dyson. 2011b. Dynamic Interaction of Hsp90 with Its Client Protein p53. *J. Mol. Biol.* 411:158-173.
- Reznick, D.N. and Ghalambor, C.K. 2005 Selection in Nature: Experimental Manipulations of Natural Populations, *Integrative and Comparative Biology*, 45(3):456–462,
- Richter, K., M. Haslbeck, and J. Buchner. 2010. The heat shock response: life on the verge of death. *Mol. Cell* 40:253-266.
- Rohl, A., J. Rohrberg, and J. Buchner. 2013. The chaperone Hsp90: changing partners for demanding clients. *Trends Biochem. Sci.* 38:253-262.
- Rutherford, S., Y. Hirate, and B. J. Swalla. 2007. The Hsp90 capacitor, developmental remodeling, and evolution: the robustness of gene networks and the curious evolvability of metamorphosis. *Crit. Rev. Biochem. Mol. Biol.* 42:355-372.
- Rutherford, S. L. 2003. Between genotype and phenotype: protein chaperones and evolvability. *Nature reviews. Genetics* 4:263-274.
- Sane, M. , Miranda, J. J. and Agashe, D. (2018), Antagonistic pleiotropy for carbon use is rare in new mutations. *Evolution*, 72: 2202-2213.
- Sangster, T. A., S. Lindquist, and C. Queitsch. 2004. Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance. *Bioessays* 26:348-362.
- Sarkisyan KS, Bolotin DA, Meer MV, Usmanova DR, Mishin AS, Sharonov GV, Ivankov DN, Bozhanova NG, Baranov MS, Soylemez O, Bogatyreva NS, Vlasov PK, Egorov ES, Logacheva MD, Kondrashov AS, Chudakov DM, Putintseva EV, Mamedov IZ, Tawfik DS, Lukyanov KA, Kondrashov FA. 2016 Local fitness landscape of the green fluorescent protein. *Nature* 533(7603):397-401.
- Sato, S., N. Fujita, and T. Tsuruo. 2000. Modulation of Akt kinase activity by binding to Hsp90. *Proc. Natl. Acad. Sci. U. S. A.* 97:10832-10837.
- Sniegowski P.D. and Gerrish P.J. 2010 Beneficial mutations and the dynamics of adaptation in asexual populations. *Philos Trans R Soc Lond B Biol Sci.* 365(1544):1255-63.
- Soskine M, Tawfik DS. 2010 Mutational effects and the evolution of new protein functions. *Nat Rev Genet*; 11(8):572–82
- Street, T. O., L. A. Lavery, K. A. Verba, C. T. Lee, M. P. Mayer, and D. A. Agard. 2012. Cross-monomer substrate contacts reposition the Hsp90 N-terminal domain and prime the chaperone activity. *J. Mol. Biol.* 415:3-15.
- Taipale, M., D. F. Jarosz, and S. Lindquist. 2010. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* 11:515-528.

- Taipale, M., I. Krykbaeva, M. Koeva, C. Kayatekin, K. D. Westover, G. I. Karras, and S. Lindquist. 2012. Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition. *Cell* 150:987-1001.
- Verba, K. A., R. Y. Wang, A. Arakawa, Y. Liu, M. Shirouzu, S. Yokoyama, and D. A. Agard. 2016. Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science* 352:1542-1547.
- Xu, W., K. Beebe, J. D. Chavez, M. Boysen, Y. Lu, A. D. Zuehlke, D. Keramisanou, J. B. Trepel, C. Prodromou, M. P. Mayer, J. E. Bruce, I. Gelis, and L. Neckers. 2019. Hsp90 middle domain phosphorylation initiates a complex conformational program to recruit the ATPase-stimulating cochaperone Aha1. *Nature communications* 10:2574.
- Xu, Y., M. A. Singer, and S. Lindquist. 1999. Maturation of the tyrosine kinase c-src as a kinase and as a substrate depends on the molecular chaperone Hsp90. *Proc. Natl. Acad. Sci. U. S. A.* 96:109-114.
- Zabinsky, R. A., G. A. Mason, C. Queitsch, and D. F. Jarosz. 2019. It's not magic - Hsp90 and its effects on genetic and epigenetic variation. *Semin. Cell Dev. Biol.* 88:21-35.
- Zhang, R., D. Luo, R. Miao, L. Bai, Q. Ge, W. C. Sessa, and W. Min. 2005. Hsp90-Akt phosphorylates ASK1 and inhibits ASK1-mediated apoptosis. *Oncogene* 24:3954-3963.
- Zhao, R., M. Davey, Y. C. Hsu, P. Kaplanek, A. Tong, A. B. Parsons, N. Krogan, G. Cagney, D. Mai, J. Greenblatt, C. Boone, A. Emili, and W. A. Houry. 2005. Navigating the chaperone network: an integrative map of physical and genetic interactions mediated by the hsp90 chaperone. *Cell* 120:715-727.