Pervasive contingency and entrenchment in a billion years of Hsp90 evolution

Tyler N. Starr (1*), Julia M. Flynn (2*), Parul Mishra (2*), Daniel N. A. Bolon (2+), Joseph W. Thornton (3+)

(1) Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA

(2) Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA

(3) Departments of Ecology & Evolution and Human Genetics, University of Chicago, Chicago, IL 60637, USA

* co-first authors

+ co-senior authors

Editorial correspondence: Joseph W. Thornton, joet1@uchicago.edu, 1-773-834-3423

Although many potential mutations within proteins modulate each other's effects[1-4], the extent to which these epistatic interactions influenced the fitness effects of the sequence changes that actually occurred during historical evolution – and thus made molecular evolution contingent and irreversible – is controversial[5-16]. We addressed this question directly by precisely measuring the fitness effects in both extant and reconstructed ancestral sequence contexts of all historical amino acid substitutions that occurred during the billion-year evolutionary history of the heat shock protein 90 (Hsp90) ATPase domain beginning from a deep eukaryotic ancestor to modern *Saccharomyces cerevisiae*. We find a pervasive influence of epistasis on historical sequence evolution: the majority of the 98 derived states that evolved during history were deleterious at times before they happened, and the vast majority also became subsequently entrenched[6], with the ancestral state becoming deleterious after its substitution. A few of these epistatic interactions were of massive fitness consequence, but the majority were of small but evolutionarily relevant effect size. We find that both the large- and small-effect epistasis were largely caused by specific interactions among sites rather than a general permissive or restrictive effect[17]. Our results highlight how epistasis continually opens and closes windows of mutational opportunity over evolutionary timescales, producing histories and biological states that reflect, in significant part, the transient internal constraints imposed by a protein's fleeting sequence states.

Epistatic interactions can affect the sequence changes that accumulate during evolution. An otherwise deleterious mutation may be fixed under purifying selection if a permissive substitution renders it neutral or beneficial[6,15,18]. Further, a neutral mutation, initially reversible, may be entrenched by a restrictive substitution that renders the ancestral state deleterious.[5,6]. The

extent to which epistasis-induced contingency and entrenchment have affected protein sequence evolution remains controversial, because there is no consensus on the prevalence, effect size, or mechanisms of epistasis among historical substitutions. Historical case studies have shown that particular substitutions were contingent[18-21] or entrenched[22], but whether these are examples of a general phenomenon is unknown. Computational analyses suggest pervasive contingency and entrenchment among substitutions[5,6,9,11,12,23,24], but some of these approaches rely on models of uncertain adequacy[7,10,13], and their claims have not been experimentally tested. Swapping sequence states among extant orthologs reveals frequent epistasis among substitutions[14], but this "horizontal" approach, unpolarized with respect to time, leaves unresolved whether permissive or restrictive interactions are at play[25]. Some experimental studies have examined epistasis among substitutions in an historical context, but they have measured effects on protein function[14,15] or stability[7,8] or have focused on large-effect interactions[16], leaving epistatic modulations of effects on fitness, particularly those of relatively small magnitude, unexamined.

We directly evaluated the roles of contingency and entrenchment on historical sequence evolution by precisely quantifying changes over time in the fitness effects of all substitutions that accumulated during the long-term evolution of heat shock protein 90 (Hsp90) from a deep eukaryotic ancestor to *S. cerevisiae*. We used a deep sequencing-based bulk fitness assay to characterize protein libraries in which each ancestral amino acid is reintroduced into an extant Hsp90 at all substituted sites and each derived state is introduced into a reconstructed ancestral Hsp90. Hsp90 is an essential molecular chaperone that facilitates folding and regulation of substrate proteins through an ATP-dependent cycle of conformational changes, modulated by co-chaperone proteins. Orthologs from other fungi, animals, and protists can complement Hsp90 deletion in *S. cerevisiae*[26,27], indicating that the protein's essential molecular function is

conserved over large evolutionary distances. We focused our experiments on the N-terminal domain (NTD) of Hsp90, which mediates ATP-dependent conformational changes.

We first reconstructed the historical trajectory of Hsp90 sequence evolution. We inferred the maximum likelihood phylogeny of Hsp90 protein sequences from 267 species of Amorphea (the clade comprising Fungi, Metazoa, Amoebozoa, and related lineages[28]), rooted using green algae and plants as an outgroup (Fig. 1a, Extended Data Fig. 1, Supplementary Data Table 1 & File 1,2). We inferred ancestral sequences at all nodes along the trajectory from the common ancestor of Amorphea (ancAmoHsp90) to extant *S. cerevisiae* (ScHsp90) and identified substitutions as differences between the most probable reconstructions at successive nodes (Supplementary Data File 3). Along this entire trajectory, substitutions occurred at 72 of the 221 sites in the NTD; because of multiple substitutions, 98 unique ancestral amino acid states existed at these sites at some point in the past and have since been replaced by the ScHsp90 state (Supplementary Data File 3). The vast majority of these 98 ancestral states are reconstructed with high confidence (posterior probability >0.95) in one or more ancestors along the trajectory (Fig. 1b), and every ancestral sequence has a mean posterior probability across sites of >0.95 (Extended Data Fig. 2a-c).

To measure the fitness effects of ancestral amino acids when re-introduced into an extant Hsp90, we created a library of ScHsp90 NTD variants, each of which contains one of the 98 ancestral states (Extended Data Fig. 3a). We determined the per-generation selection coefficient (s, log relative fitness) of each mutation to an ancestral state relative to ScHsp90 via bulk competition monitored by deep sequencing (Extended Data Fig. 3a), a technique with highly reproducible results (Extended Data Fig. 3b,c). We found that the vast majority of reversions to ancestral states in ScHsp90 are deleterious (Fig. 1c): after accounting for experimental noise in

fitness measurements, an estimated 93% of all ancestral states reduce ScHsp90 fitness (95% CI 83% to 100%; Extended Data Fig. 4). Two ancestral states cause very strong fitness defects (s = -0.38 and -0.54), but the typical reversion is mildly deleterious (median s = -0.010, Fig. 1c), a conclusion that is robust to excluding ancestral states that are reconstructed with any statistical ambiguity. The magnitude of each mutation's negative effect on fitness correlates with indicators of site-specific evolutionary, structural, and functional constraint, corroborating the view that they are authentically deleterious (Extended Data Fig. 5). These results do not imply that reversions can never happen—12 sites did undergo substitution and reversion from ancAmoHsp90 to ScHsp90. Rather, our observations indicate that at the current moment in time, the vast majority of ancestral states are selectively inaccessible, irrespective of whether they were available at some moment in the past or might become so in the future.

Reversions to ancestral states might be deleterious because the derived states were entrenched by subsequent substitutions within Hsp90 (intramolecular epistasis)[5,6]; they might also be incompatible with derived states at other loci in the *S. cerevisiae* genome (intermolecular epistasis), or the derived state might unconditionally increase fitness. Entrenchment because of intramolecular epistasis predicts that introducing into ScHsp90 sets of deleterious ancestral states that existed together at ancestral nodes should not reduce fitness as drastically as predicted from the individual mutations' effects. We therefore reconstructed complete ancestral NTDs from two widely spaced ancestral Hsp90s on the phylogenetic trajectory (Fig. 2a) and assayed their relative fitness in *S. cerevisiae.* The first, ancAmoHsp90 (estimated age ~1 billion years) differs from ScHsp90 at 60 sites; if the fitness effects of the ancestral states when combined were the same as when introduced individually, they would confer an expected fitness of just 0.23 (95% CI 0.21 to 0.26; Fig. 2b, Extended Data Fig. 6a). When introduced together, however, the actual

fitness is 0.96 (Fig. 2d, Extended Data Fig. 2d, 6a). The second reconstruction, ancAscoHsp90, from the ancestor of Ascomycota fungi (estimated age ~450 million years) differs from ScHsp90 at 42 sites and has expected fitness 0.65 (95% CI 0.61 to 0.69; Fig. 2b, Extended Data Fig. 6a), but its actual fitness is 0.99 (Fig. 2d, Extended Data Fig. 2d, 6a). The near-universal fitness defects we observed when individual ancestral states were introduced into the extant ScHsp90 are therefore caused almost entirely by intramolecular epistasis: derived states that emerged along the Hsp90 trajectory have been entrenched by subsequent substitutions within the same protein, which closed the direct path back to the ancestral amino acid without causing major changes in function or fitness[6].

We next determined whether paths to derived states were contingent on prior permissive substitutions during the protein's evolutionary history. We constructed a library of variants of ancAmoHsp90, the deepest ancestor of the trajectory, each of which contains one of the 98 forward mutations to a derived state. We cloned this library into yeast and used our deep sequencing-based bulk fitness assay to measure the selection coefficient of each mutation relative to ancAmoHsp90 (Extended Data Fig. 3d,e). We found that most mutations to derived states were selectively unfavorable (Fig. 3a): after accounting for experimental noise in fitness measurements, an estimated 53% of derived states reduce ancAmoHsp90 fitness (95% CI 27% to 96%), and 32% are neutral (CI 0% to 59%; Extended Data Fig. 7). Fifteen percent of the derived states are beneficial in our assay (95% CI 3% to 57%), which could be because they are unconditionally advantageous or because of epistatic interactions with other loci in the yeast system we used for our fitness assay. Two derived states had very strong fitness defects, but the typical derived state is weakly deleterious (median s = -0.005, Fig. 3a). As with the reversions to ancestral states, the effect of all the individual derived states, as measured in the ancestral

background, predicts fitness consequences far greater than observed when the derived states are combined into the Hsp90 genotypes that occurred along the phylogeny (Fig. 2c,d, Extended Data Fig. 6b). Thus, most derived states would have been deleterious if they had occurred in the ancestral background, but they became accessible following subsequent permissive substitutions that occurred within Hsp90. The data from the two libraries, taken together, indicate that 83% of the amino acid states that occurred along this evolutionary trajectory were contingent on prior permissive substitutions, entrenched by subsequent restrictive substitutions, or both (Fig. 3b).

Epistatic effects on fitness can emerge from specific genetic interactions between substitutions that directly modify each other's effect on some important molecular property, or from nonspecific interactions between substitutions that are additive with respect to bulk molecular properties (e.g. stability[15,29]) if those properties nonlinearly affect fitness[17,30]. To explore which type of epistasis predominates in the long-term evolution of Hsp90, we first investigated the two strongest cases of entrenchment, the strongly deleterious reversions V23f and E7a (with upper-case letters indicating the ScHsp90 state and lower-case the ancestral states). We sought candidate restrictive substitutions for each of these large-effect reversions by examining patterns of phylogenetic co-occurrence. Substitution f23V occurred not only along the trajectory from ancAmoHsp90 to ScHsp90 but also in parallel on another fungal lineage; in both cases, candidate epistatic substitution i378L co-occurred on the same branch (Extended Data Fig. 8a,b). As predicted if i378L entrenched f23V, we found that introducing the ancestral state i378 in ScHsp90 relieves the deleterious effect of the ancestral state f23 (Fig. 4a). These two residues directly interact in the protein's tertiary structure to position a key residue in the ATPase active site (Extended Data Fig. 8c,d). In the case of E7a, the other strongly deleterious reversion in ScHsp90, the ancestral state was actually reacquired in a closely related fungal lineage; on this

branch, two candidate epistatic modifiers that occurred previously (n13T and a151N) either reverted or were further modified on the same branch (Extended Data Fig. 9a,b,c). As predicted, experimentally introducing the ancestral states n13 or a151 into ScHsp90 relieves much of the fitness defect caused by the ancestral state a7, indicating that substitutions n13T and a151N entrenched a7E (Fig. 4b). These three sites are on interacting secondary structural elements that are conformationally rearranged when Hsp90 converts between ADP- and ATP-bound states. (Extended Data Fig. 9d,e). To test whether these modifiers specifically restrict particular substitutions or are general epistatic modifiers, we asked whether the restrictive substitutions that entrenched one substitution also modify the effects of the other[15]. As predicted if the interactions among these sets of substitutions are specific, introducing L378i does not ameliorate the fitness defect caused by E7a, and introducing T13n or N151a does not ameliorate the fitness defect caused by V23f (Fig. 4c,d). These data indicate that specific biochemical mechanisms underlie the restrictive interactions for these large-effect examples of epistatic entrenchment.

