Chance, contingency, and necessity in the experimental evolution of ancestral proteins

Victoria Cochran Xie[1,4], Jinyue Pu[1,4,*], Brian P.H. Metzger[2,4], Joseph W. Thornton[2,3,*], Bryan C. Dickinson[1,*]

[1]Department of Chemistry, University of Chicago, Chicago, Illinois, USA
[2]Department of Ecology and Evolution, University of Chicago, Chicago, Illinois, USA
[3]Department of Human Genetics, University of Chicago, Chicago, Illinois, USA
[4]These authors contributed equally: Victoria Cochran Xie, Jinyue Pu, Brian P.H. Metzger

*email: pujy@uchicago.edu, joet1@uchicago.edu, dickinson@uchicago.edu

Keywords: PACE, continuous directed evolution, protein-protein interactions, genetic variance, epistasis, BCL-2, ancestral protein reconstruction

SUMMARY

The extent to which chance and contingency shaped the sequence outcomes of protein evolution is largely unknown. To directly characterize the causes and consequences of chance and contingency, we combined directed evolution with ancestral protein reconstruction. By repeatedly selecting a phylogenetic series of ancestral proteins in the B-cell lymphoma-2 family to evolve the same protein-protein interaction specificities that existed during history, we show that contingency and chance interact to make sequence evolution almost entirely unpredictable over the timescale of metazoan evolution. At any historical moment, multiple sets of mutations can alter or maintain specificity, and chance decides which ones occur. Contingency arises because historical sequence substitutions epistatically altered which mutations are compatible with new or ancestral functions. Evolutionary trajectories launched from different ancestors therefore lead to dramatically different outcomes over phylogenetic time, with virtually no mutations occurring repeatedly in distantly related proteins, even under identical selection conditions.

INTRODUCTION

Whether the world is a necessary outcome of fate-like processes has long fascinated philosophers, historians, and scientists (Aristotle, 1938; Gould, 1989; Jablonski, 2017; Ramsey and Pence, 2016; Travisano et al., 1995). In biology, if a single optimal state is accessible from all starting points, then natural selection is expected to predictably drive the realization of that state, deterministically and independently of initial conditions. Chance and contingency reduce the necessity of evolutionary outcomes in distinct ways and arise from different aspects of the genotype-phenotype relationship. If multiple adaptive outcomes are accessible from a given genetic starting point, then which one is realized depends on chance; a process free of chance

is deterministic and predictably leads to the same outcome again and again (Lobkovsky and Koonin, 2012; Monod, 1972). If the accessibility of outcomes differs among starting genotypes, then the realized outcome of evolution is contingent on the particular starting point from which a historical trajectory is initiated; even in the absence of chance, the course of evolution can be predicted only given full knowledge of the starting genotype and the particular outcomes to which it can lead (Blount et al., 2018).

To systematically characterize how chance and contingency affected historical protein evolution, we would have to travel back in time, re-launch evolution repeatedly from each of various starting points that arose during history, and compare the outcomes realized in replicates from the same starting point (for chance) and from different starting points (for contingency) (Figure 1A). This design cannot be executed directly, but we can come close by reconstructing ancestral proteins as they existed in the deep past (Thornton, 2004) and using them to launch replicated evolutionary trajectories in the laboratory under selection to evolve the same molecular functions that they acquired during history. Studies to date have used less direct approaches to provide partial insights into the effects of chance and contingency on the repeatability of sequence evolution (Orgogozo, 2015; Storz, 2016). Experimental studies have typically replicated trajectories from just one or a few starting points (Bollback and Huelsenbeck, 2009; Counago et al., 2006; Dickinson et al., 2013; Kacar et al., 2017; Meyer et al., 2012; Zheng et al., 2019) – a design that can illuminate the effect of chance but is limited in addressing contingency; moreover, these studies have not used ancestral proteins ordered across time to make their findings historically relevant (Baier et al., 2019; Blount et al., 2012; Kryazhimskiy et al., 2014; Salverda et al., 2011; Wunsche et al., 2017). Mechanistic studies have established that particular historical mutations have different effects when introduced into different ancestral backgrounds, suggesting contingency, but history happened only once, so these studies do not elucidate the role of chance (Gong et al., 2013; Harms and Thornton, 2014; Ortlund et al., 2007; Starr et al., 2018). Case studies show that phenotypic convergence by closely related populations or species in nature sometimes involve the same gene, and occasionally the same mutations, suggesting some degree of necessity; however, these studies do not involve replicate lineages from multiple starting points, so they do not allow chance and contingency to be disentangled (Orgogozo, 2015; Storz, 2016). It is therefore clear that chance and contingency can affect the sequence outcomes of protein evolution, but the extent of their effects, interactions, and causes at historical timescales remain largely unknown.