Finally, we investigated whether the epistatic interactions among the large set of small-effect substitutions in this trajectory are also specific or the nonspecific result of a threshold-like relationship between fitness and some bulk property such as stability[15,29]. If epistasis is mediated by a nonspecific threshold relationship, mutations that decrease fitness in one background will never be beneficial in another, although they can be neutral if buffered by the threshold (Fig. 4e)[7,15,16]. In contrast, specific interactions can switch the sign of a mutation's selection coefficient in different sequence contexts (Fig. 4f)[17]. As predicted under specific epistasis, we found that for many differences between ancAmoHsp90 and ScHsp90 (65%), the ancestral state confers increased fitness relative to the derived state in the ancestral background but decreases it in the extant background (Fig. 4g). Further, selection coefficients are negatively correlated

between backgrounds (P=0.009), indicating that the substitutions that have become most entrenched in the present also required the strongest permissive effect in the past. This pattern is expected if the structural constraints that determine the selective cost of having a suboptimal state at some site are conserved over time, but the specific states preferred at the site depend on the broader protein sequence context.

The widespread and specific epistasis that we observed took place over the course of a billion years during which function and fitness were conserved. The fraction of historical substitutions in Hsp90 that were either contingent on permissive substitutions or entrenched by restrictive substitutions – about 83% – is considerably higher than suggested by previous experimental studies [8,14,15], and some computational analyses[23], rivaling some of the highest estimates from computational modeling[6,9,12]. One explanation for the greater prevalence of epistatic interactions that we observed may be our method's capacity to detect much smaller fitness deficits than have been discernable in previous experimental studies. Another difference from previous research is that we observed primarily specific epistasis, whereas several studies have found a dominant role for nonspecific epistasis, particularly in short-term viral evolution[7,15]. This difference could be attributable to a difference in selective regime or in time scale: the epistatic constraints caused by specific interactions are expected to be maintained over far longer periods of time than those caused by nonspecific interactions, which are easily replaced by other substitutions because of the many-to-many relationship between permissive and permitted amino acid states[7,17].

Our observation of pervasive and specific contingency and entrenchment suggest a daisy-chain model by which genetic interactions structured long-term Hsp90 evolution (Fig. 4h). A permissive mutation becomes entrenched and irreversible once a substitution contingent upon it

occurs; if the contingent substitution subsequently permits a third substitution, it then becomes entrenched, as well[6,12]. Most of the substitutions along the trajectory from ancAmoHsp90 to ScHsp90 are both contingent and entrenched, suggesting they occupy an internal position in this daisy chain of contingency and entrenchment. Thus, virtually every historical substitution altered the set of accessible evolutionary trajectories, closing reverse paths at some sites and opening forward paths at others, which – if taken – would then entrench the previous step. Evolving this way over long periods of time, proteins come to appear exquisitely well adapted to the conditions of their existence, with virtually all present states superior to past ones, but the conditions that make these states fit include – and may even be dominated by – the transient internal organization of the protein itself.

Methods

Phylogenetic analysis and ancestral reconstruction. We obtained Hsp90 protein sequences from the Amorphea clade[28] from NCBI, the JGI Fungal Program, the Broad Institute Multicellularity Project, the literature[31] and Iñaki Ruiz-Trello (personal communication). Full identifiers and sources of sequences are listed in Supplementary Table 1. Each protein was used as a query in a BLASTp search against the human proteome to identify and retain Hsp90A paralogs. We used CD-HIT[32] to filter proteins with high sequence similarity. We removed sequences with >67% missing characters and highly diverged, unalignable sequences. Remaining sequences were aligned with Clustal Omega[33]. Lineage-specific insertions were removed, as were unalignable linker regions (ScHsp90 sites 1-3, 225-237, 686-701). We added six Hsp90A sequences from Viridiplantae as an outgroup, resulting in a final alignment of 267 protein sequences and 680 sites (Supplementary Data File 1).

We inferred the maximum likelihood (ML) phylogeny (Supplementary Data File 2) given our alignment and the LG model[34] with gamma-distributed among-site rate variation (4 categories) and ML estimates of amino acid frequencies, which was the best-fit model as judged by AIC. The phylogeny was inferred using RAxML version 8.1.17 (ref 35). The ML phylogeny reproduces accepted relationships between major taxonomic lineages[28,36-40]. Most probable ancestral sequences (Supplementary Data File 3) were reconstructed on the maximum likelihood phylogeny using the AAML module of PAML version 4.4 (ref 41) given the alignment, ML phylogeny, and LG+Γ model. The trajectory of sequence change was enumerated from the amino acid sequence differences between successive ancestral nodes on the lineage from the common ancestor of Amoebozoa + Opisthokonta (ancAmorphea) to *S. cerevisiae* Hsp82 (ScHsp90, Uniprot P02829). Ancestral states are defined as amino acid states not present in

ScHsp90 that occurred in at least one ancestral node on the lineage from ancAmorphea to ScHsp90. Derived states are defined as amino acid states not present in the reconstructed ancAmorphea sequence that occurred in at least one descendent node on the lineage to ScHsp90.

Coding sequences for the most probable ancestral amino acid sequences of the Hsp90 N-terminal domain (NTD) from ancAmorphea (ancAmoHsp90, estimated age at least 1 billion years[42]) and the common ancestor of Ascomycota yeast (ancAscoHsp90, estimated age ~450 million years[43]) were synthesized by IDT (Supplementary Data Table 2). These sequences were cloned as chimeras with the ScHsp90 middle and C-terminal domains and intervening linkers via Gibson Assembly. AncAmoHsp90 also carries an additional reversion to the ancAmorphea state at site 378 in the middle domain (Extended Data Fig. 2d), which is part of a loop that extends down and interacts with ATP and the NTD[44,45].

Generating mutant libraries. ScHsp90 and ancAmoHsp90 gene constructs were expressed from the p414ADHΔTer plasmid[46]. The ScHsp90 library consists of variants of the ScHsp90 NTD, each containing one mutation to an ancestral amino acid state. The ancAmoHsp90 library consists of variants of the ancAmoHsp90 NTD, each containing one mutation to a derived state. Two sets of PCR primers were designed for each mutation, to amplify Hsp90 NTD fragments N-terminal and C-terminal to the mutation of interest; primers introduce the mutation of interest and generate a 25-bp overlap between fragments, as well as 20-bp overlaps between each fragment and the destination vector for gene re-assembly (Supplementary Data Table 2). PCR was conducted with Pfu Turbo polymerase (Agilent) for 15 amplification cycles. The resulting PCR fragments were stitched together with a 10-cycle assembly PCR, pooled, and combined via Gibson Assembly (NEB) with a linearized p414ADHΔTer Hsp90 destination vector excised of the NTD.

Barcode labeling of library genotypes. Following construction of the plasmid libraries, each library variant was tagged with a unique barcode to simplify sequencing steps during bulk competition[47]. A pool of DNA constructs containing a randomized 18 base-pair barcode sequence (N18) and Illumina sequencing primer annealing regions (IDT; Supplementary Data Table 2) was cloned 200 nucleotides downstream from the hsp90 stop codon via restriction digestion, ligation, and transformation into chemically-competent E. *coli*. Cultures with different amounts of the transformation reaction were grown overnight and the colony forming units in each culture were assessed by plating a small fraction. We isolated DNA from the transformation that contained approximately 10-20 fold more colony-forming units than mutants, with the goal that each mutant would be represented by 10-20 unique barcodes.

To associate barcodes with Hsp90 mutant alleles, we conducted paired end sequencing of each library using primers that read the N18 barcode in the first read and the Hsp90 NTD in the other (Supplementary Data Table 2). To generate short DNA fragments from the plasmid library that would be efficiently sequenced, we excised the gene region between the NTD and the N18 barcode via restriction digest, followed by blunt ending with T4 DNA polymerase (NEB) and plasmid ligation at a low concentration (3 ng/μL) that favors circularization over bi-molecular ligations. The resulting DNA was re-linearized by restriction digest, and Illumina adapter sequences were added via an 11-cycle PCR (Supplementary Data Table 2). The resulting PCR products were sequenced using an Illumina MiSeq instrument with asymmetric reads of 50 bases and 250 bases for Read1 and Read2 respectively. After filtering low quality reads (Phred scores < 10), the data were organized by barcode sequence. For each barcode that was read more than 3 times, we generated a consensus sequence of the N-domain indicating the mutation that it contained.

Bulk growth competitions. For bulk fitness assessments, we transformed *S. cerevisiae* with the ScHsp90 library along with wildtype ScHsp90 and a no-insert control; we also transformed *S.* cerevisiae with the ancAmoHsp90 library along with wildtype ScHsp90, wildtype ancAmoHsp90, and a no-insert control. Concentrations of plasmids were adjusted to yield a 2:6:1 molar ratio of wildtype: no-insert control: average library variant. Plasmid libraries and corresponding controls were transformed into the DBY288 Hsp90 shutoff strain[48,49], resulting in ~150,000 unique yeast transformants representing 50-fold sampling for the average barcode. Following recovery, transformed cells were washed 5 times in SRGal-W (synthetic 1% raffinose and 1% galactose lacking tryptophan) media to remove extracellular DNA, and then transferred to plasmid selection media SRGal-W and grown at 30°C for 48 hours with repeated dilution to maintain the cells in log phase of growth. To select for function of the plasmid-borne Hsp90 allele, cells were shifted to shutoff conditions by centrifugation, washing and re-suspension in 200 mL SD-W (synthetic 2% dextrose lacking tryptophan) media and ampicillin (50µg/mL), and growth at 30°C 225 rpm. Following a 16-hour growth period required to shut off expression of the wildtype chromosomal Hsp90, we collected samples of ~10^8 cells at 8 or more time points over the course of 48 (ScHsp90 library) or 31 (ancAmoHsp90 library) hours and stored them at -80°C. Cultures were maintained in log phase by regular dilution with fresh media, maintaining a population size of 10^9 or greater throughout the bulk competition. Bulk competitions of each library were conducted in duplicate from independent transformations.

DNA preparation and sequencing. We collected plasmid DNA from each bulk competition time point as previously reported[50]. Purified plasmid was linearized with AscI. Barcodes were amplified by 18 cycles of PCR using Phusion polymerase and primers that add Illumina adapter sequences, as well as an 8-bp identifier used to distinguish among libraries and

time points (Supplementary Data Table 2). Identifiers were designed so that each differed by more than two bases from all others to avoid misattributions due to sequencing errors. PCR products were purified two times over silica columns (Zymo research), and quantified using the KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) on a Bio-Rad CFX machine. Samples were pooled and sequenced on an Illumina NextSeq (ancAmoHsp90 library) or HiSeq 2000 (ScHsp90 library) instrument in single-end 100 bp mode.

Analysis of bulk competition sequencing data. Illumina sequence reads were filtered for Phred scores >20, strict matching of the sequence of the intervening bases to the template, and strict matching of the N18 barcode and experimental identifier to those that were expected in the given library. Reads that passed these filters were parsed based on the identifier sequence. For each identifier, the data was condensed by generating a count of each unique N18 read. The unique N18 count file was then used to identify the frequency of each mutant using the variant-barcode association table. For each library variant, the counts of each associated barcode were summed to generate a cumulative count for that mutant.

Determination of selection coefficient. The ratio of the frequency of each library variant relative to wildtype (ancAmoHsp90 or ScHsp90) was determined at each time point, and the slope of the logarithm of this ratio versus time (in number of generations) was determined as the raw per-generation selection coefficient (s)[51]:

$$s = \mathrm{d}/\mathrm{dt}\ [\ \ln(\mathrm{n_m}\ /\ \mathrm{n_{wt}})\ ]$$

where n_m and n_{wt} are the number of sequence reads of mutant and wildtype, respectively, and time is measured in number of wildtype generations. No-insert plasmid selection coefficients were determined from the first three time points because their counts drop rapidly over time. Mutants with selection coefficients within three standard deviations of the mean of no-insert

variants were considered null-like and also analyzed based on the first three time points. For all other variants, selection coefficients were determined from all time points. Final selection coefficients for each variant were scaled in relative fitness space ($w = e^s$) such that the Hsp90 null allele, which is lethal, has a relative fitness of 0 ($s = -\infty$)[52].

Generation of individual mutants and monoculture analysis of yeast growth. To measure the relative fitness of ancAscoHsp90, mutations missed in the bulk libraries, and genotypes in mutant cycles that we sought to test in combination for epistatic interactions, we assayed growth rate in monoculture and related this to fitness, which assumes the relative rate of growth of two genotypes is the same in isolation as in direct competition[51]. The growth rate of individually cloned mutants was estimated over 30 hours of growth with periodic dilution to maintain log-phase growth, as per Jiang et al.[46]. Growth rates were determined as the slope of the linear model relating the log-transformed dilution-corrected cell density to time. The growth rate was converted to an estimate of the selection coefficient by taking the difference in growth rate (Malthusian parameter) between mutant and wildtype and multiplying this by the wildtype generation time[51], then rescaling selection coefficients in relative fitness space such that a null mutant has relative fitness 0 ($s = -\infty$).

We assessed relative fitness for six genotypes (ScHsp90+E7a, ScHsp90+V23f, ScHsp90+N151a, ScHsp90+T13n, ancAmoHsp90, and ancAmoHsp90+i378L) both by monoculture and by bulk competition; these two measures are well correlated (Pearson R^2 = 0.95), although the magnitude of a fitness defect is smaller when measured by monoculture growth assays (Extended Data Fig. 3f). Individual mutants of ancAmoHsp90 and ScHsp90 were generated in the p414ADHΔTer background by Quikchange site-directed mutagenesis (Supplementary Data Table 2), confirmed by Sanger sequencing. Mutations that were generated

and assayed in ancAmoHsp90 (with number of replicate measurements in parentheses) include: S49A (n=1), T137I (n=1), V147I (n=1), I158V (n=1), R160L (n=1), G164N (n=1), E165P (n=1), L167I (n=1), K172I (n=1), L193I (n=1), and V194I (n=1). Mutations generated and assayed in ScHsp90 include: T5S (n=3), E7A (n=4), T13N (n=3), V23F (n=2), N151A (n=3), L378I (n=2), double mutants E7A/T13N (n=3), E7A/N151A (n=3), T13N/N151A (n=3), V23F/T13N (n=1), V23F/N151A (n=1), E7A/L378I (n=1), V23F/L378I (n=1) and triple mutants E7A/T13N/N151A (n=2) and V23F/T13N/N151A (n=1).

Robustness of results to statistical uncertainty and technical variables. The conclusion that the typical ancestral state is deleterious in ScHsp90 is robust to the exclusion of 20 ancestral states that have posterior probability < 1.0 at all ancestral nodes along the trajectory ($P = 4.5 \times 10^{-14}$, Wilcoxon rank sum test with continuity correction). The mutation to one ancestral state was missed in the bulk competition: its selection coefficient was inferred separately via monoculture, and including it in the analysis still leads to the conclusion that the typical ancestral state is deleterious ($P = 7.8 \times 10^{-17}$, Wilcoxon rank sum test with continuity correction).

The conclusion that the average derived state is deleterious in ancAmoHsp90 is retained when we include only the 32 mutations for which the ancAmoHsp90 state is inferred with a posterior probability of 1.0 and the derived state is inferred with posterior probability 1.0 in at least one node along the trajectory ($P = 1.1 \times 10^{-4}$, Wilcoxon rank sum test with continuity correction). The conclusion is also robust if we include selection coefficients as determined separately via monoculture for mutations to 11 derived states that were missed in the bulk competition ($P = 5.4 \times 10^{-4}$, Wilcoxon rank sum test with continuity correction).

Our assay reduces Hsp90 expression to ~1% of the endogenous level[46], which amplifies the fitness consequences of Hsp90 defects. Based on the relationship between Hsp90 function, expression, and growth rate[46], we estimate that the average selection coefficient of -0.01 we observed associated with contingency or entrenchment corresponds to a deficit of approximately $s = -5 \times 10^{-6}$ under native-like expression levels.

Expected versus observed fitness. To identify epistasis between candidate interacting sites (e.g. Fig. 4a-d) or among the broader set of substitutions (e.g. Fig. 2), we compared the observed fitness of genotypes with multiple mutations to that expected in the absence of epistasis. In the absence of epistatic interactions, selection coefficients combine additively[52]. We therefore calculated the expected selection coefficient of a genotype as the sum of selection coefficients of its component mutations as measured independently in a reference background (ancAmoHsp90 or ScHsp90). The standard error of a predicted fitness given the sum of selection coefficients was calculated as the square root of the sum of squared standard errors of the individual selection coefficient estimates, as determined from the duplicate bulk competition measurements. Epistasis was implicated if the observed fitness of a genotype differed from that predicted from the sum of its corresponding single-mutant selection coefficients.

Estimating the fraction of deleterious mutations. We sought to determine the fraction of mutations in each dataset that are deleterious using a modeling approach that incorporates measurement error and which does not require individual mutations to be classified as deleterious, neutral, or beneficial. We used the mixtools package[53] in R to estimate mixture models of underlying Gaussian distributions that best fit the observed distributions of mutant selection coefficients in each library. First, to incorporate measurement error in the model for each library, we fit a single Gaussian distribution to the measured selection coefficients of

replicate wildtype sequences that were present in the library but represented by independent barcodes. We then required one of the Gaussian distributions in each mixture model to have a mean and standard deviation fixed to that of the wildtype measurements, with a freely estimated mixture proportion. The other Gaussian components in each mixture model had a freely fit mean, standard deviation, and mixture proportion. Mixture models were fit to all non-outlier selection coefficients, because the presence of strongly deleterious selection coefficients ($s < -0.04$), which are unambiguously deleterious, interfered with model convergence. We assessed mixture models with a variable number of mixture components (k = 2 to 6 for the ancAmoHsp90 library and 2 to 5 for the ScHsp90 library, because the 6-component model would not converge), and obtained the maximum likelihood estimate of each component's mean, standard deviation, and mixture proportion via an expectation-maximization algorithm as implemented in mixtools. We compared the models built for each k using AIC. For ScHsp90, the 3-component mixture model was favored by AIC (Extended Data Fig. 4a). For ancAmoHsp90, the 2-component and 5-component mixture models had virtually indistinguishable AIC (Extended Data Fig. 6a), but the 2-component mixture model had a visually suboptimal fit (Extended Data Fig. 6c,d) and attributed a larger proportion of mutations as belonging to a deleterious sampling distribution (0.78 versus 0.53 for the 5-component mixture model), so we selected the more conservative and visually superior 5-component mixture model.

The mixture component derived from the wildtype sampling distribution was taken to represent genotypes in the library with fitness indistinguishable from wildtype; mixture components with mean < 0 were taken to reflect deleterious variants; and mixture components with mean > 0 were taken to reflect beneficial variants. For each variant, the posterior probability of being deleterious, neutral, or beneficial was determined from the relative probability density

function for mixture components in each category at the selection coefficient measured for that mutation; for variants with $s < -0.04$ that were excluded from model inference, the posterior probability of being deleterious was 1. The total fraction of variants in the library that are deleterious (or beneficial) was determined by summing the posterior probabilities of being deleterious (or beneficial) over all mutants. To generate the representations in Figures 1c and 3a, posterior probabilities were summed separately for the set of measurements that fall within each histogram bin.

Uncertainty in the estimated fraction of mutations that are deleterious or beneficial was determined via a bootstrapping procedure. For each of 10,000 bootstrap replicates, measured selection coefficients from the bulk competition were resampled with replacement. Mixture models with fixed k were fit to each bootstrap sample, and the estimated fractions of mutations in deleterious, neutral, or beneficial sampling distributions were determined as above.

To estimate the probability that a pair of states exhibit contingency and/or entrenchment, we calculated the joint posterior probability as the product of the probabilities that each pair of sites is in the relevant selection category (ancestral state with fitness greater than, less than, or indistinguishable from the derived state) in the ScHsp90 and the ancAmoHsp90 backgrounds. For sites that substituted from the ancAmoHsp90 state i to the ScHsp90 state j ($i \rightarrow j$, n = 35), i is the ancestral state and j the derived state for measurements in both backgrounds. For sites that substituted from the ancAmoHsp90 state i to an intermediate state j before substituting back to i in ScHsp90 ($i \rightarrow j \rightarrow i$, n=12), then i is the ancestral state and j derived in ancAmoHsp90 assay, and j is the ancestral state and i derived in ScHsp90. For sites that substituted from the ancAmoHsp90 state i to an intermediate state j that was further modified to k in ScHsp90 ($i \rightarrow j \rightarrow k$, n=25), two comparison were made: in the first, i was ancestral and k was derived for

measurements in both backgrounds, while in the second comparison, *i* was ancestral and *j* derived in ancAmoHsp90, and *j* ancestral and *k* derived in ScHsp90.

Data and code availability. Processed sequencing data and scripts to reproduce all analyses are available at github.com/JoeThorntonLab/Hsp90_contingency-entrenchment. All other data are available from the corresponding author upon reasonable request.

Acknowledgements We thank members of the Thornton lab for comments on the manuscript and Jeff Boucher for technical assistance. This work was supported by National Institutes of Health R01GM104397 and R01GM121931 (J.W.T.), R01GM112844 (D.N.A.B.), F32GM119205-2 (J.M.F.), T32-GM007183 (T.N.S.), and a National Science Foundation Graduate Research Fellowship (T.N.S.)

Author Contributions All authors conceived the project and designed experiments. T.N.S., J.M.F., and P.M. performed experiments and analyzed data. T.N.S. and J.W.T. wrote the paper, with contributions from all authors.

Author Information The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.W.T. (joet1@uchicago.edu).

References

1. Olson, C. A., Wu, N. C. & Sun, R. A comprehensive biophysical description of pairwise epistasis throughout an entire protein domain. *Curr Biol* **24,** 2643–2651 (2014).
2. Bank, C., Hietpas, R. T., Jensen, J. D. & Bolon, D. N. A. A systematic survey of an intragenic epistatic landscape. *Mol Biol Evol* **32,** 229–238 (2015).
3. Podgornaia, A. I. & Laub, M. T. Pervasive degeneracy and epistasis in a protein-protein interface. *Science* **347,** 673–677 (2015).
4. Sarkisyan, K. S. *et al.* Local fitness landscape of the green fluorescent protein. *Nature* **533,** 397–401 (2016).
5. Pollock, D. D., Thiltgen, G. & Goldstein, R. A. Amino acid coevolution induces an evolutionary Stokes shift. *P Natl Acad Sci USA* **109,** E1352–9 (2012).
6. Shah, P., McCandlish, D. M. & Plotkin, J. B. Contingency and entrenchment in protein evolution under purifying selection. *P Natl Acad Sci USA* **112,** E3226–E3235 (2015).
7. Ashenberg, O., Gong, L. I. & Bloom, J. D. Mutational effects on stability are largely conserved during protein evolution. *P Natl Acad Sci USA* **110,** 21071–21076 (2013).
8. Risso, V. A. *et al.* Mutational studies on resurrected ancestral proteins reveal conservation of site-specific amino acid preferences throughout evolutionary history. *Mol Biol Evol* **32,** 440–455 (2015).
9. Breen, M. S., Kemena, C., Vlasov, P. K., Notredame, C. & Kondrashov, F. A. Epistasis as the primary factor in molecular evolution. *Nature* **490,** 535–538 (2012).
10. McCandlish, D. M., Rajon, E., Shah, P., Ding, Y. & Plotkin, J. B. The role of epistasis in protein evolution. *Nature* **497,** E1–2 (2013).
11. Goldstein, R. A., Pollard, S. T., Shah, S. D. & Pollock, D. D. Non-adaptive amino acid convergence rates decrease over time. *Mol Biol Evol* (2015). doi:10.1093/molbev/msv041
12. Povolotskaya, I. S. & Kondrashov, F. A. Sequence space and the ongoing expansion of the protein universe. *Nature* **465,** 922–926 (2010).
13. Mendes, F. K., Hahn, Y. & Hahn, M. W. Gene Tree Discordance Can Generate Patterns of Diminishing Convergence over Time. *Mol Biol Evol* **33,** 3299–3307 (2016).
14. Lunzer, M., Golding, G. B. & Dean, A. M. Pervasive cryptic epistasis in molecular evolution. *PLOS Genet* **6,** e1001162 (2010).
15. Gong, L. I., Suchard, M. A. & Bloom, J. D. Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife* **2,** e00631 (2013).
16. Doud, M. B., Ashenberg, O. & Bloom, J. D. Site-Specific Amino Acid Preferences Are Mostly Conserved in Two Closely Related Protein Homologs. *Mol Biol Evol* **32,** 2944–2960 (2015).
17. Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci* **25,** 1204–1218 (2016).
18. Ortlund, E. A., Bridgham, J. T., Redinbo, M. R. & Thornton, J. W. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* **317,** 1544–1548 (2007).
19. Bloom, J. D., Gong, L. I. & Baltimore, D. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* **328,** 1272–1275 (2010).
20. Natarajan, C. *et al.* Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* **354,** 336–339 (2016).
21. McKeown, A. N. *et al.* Evolution of DNA specificity in a transcription factor family produced a new gene regulatory module. *Cell* **159,** 58–68 (2014).