Here we combine ancestral protein reconstruction with repeated experimental evolution in the B-cell lymphoma-2 (BCL-2) family of proteins. These proteins promote or inhibit apoptosis in eukaryotes (Chipuk et al., 2010; Danial and Korsmeyer, 2004) and are regulated by specific protein-protein interactions (PPIs) with coregulators (Chen et al., 2005; Lomonosova and Chinnadurai, 2008), the structural and biochemical basis of which is well understood (Kale et al., 2018; Petros et al., 2004). BCL-2 family proteins are found throughout Metazoa (Banjara et al., 2020; Lanave et al., 2004) and differ from each other in PPI specificity: proteins in the Myeloid Cell Leukemia Sequence 1 (MCL-1) class bind both the BID and NOXA coregulators, whereas proteins in the BCL-2 class (a subset of the larger BCL-2 protein family) bind BID but not NOXA (Figure 1B) (Certo et al., 2006). The two classes are structurally similar and use the same cleft to interact with their coregulators (Figures 1C and S1). Prior work has identified some determinants of coregulator binding and specificity (Dutta et al., 2010), but little is known

about how the MCL-1 and BCL-2 classes evolved their distinct binding profiles during history. We used ancestral protein reconstruction to characterize the historical trajectory by which these proteins evolved their coregulator specificities and then applied a rapid experimental evolution technique to repeatedly select reconstructed ancestral family members to acquire these same functions. This approach allowed us to quantitatively dissect the roles of chance, contingency, and necessity in the evolution of PPI specificity in this protein family.

RESULTS

BID specificity is derived from an ancestor that bound both BID and NOXA

We first characterized the historical evolution of PPI specificity in the BCL-2 family using ancestral protein reconstruction. We inferred the maximum likelihood phylogeny of the family, recovering the expected sister relationship between the metazoan BCL-2 and MCL-1 classes (Figure 2, Figure S2A). We then reconstructed the most recent common ancestor (AncMB1) of the two classes, which represents a gene duplication that occurred before the last common ancestor of all animals, by inferring the posterior probability distribution of ancestral states and the maximum a posteriori (MAP) amino acid sequence (Supplementary Table S1). We also reconstructed 11 other ancestral proteins that existed along the lineages leading from AncMB1 to human BCL-2 (hsBCL-2) and to human MCL-1 (hsMCL-1).

We then synthesized genes coding for these proteins and experimentally assayed their ability to bind BID and NOXA using a proximity-dependent split RNA polymerase (RNAP) luciferase assay (Figures 1D and 1E) (Pu et al., 2017). AncMB1 – the deepest ancestor reconstructed -- bound both BID and NOXA, as did all ancestral proteins in the MCL-1 clade and extant MCL-1 from humans (Figure 2, Figure S2, Table S1). Ancestral proteins in the BCL-2 clade that existed before the last common ancestor of deuterostomes also bound both BID and NOXA, whereas BCL-2 ancestors within the deuterostomes bound only BID, just as human BCL-2 does. BID specificity therefore evolved when the ancestral ability to bind NOXA was lost. This event occurred between AncB2 (in the ancestral eumetazoan) and AncB4 (in the ancestral deuterostome): the precise timing of the loss of NOXA binding during this interval cannot be resolved by these experiments, because AncB3 (the protostome-deuterostome ancestor) had an intermediate phenotype, binding NOXA weakly.

To further test this history, we characterized the coregulator specificity of extant BCL-2 class proteins from taxonomic groups in particularly informative phylogenetic positions. Those from Cnidaria were activated by both BID and NOXA, whereas those from protostomes and invertebrate deuterostomes were BID-specific (Figure 2, Figure S2, Table S1). These results corroborate the inferences made from ancestral proteins and suggest that the loss of NOXA binding occurred in the BCL-2 class between the last common ancestor of Eumetazoa and before the protostome-deuterostome ancestor. This reconstruction of history was also robust to uncertainty in the ancestral sequence reconstruction, because experiments on “AltAll” proteins at each ancestral node – which combine all plausible alternative amino acid states (PP>0.2) in a single “worst-case” alternative reconstruction – also showed that BID specificity arose between AncB2 and AncB4 (Table S2).