22. Bridgham, J. T., Ortlund, E. A. & Thornton, J. W. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* **461,** 515–519 (2009).
23. Soylemez, O. & Kondrashov, F. A. Estimating the rate of irreversibility in protein evolution. *Genome Biol Evol* **4,** 1213–1222 (2012).
24. Jordan, D. M. *et al.* Identification of cis-suppression of human disease mutations by comparative genomics. *Nature* **524,** 225–229 (2015).
25. Hochberg, G. K. A. & Thornton, J. W. Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. *Annu Rev Biophys* **46,** 247–269 (2017).
26. Piper, P. W. *et al.* Yeast is selectively hypersensitised to heat shock protein 90 (Hsp90)-targetting drugs with heterologous expression of the human Hsp90β, a property that can be exploited in screens for new Hsp90 chaperone inhibitors. *Gene* **302,** 165–170 (2003).
27. Wider, D., Péli-Gulli, M.-P., Briand, P.-A., Tatu, U. & Picard, D. The complementation of yeast with human or Plasmodium falciparum Hsp90 confers differential inhibitor sensitivities. *Mol Biochem Parasit* **164,** 147–152 (2009).
28. Adl, S. M. *et al.* The Revised Classification of Eukaryotes. *J Eukaryot Microbiol* **59,** 429–514 (2012).
29. Tokuriki, N. & Tawfik, D. S. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol* **19,** 596–604 (2009).
30. Sailer, Z. R. & Harms, M. J. Detecting High-Order Epistasis in Nonlinear Genotype-Phenotype Maps. *Genetics* (2017). doi:10.1534/genetics.116.195214/-/DC1
31. Pantzartzi, C. N., Drosopoulou, E. & Scouras, Z. G. Assessment and Reconstruction of Novel HSP90 Genes: Duplications, Gains and Losses in Fungal and Animal Lineages. *PLOS One* **8,** e73217 (2013).
32. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22,** 1658–1659 (2006).
33. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7,** 1–6 (2011).
34. Le, S. Q. & Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Mol Biol Evol* **25,** 1307–1320 (2008).
35. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).
36. Brown, M. W. *et al.* Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. *P Roy Soc B: Biol Sci* **280,** 20131755–20131755 (2013).
37. Brown, M. W., Spiegel, F. W. & Silberman, J. D. Phylogeny of the 'Forgotten' Cellular Slime Mold, Fonticula alba, Reveals a Key Evolutionary Branch within Opisthokonta. *Mol Biol Evol* **26,** 2699–2709 (2009).
38. Paps, J., Medina-Chacón, L. A., Marshall, W., Suga, H. & Ruiz-Trillo, I. Molecular Phylogeny of Unikonts: New Insights into the Position of Apusomonads and Ancyromonads and the Internal Relationships of Opisthokonts. *Protist* **164,** 2–12 (2013).
39. Kurtzman, C. P. & Robnett, C. J. Relationships among genera of the Saccharomycotina(Ascomycota) from multigene phylogenetic analysis of type species. *FEMS Yeast Res* **13,** 23–33 (2013).
40. Shen, X.-X. *et al.* Reconstructing the Backbone of the Saccharomycotina Yeast Phylogeny Using Genome-Scale Data. *G3* **6,** 3927–3939 (2016).
41. Yang, Z., Kumar, S. & Nei, M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141,** 1641–1650 (1995).

42. Eme, L., Sharpe, S. C., Brown, M. W. & Roger, A. J. On the Age of Eukaryotes: Evaluating Evidence from Fossils and Molecular Clocks. *CSH Perspect Biol* **6,** a016139–a016139 (2014).
43. Taylor, J. W. & Berbee, M. L. Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia* **98,** 838–849 (2006).
44. Ali, M. M. U. *et al.* Crystal structure of an Hsp90–nucleotide–p23/Sba1 closed chaperone complex. *Nature* **440,** 1013–1017 (2006).
45. Cunningham, C. N., Southworth, D. R., Krukenberg, K. A. & Agard, D. A. The conserved arginine 380 of Hsp90 is not a catalytic residue, but stabilizes the closed conformation required for ATP hydrolysis. *Protein Sci* **21,** 1162–1171 (2012).
46. Jiang, L., Mishra, P., Hietpas, R. T., Zeldovich, K. B. & Bolon, D. N. A. Latent Effects of Hsp90 Mutants Revealed at Reduced Expression Levels. *PLOS Genet* **9,** e1003600 (2013).
47. Hiatt, J. B., Patwardhan, R. P., Turner, E. H., Lee, C. & Shendure, J. Parallel, tag-directed assembly of locally derived short sequence reads. *Nat Methods* **7,** 119–122 (2010).
48. Hietpas, R. T., Bank, C., Jensen, J. D. & Bolon, D. N. A. Shifting fitness landscapes in response to altered environments. *Evolution* **67,** 3512–3522 (2013).
49. Mishra, P., Flynn, J. M., Starr, T. N. & Bolon, D. N. A. Systematic Mutant Analyses Elucidate General and Client-Specific Aspects of Hsp90 Function. *Cell Rep* **15,** 588–598 (2016).
50. Hietpas, R., Roscoe, B., Jiang, L. & Bolon, D. N. A. Fitness analyses of all possible point mutations for regions of genes in yeast. *Nat Protoc* **7,** 1382–1396 (2012).
51. Chevin, L. M. On measuring selection in experimental evolution. *Biol Letters* **7,** 210–213 (2011).
52. Sella, G. & Hirsh, A. E. The application of statistical physics to evolutionary biology. *P Natl Acad Sci USA* **102,** 9541–9546 (2005).
53. Benaglia, T., Chauveau, D. & Hunter, D. mixtools: An R package for analyzing finite mixture models. *J Stat Softw* **32,** (2009).

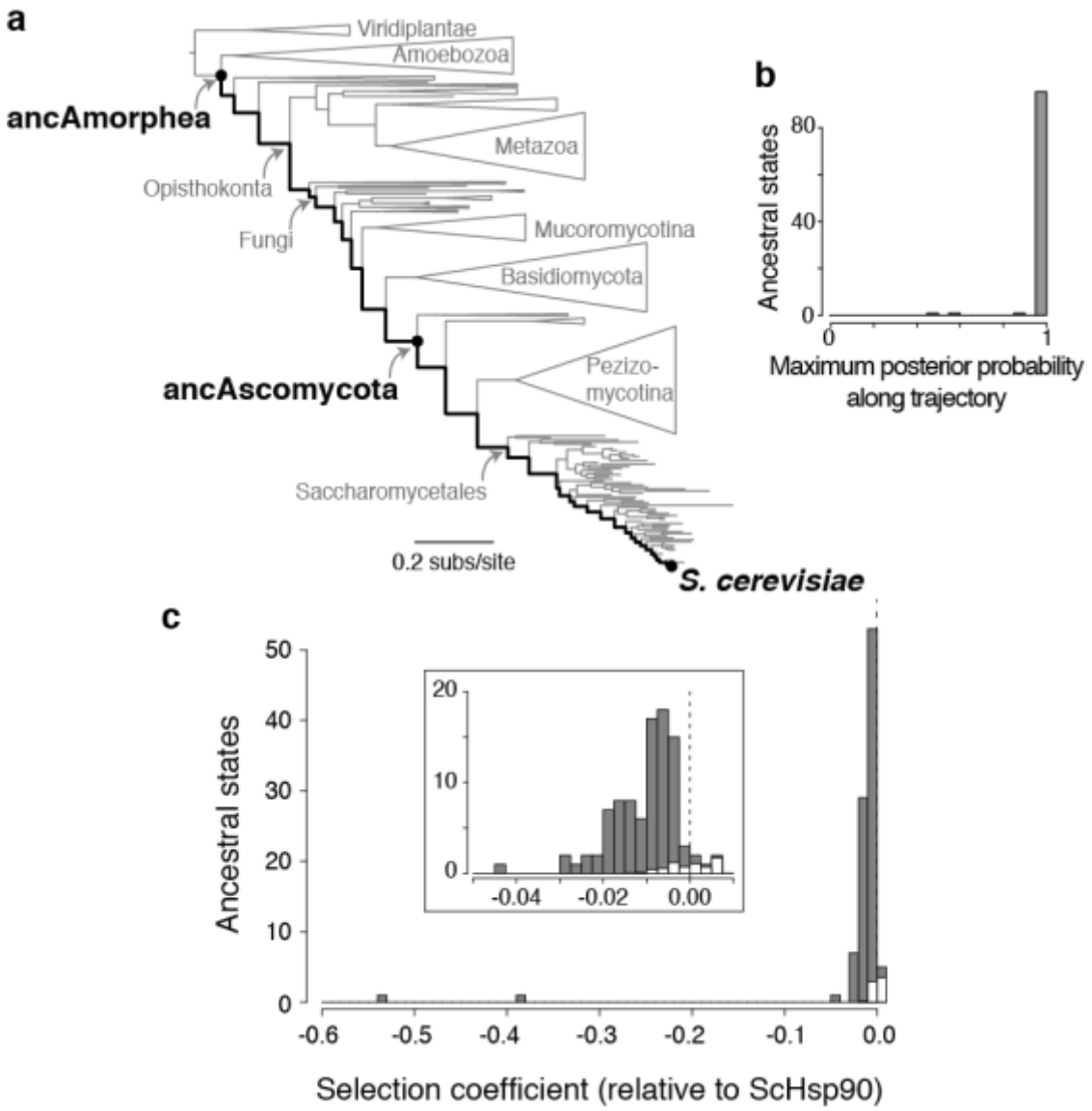


Figure 1 | Ancestral states are deleterious in yeast Hsp90. a, Maximum likelihood phylogeny of Hsp90 protein sequences from Amorphea. The evolutionary trajectory studied, from the last common ancestor of Amorphea to modern *S. cerevisiae*, is indicated by a dark black line. Major taxonomic groups are labeled in gray. Ancestral and extant genotypes characterized in this study are in black. Complete phylogeny with taxon names is in Extended Data Fig. 1. **b**, Statistical confidence in ancestral amino acid states. For each of the 98 inferred ancestral states, the highest posterior probability of the state at any internal node along the trajectory is shown. **c**, Distribution of selection coefficients of individual ancestral states when introduced into ScHsp90, measured as the log relative fitness compared to ScHsp90 in a deep-sequencing based bulk competition assay. Dashed line indicates neutrality. Inset, close view of the region near s = 0. In each histogram bin, white and grey show the proportion of ancestral states with selection coefficients in that range that are estimated to be neutral or deleterious, respectively, when measurement error is taken into account. Median of the distribution is significantly less than zero ($P = 1.2 \times 10^{-16}$, Wilcoxon rank sum test with continuity correction), a conclusion robust to excluding ancestral states that are reconstructed with any statistical ambiguity ($P = 4.5 \times 10^{-14}$).

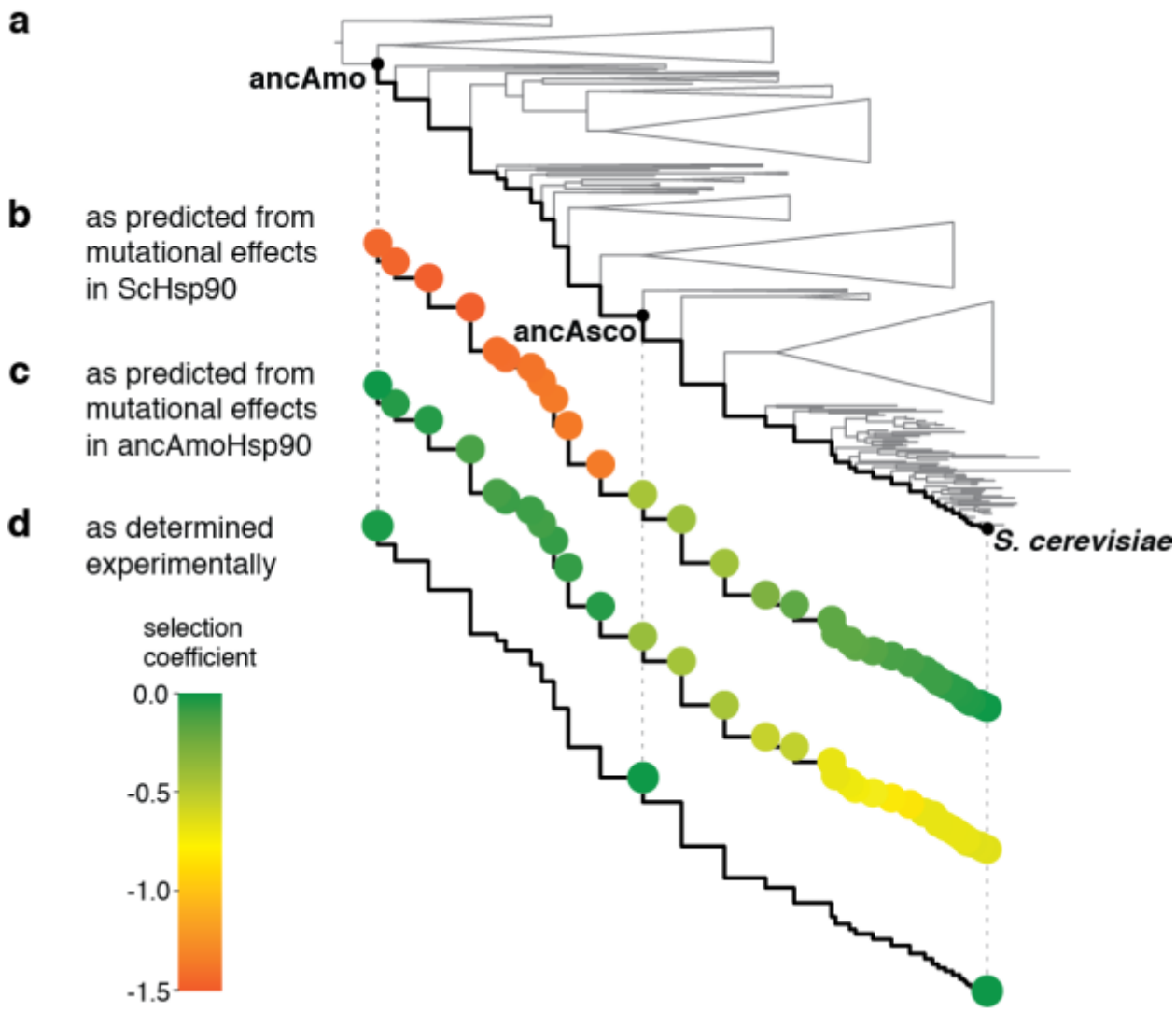


Figure 2 | Fitness effects of historical substitutions are modified by intramolecular epistasis. For each node along the trajectory from ancAmoHsp90 to ScHsp90 (black line), the predicted or actual selection coefficient of the entire genotype is represented from green (s = 0) to orange (s = -1.5). **a**, The Hsp90 phylogeny, represented as in Fig. 1a. **b**, The predicted selection coefficient of each ancestral sequence relative to ScHsp90 was calculated as the sum of the selection coefficients of each ancestral state present in that ancestor when measured individually in ScHsp90. **c**, The predicted selection coefficient of each sequence relative to ancAmoHsp90 was calculated as the sum of the selection coefficients of each derived state present at that node when measured individually in ancAmoHsp90. **d**, The experimentally determined selection coefficients for ancAmoHsp90 and ancAscoHsp90 relative to ScHsp90. For selection coefficients of each genotype, see Extended Data Fig. 6.

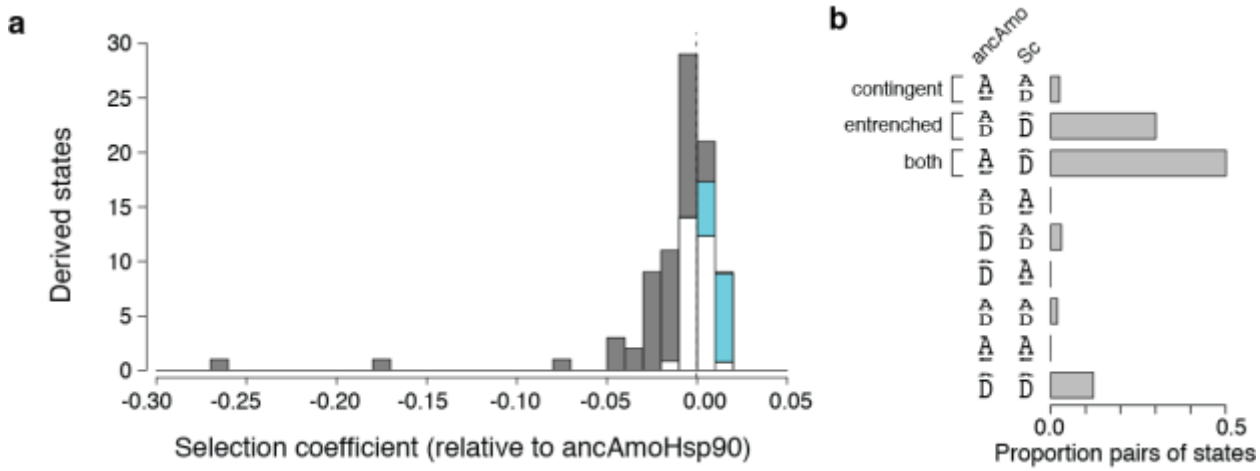


Figure 3 | Widespread contingency and entrenchment. a, Distribution of measured selection coefficients of derived states when introduced singly into ancAmoHsp90. Dashed line indicates neutrality. In each histogram bin, white shows the proportion of derived states with selection coefficients in that range that are estimated to be neutral; gray, deleterious; blue, beneficial. The median of the distribution is significantly less than zero ($P = 5.8 \times 10^{-4}$, Wilcoxon rank sum test with continuity correction). **b**, The fraction of pairs of ancestral and derived states that are inferred to be contingent, entrenched or both. Pairs of ancestral and derived states at each site can be classified by the relative fitness of the two states when measured in ancAmoHsp90 or in ScHsp90: ancestral state more fit (A larger than D), derived state more fit (D larger than A), or fitnesses indistinguishable (A and D same size). The fraction of pairs in each category was estimated from the joint distribution of selection coefficients in both backgrounds.

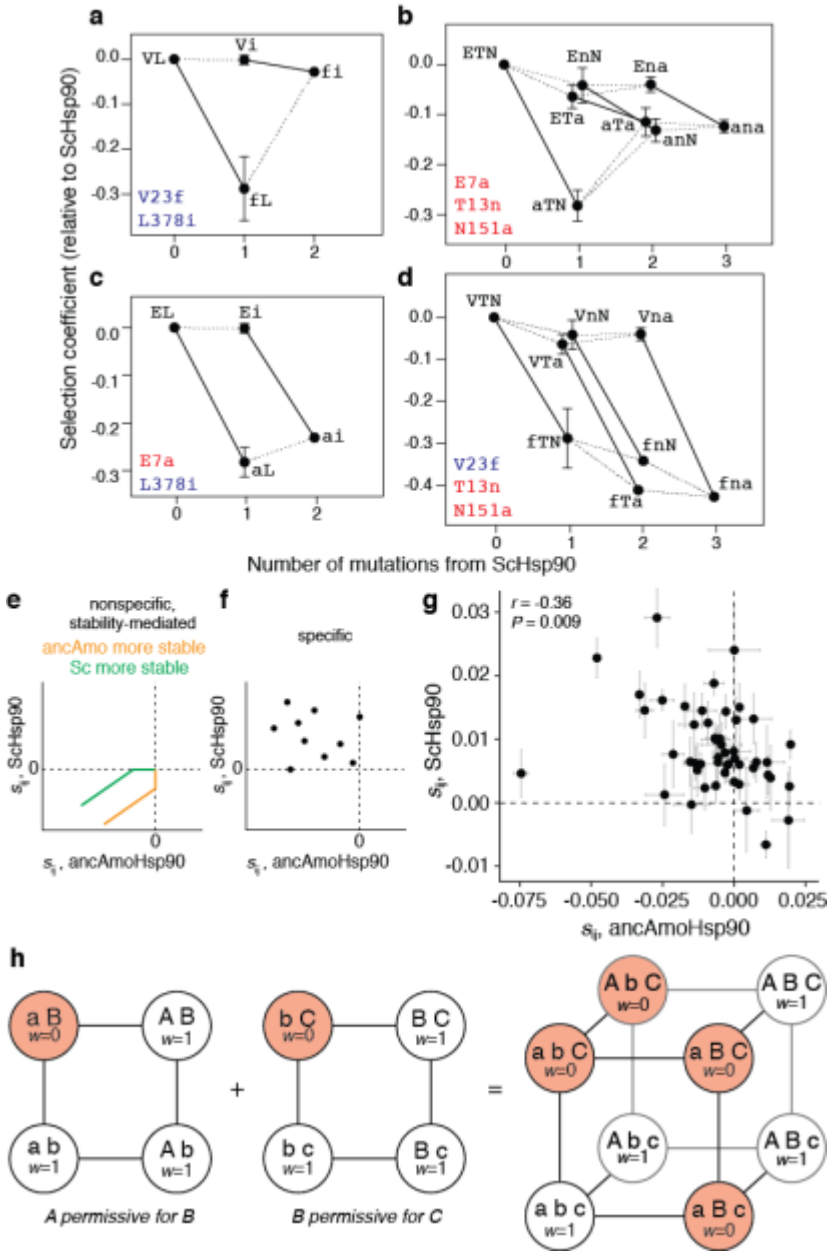
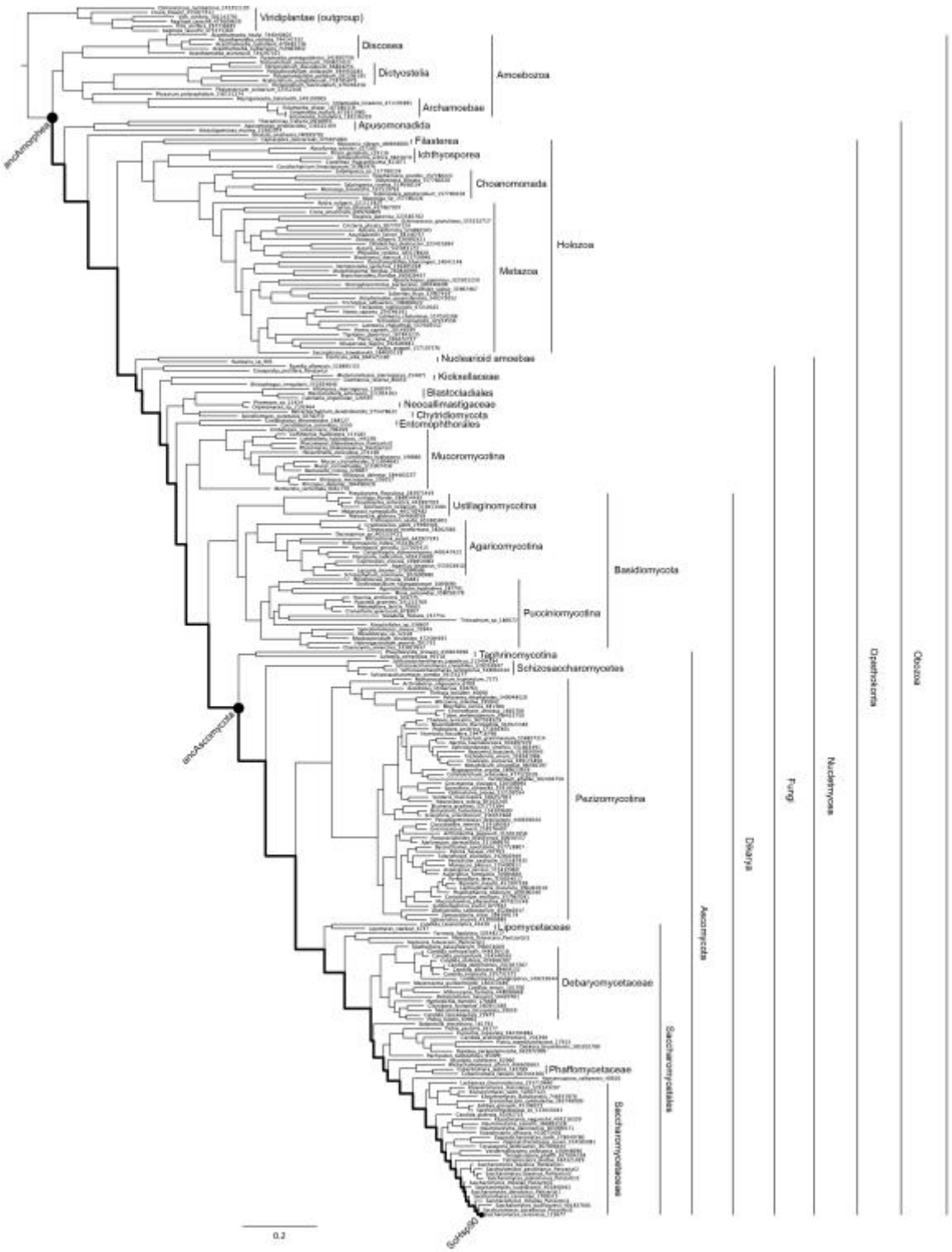
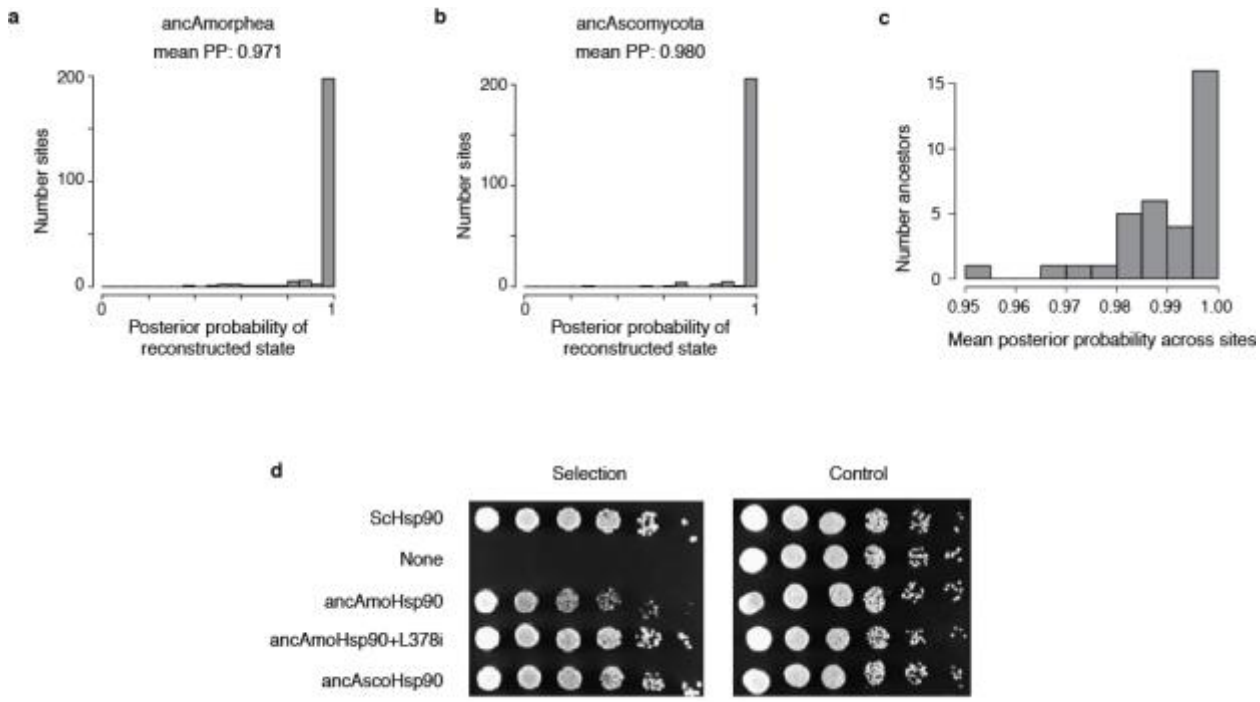


Figure 4 | Epistatic interactions are specific. a-d, Large-effect deleterious reversions and restrictive substitutions that contributed to their irreversibility. For each single, double, or triple mutant in ScHsp90, the selection coefficient relative to ScHsp90 is shown, as assessed in monoculture growth assays. Lines connect genotypes that differ by a single mutation; solid lines indicate the effect of the large-effect reversions in each background. Error bars, SEM for 2 to 4 replicates (see Methods). Data points are labeled by amino acid states: lower-case, ancestral state; upper-case, derived state. **a**, Deleterious reversion V23f is ameliorated by L378i. **b**, Deleterious reversion E7a is partially ameliorated by N151a or T13n. **c**, L378i does not ameliorate E7a. **d**, N151a and T13n do not ameliorate V23f. **e-f**, Expected relationship under two models of epistasis between selection coefficients of ancestral-to-derived mutations (s_{ij}) when introduced into ancestral (*x*-axis) or derived (*y*-axis) backgrounds. **e**, Nonspecific epistasis: if genetic interactions are the nonspecific result of a threshold-like, buffering relationship between stability (or another bulk property) and fitness[15,29], then the effects of strongly deleterious

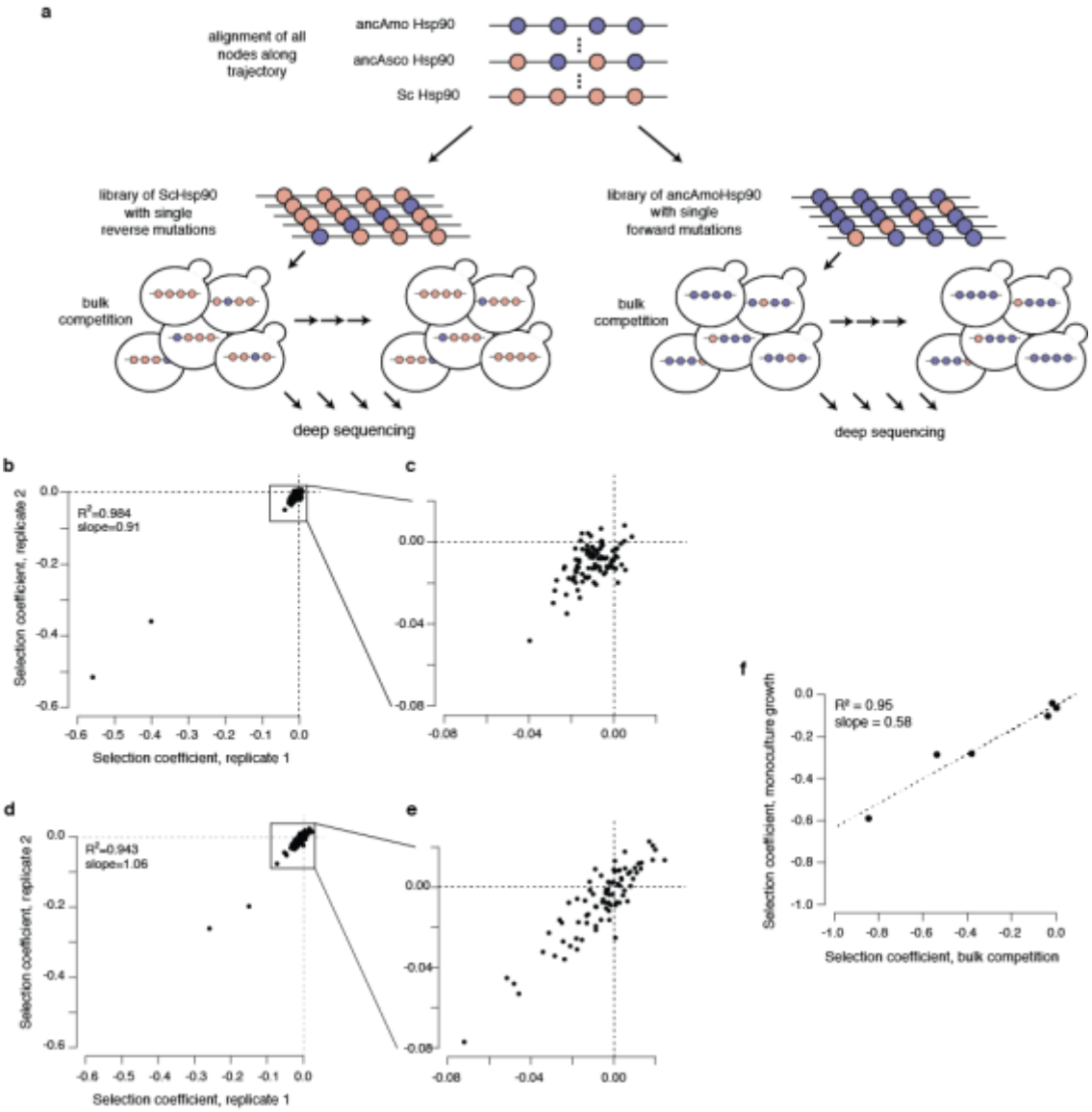
mutations will be positively correlated between the two backgrounds, but weakly deleterious mutations in the less stable background may be neutral in the more stable background (yellow, ancAmoHsp90 more stable; green, ScHsp90 more stable). **f**, Specific epistasis: if interactions reflect specific couplings between sites, then mutations from ancestral to derived states can be deleterious in the ancestral background but beneficial in the derived background (upper left quadrant). **g**, Measured selection coefficients for ancestral-derived state pairs that differ between ancAmoHsp90 and ScHsp90. Dashed lines, $s = 0$. Error bars, SEM from two replicate bulk competition measurements. r, Pearson correlation coefficient and associated P value. Two additional points that are strongly deleterious outliers in the ScHsp90 or ancAmoHsp90 data are not shown for clarity and are not included in the correlation; both fall in the upper-left quadrant. **h**, Daisy-chain model of specific epistatic interactions. Each square shows the mutant cycle for a pair of substitutions (A and B or B and C; lower-case, ancestral state; upper-case, derived), one of which is permissive for the other. Each circle is a genotype colored by its fitness (w): white, neutral; orange, deleterious. Edges are single-site amino acid changes. The cube shows the combined mutant cycle for all three substitutions. Permissive substitutions become entrenched when the mutation that was contingent upon it occurs. Substitutions in the middle of the daisy-chain, which require a permissive mutation and are permissive for a subsequent mutation, are both contingent and entrenched.



Extended Data Figure 1 | Hsp90 phylogeny. The maximum likelihood phylogeny of 267 Hsp90 protein sequences, with major taxonomic groups labeled. Taxon names indicate genus, species, and an accession number or sequence identifier; complete sequence identification information is given in Supplementary Data Table 1. Nodes characterized in this study are shown as black dots; the trajectory studied is shown as a thick black line.



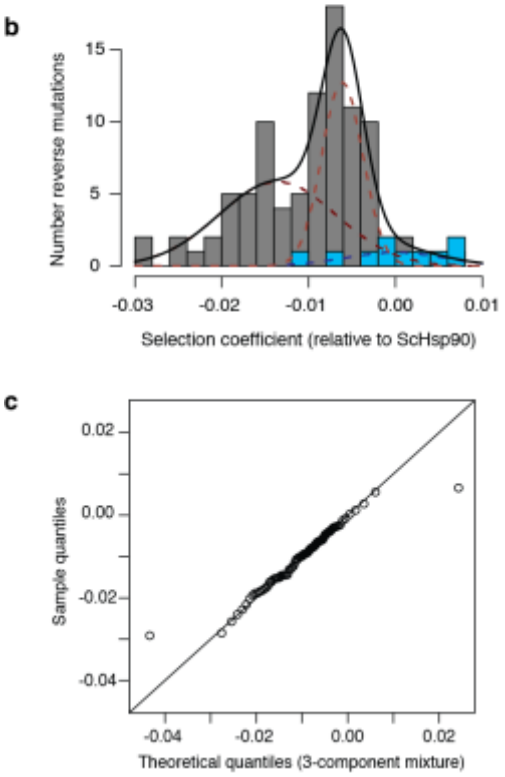
Extended Data Figure 2 | Ancestral Hsp90 sequences have high statistical support and complement yeast growth. a,**b** For the ancestral sequences reconstructed in this study, the distribution of posterior probability of ancestral states across NTD sites is shown as a histogram. The mean posterior probability of the most probable state across sites (mean PP) is shown for each ancestor. **c**, The distribution of mean PP for reconstructed ancestral sequences along the trajectory from ancAmoHsp90 to ScHsp90. **d**, Growth of *S. cerevisiae* Hsp90 shutoff strains complemented with ancestral Hsp90 variants. Spots from left to right are 5-fold serial dilutions. Control plates represent conditions in which the native ScHsp90 allele is expressed. Under selection conditions, the native ScHsp90 allele is turned off, and growth can only persist when a complementary Hsp90 allele is provided. The ancAmoHsp90 NTD expressed as a chimera with the Sc middle and C-terminal domains exhibits a slight growth defect; this is rescued by adding an additional reversion to the ancAmoHsp90 state in the middle domain (L378i), which occurs on a middle domain loop that extends down and interacts directly with the N-terminal domain and contributes to the NTD ATP-binding pocket. We subsequently refer to ancAmoHsp90+L378i as ancAmoHsp90.



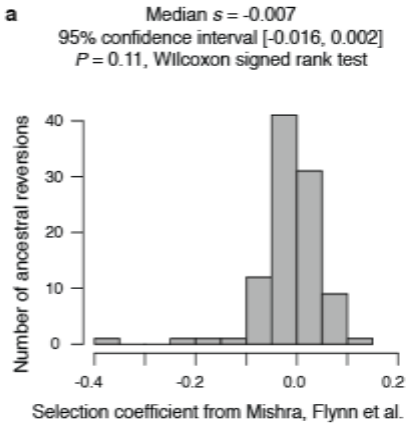
Extended Data Figure 3 | Experimental scheme and reproducibility. a, Experimental scheme for testing the fitness effects of individual mutations to ancestral states in ScHsp90 (left) or individual mutations to derived states in ancAmoHsp90 (right). An alignment of all ancestors along the focal trajectory was constructed to identify the trajectory of Hsp90 sequence change from ancAmoHsp90 to ScHsp90. In each background, a library was constructed consisting of the wildtype sequence and all individual mutations to ancestral or derived states. This library was transformed into yeast, which grew through a bulk competition. The frequency of each genotype at each time point was determined by deep sequencing, allowing us to calculate a selection coefficient for each mutation relative to the respective wildtype sequence. **b**, Reproducibility in selection coefficient estimates for replicate bulk competitions of the ScHsp90 library. R^2, Pearson coefficient of determination. **c**, For visual clarity, zoomed in representation of the boxed region in (**b**). **d**, Reproducibility in selection coefficient estimates for replicate bulk competitions of the ancAmoHsp90 library. R^2, Pearson coefficient of determination. **e**, For visual clarity, zoomed in representation of the boxed region in (**d**). **f**, Correlation in fitness as measured via bulk competition or monoculture growth assay. R^2, Pearson coefficient of determination.

a

Number components	log-likelihood	Number parameters	AIC
2	334.4	3	-662.7
3	341.0	6	-669.9
4	341.8	9	-665.6
5	343.2	12	-662.3



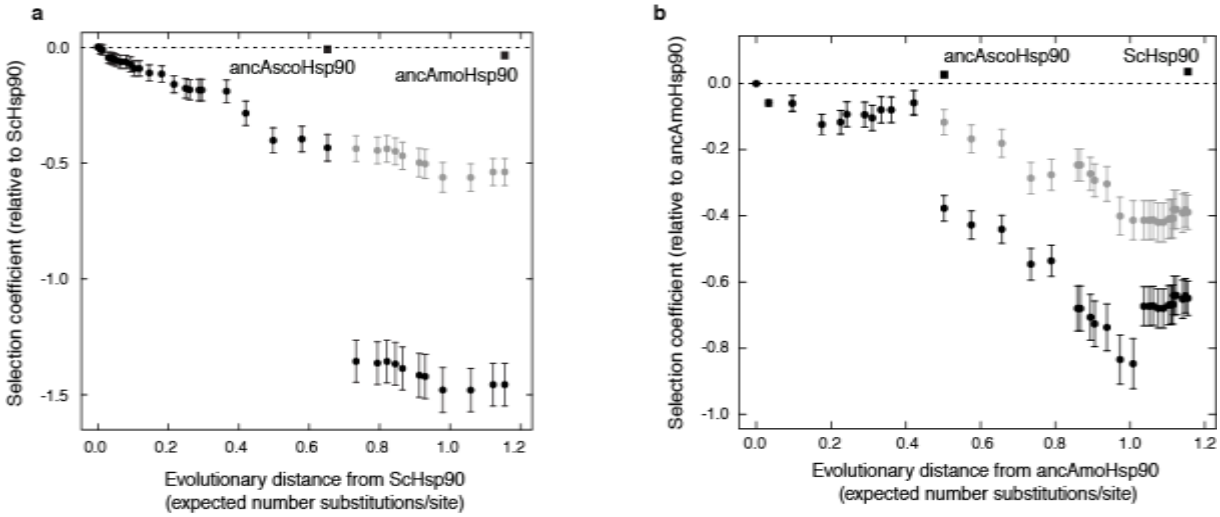
Extended Data Figure 4 | Estimating the proportion of mutations to ancestral states that are deleterious with a mixture model. a, Observed selection coefficients of reversions were fit to mixture models containing a variable number of Gaussian distributions; in each case, one distribution is fixed to have the mean and standard deviation of the sampling distribution of independent wildtype ScHsp90 sequences present in the library, the mixture proportion of which is a free parameter; each additional mixture component has a free mean, standard deviation, and mixture proportion. The empirical data were best fit by a 3-component mixture model, as assessed by AIC. **b**, The best-fit mixture model. Gray bars, observed distribution of selection coefficients of ancestral reversions; blue bars, distribution of observed selection coefficients of wildtype ScHsp90 sequences present in the library. Black line, best-fit mixture model; red dashed lines, individual mixture components centered below zero; blue dashed line, wildtype mixture component. The area under the curve for each mixture component corresponds to the proportion it contributes to the overall mixture model. **c**, Quantile-quantile plot showing the quality of fit of the 3-component mixture model (*x*-axis) to the empirical distribution of selection coefficients of ancestral reversions (*y*-axis). The mixture model assigns more extreme selection coefficients to the tails than is observed in the empirical distribution, but provides a reasonable fit along the bulk of the distribution.



b

	Spearman's rank correlation, $s > -0.1$		Spearman's rank correlation, all mutations	
Properties of sites	rho	*P*-value	rho	*P*-value
Average effect of all mutations at position (ref 49)	0.267	0.009	0.308	0.002
Relative evolutionary rate	0.294	0.004	0.326	0.001
Relative solvent accessibility (PDB 2CG9, ref 44)	0.331	0.002	0.313	0.003
Distance to ATP gamma-phosphate	0.380	0.0003	0.395	0.0001
Properties of mutations				
BLOSUM62 score	0.048	0.65	0.077	0.45
Difference in hydrophobicity	-0.209	0.04	-0.178	0.08
Absolute difference in hydrophobicity	-0.121	0.24	-0.136	0.18
Difference in side chain volume	0.134	0.19	0.129	0.21
Absolute difference in side chain volume	-0.042	0.68	-0.072	0.48

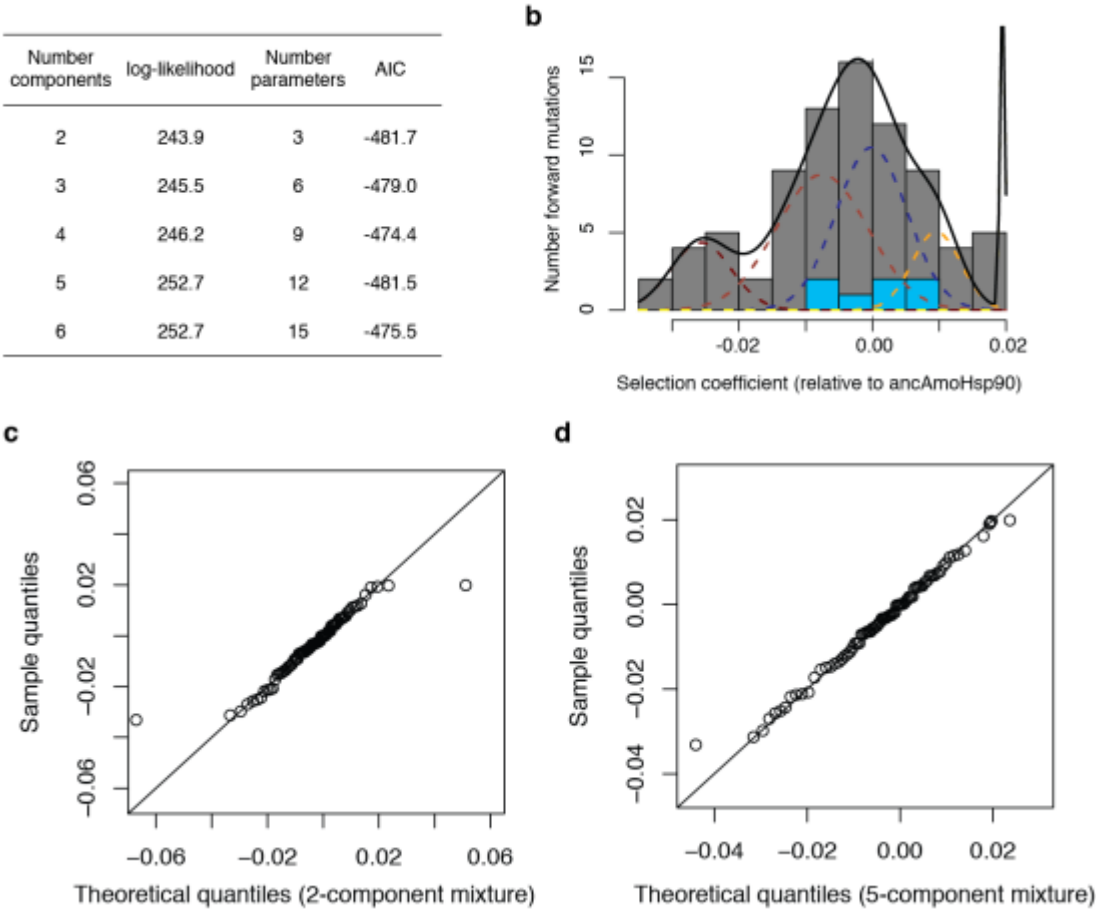
Extended Data Figure 5 | Ancestral states are deleterious in yeast Hsp90. a, The signature of deleterious ancestral states is also present in the independent but lower-resolution dataset of Mishra, Flynn et al.[49]. For each mutation to an ancestral state, the selection coefficient as determined by Mishra, Flynn et al. is shown. The median selection coefficient is -0.007, close to that estimated in the current study; however, this median selection coefficient is not significantly different than zero (P = 0.11). Because Mishra, Flynn et al. tested a much larger panel of mutations (all single mutations across the entire NTD), experimental variability of estimated selection coefficients was much larger, possibly explaining the lack of significance of this result in this dataset. **b**, Deleterious reversions exhibit properties typical of genuinely deleterious mutations. For various properties of sites at which we measured the fitness of ancestral variants (top) or properties of the specific amino acids mutated (bottom), we asked whether there was a significant correlation between the property and the selection coefficients of mutations via Spearman's rank correlation. Ancestral states tend to be more deleterious at positions that are less robust to any mutation, evolve more slowly, are less solvent accessible, and are closer to the gamma-phosphate of bound ATP. These properties are not completely independent; for example, there is a significant positive correlation between relative solvent accessibility and distance to ATP gamma-phosphate.



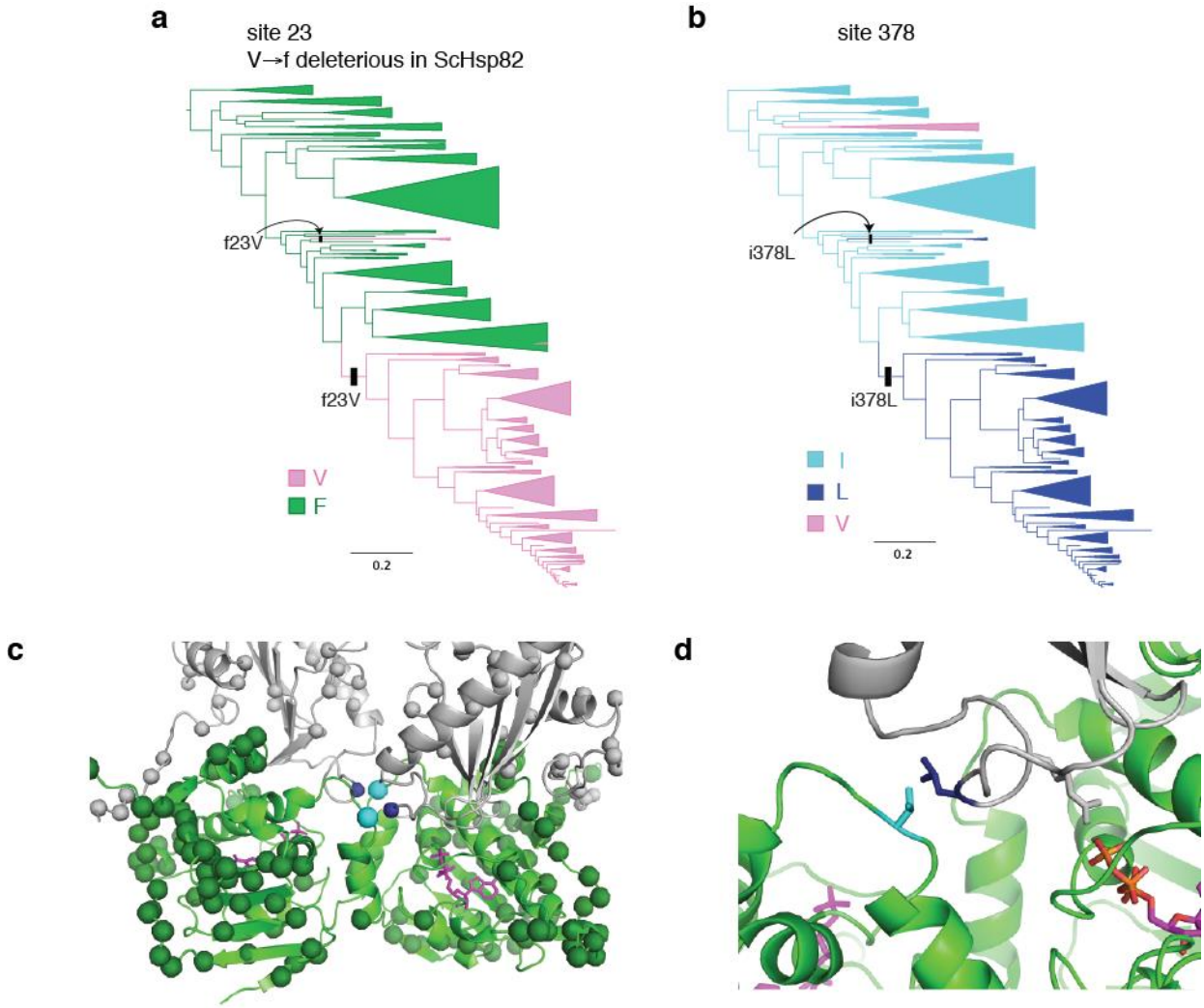
Extended Data Figure 6 | Fitness effects of historical substitutions are modified by intramolecular epistasis. Each black circle represents an ancestral protein along the trajectory from ancAmoHsp90 to ScHsp90. Position along the *x*-axis shows the evolutionary distance that separates it from ScHsp90 (**a**) or ancAmoHsp90 (**b**); *y*-axis position shows the predicted selection coefficient assuming no epistasis relative to ScHsp90 (**a**) or ancAmoHsp90 (**b**). Predicted selection coefficients were calculated as the sum of individual selection coefficients for all sequence differences present in its sequence as measured in ScHsp90 (**a**) or ancAmoHsp90 (**b**). Error bars show the standard error of the predicted value, calculated by propagating the standard errors of individual site-specific selection coefficient measurements. Light gray dots show the same data, but excluding the effects of the two strongly deleterious outliers in each library. Labeled squares indicate experimentally determined selection coefficients for complete genotypes: ancAscoHsp90, ancestral Ascomycota (fitness determined via monoculture growth); ancAmoHsp90, ancestral Amorphea (fitness determined via bulk competition). Dashed line, s = 0.

a

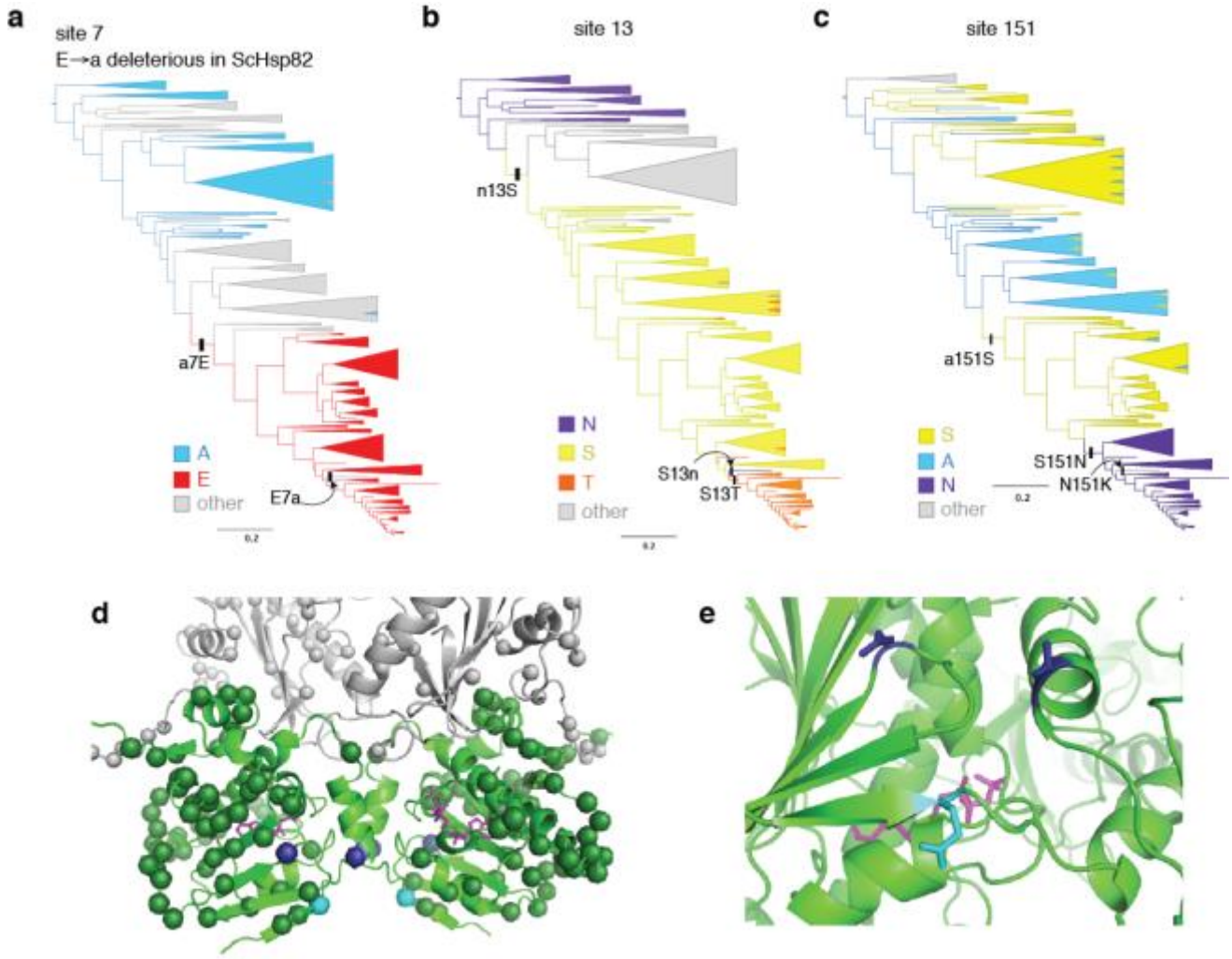
Number components	log-likelihood	Number parameters	AIC
2	243.9	3	-481.7
3	245.5	6	-479.0
4	246.2	9	-474.4
5	252.7	12	-481.5
6	252.7	15	-475.5



Extended Data Figure 7 | Estimating the proportion of mutations to derived states that are deleterious with a mixture model. a, The distribution of selection coefficients of mutations to derived states was fit by mixture models containing a variable number of Gaussian distributions; in each case, one distribution is fixed to have the mean and standard deviation of the sampling distribution of independent wildtype ancAmoHsp90 alleles in the library, the mixture proportion of which is a free parameter; each additional mixture component has a free mean, standard deviation, and mixture proportion. The empirical data were best fit by a 2-component mixture model, as judged by AIC, with a 5-component mixture being almost equally well fit; the 5-component mixture resulted in a more conservative estimate of the proportion of mutations that were deleterious than the 2-component mixture, and so was chosen despite the AIC difference of 0.2. **b**, The fit of the 5-component mixture model. Gray bars, distribution of selection coefficients of mutations to derived states; blue bars, distribution of selection coefficients of independent ancAmoHsp90 alleles present in the library. Black line, five-component mixture model. Red dashed lines, individual mixture components centered below zero; blue dashed line, wildtype mixture component; yellow dashed lines, individual mixture components centered above zero; relative integrated areas of mixture components correspond to the relative proportions they contribute to the overall mixture model. **c**,**d**, Quantile-quantile plot showing the quality of fit of the 2-component (**c**), or 5-component (**d**), mixture models (*x*-axis) to the empirical distribution of selection coefficients of mutations to derived states (*y*-axis).



Extended Data Figure 8 | The deleterious V23f reversion is ameliorated by L378i. a,**b** Character state patterns at sites 23 (**a**) and 378 (**b**). On the lineage to ScHsp90, f23V co-occurred with i378L before the common ancestor of Ascomycota. The same two substitutions also co-occur on an independent lineage on this phylogeny (Kickxellaceae fungi), and in the distantly related Rhodophyta red algae (not shown). **c**, The locations of sites 23 and 378 on the ATP-bound Hsp90 dimer structure (PDB 2CG9). Cyan spheres, site 23; dark blue; site 378; dark green, other variable NTD sites; gray, other variable middle and C-terminal domain sites. Magenta sticks, ATP. **d**, Zoomed view of sites 23 and 378. These side chains are in direct structural contact, and may be important for the positioning of the middle domain loop that bears R380 (gray sticks), which forms a salt bridge with the ATP gamma-phosphate and is critical for ATP binding and hydrolysis[44,45].



Extended Data Figure 9 | The deleterious E7a reversion is partially ameliorated by N151a or T13n. a,**b**,**c**, Character state patterns at sites 7 (**a**), 13 (**b**) and 151 (**c**). On the trajectory to ScHsp90, a7E occurred before the common ancestor of Ascomycota, then later reverted in the lineage leading to *Ascoidea rubescens* (arrow)*;* on this latter lineage, site 13 also reverted to the ancestral state asparagine, and site 151 substituted to a third state lysine. **d**, The locations of sites 7, 13, and 151 on the ATP-bound Hsp90 structure (2CG9), represented as in Extended Data Fig. 8c. Cyan spheres, site 7; dark blue, sites 13 and 151. **e**, Zoomed in view of sites 7, 13, and 151. These side chains are not in direct physical contact; however, site 7 is on a beta strand that undergoes extensive conformational movement when Hsp90 converts between ADP- and ATP-bound states.