A directed continues evolution system for the rapid, repeated selection of PPI specificity

We next developed a new phage-assisted continuous evolution (PACE) system (Esvelt et al., 2011) to rapidly evolve ancestral and extant BCL-2 family proteins to acquire the same PPI specificities that existed during the family's history (Figures 3A, S3A, and S3B). In general, PACE links the life cycle of M13 bacteriophage to the evolution of a desired function through the inducible expression of an essential phage gene, gIII. By focusing variation and selection on a particular protein of interest, PACE typically allows evolution of a new function in very large, replicated populations over the course of just a few days.

Previous PACE experiments have evolved binding to new protein partners using a bacterial 2-hybrid approach (Badran et al., 2016). To evolve novel PPI specificity, however, requires simultaneous selection for a desired PPI and against an undesired PPI. We therefore developed a PACE system that uses two orthogonal proximity-dependent split RNAPs that imposes both selection and counterselection on defined PPIs (Pu et al., 2017) (Figure 3A). In our system, the N-terminal portion of RNAP is fused to the BCL-2 family protein of interest. The target protein to which binding is desired is fused to the C-terminal portion of an RNAP that recognizes a promoter linked to gIII; binding of the BCL-2 family protein to the target protein reconstitutes this RNAP, causing gIII to be expressed and infectious phage to be produced. To achieve specificity, the protein to which no binding is desired is fused to the C-terminal portion of a different RNAP, which recognizes a promoter that drives expression of a dominant negative form of gIII ($gIII_{neg}$); binding of the BCL-2 family protein to this counterselection protein reconstitutes this second RNAP, producing $gIII_{neg}$ and causing virtually all phage produced in the system to be incapable of infection (Carlson et al., 2014). To increase the mutation rate within the PACE system an arabinose-inducible mutagenesis plasmid (MP) was included in host cells.

We optimized this system using activity-dependent plaque assays and phage growth assays to drive acquisition of the natural binding profiles of BCL-2 and MCL1 proteins. We showed that phage carrying either hsBCL-2 or hsMCL-1 could replicate when BID binding was selected for. Conversely, phage carrying hsBCL-2 could replicate when BID binding was selected for and NOXA binding was selected against, but phage carrying hsMCL-1 could not (Figures 3B and 3C). Likewise, phage carrying hsMCL-1 could replicate when binding NOXA was selected for, whereas phage carrying hsBCL-2 could not. Finally, neither phage could replicate when NOXA binding was selected for and BID binding was selected against.

Characterizing chance, contingency, and necessity

To characterize the extent of chance, contingency, and necessity on the outcomes of evolution, we used our PACE system to drive extant and reconstructed ancestral proteins to repeatedly recapitulate or reverse the historical loss of NOXA binding. Three proteins with the ancestral phenotype—hsMCL-1, AncM6, and AncB1—were selected to acquire the derived BCL-2 phenotype, losing NOXA binding but retaining BID binding (Figure S3C). Conversely, hsBCL-2, AncB5, and AncB4 were evolved to revert to the ancestral phenotype, gaining NOXA binding (Figure S3D). For each starting genotype, we performed three to four replicate experimental evolution trajectories, for a total of 23 separate trajectories (Table S3). All trajectories produced the target phenotype, and we confirmed that the selected PPI specificity had been acquired by randomly isolated phage clones using activity-dependent plaque assays as well as *in vivo* and *in vitro* binding assays (Figures 3D and S3E-L).

We found very limited necessity in the sequence outcomes of directed evolution. High-throughput sequencing of the phage populations revealed that 100 mutant amino acid states (at 75 different sites) evolved to frequency >5% in at least one replicate (Figures 4A-B and S4A-D and Table S4). Of these acquired states, 73 occurred in only a single trajectory, although the 27 that appeared more than once is significantly greater than expected by chance alone ($P<10^{-5}$ by permutation test, given the same number of mutations and replicates). Most of the repeated mutations were observed in replicate trajectories from the same starting genotype, indicating some degree of determinism in evolutionary outcomes from a given starting genotype (Figure 4C). Only four mutations were observed in more than one replicate from different starting genotypes, however, suggesting considerable contingency. Only a single mutation was observed in every replicate from more than one starting point, and this mutation was not experimentally sufficient to confer the desired phenotype (Figure S4E). When mapped onto the protein structure, all repeatedly mutated sites either directly contact the bound peptide or are on secondary structural elements that do so (Figures 4D and 4E), suggesting that evolutionary necessity in this protein family, to the extent that it exists, reflects a limited number of structural mechanisms by which PPIs can be altered.

To quantify the effects of chance and contingency on the genetic outcomes of evolution, we analyzed the genetic variance -- the probability that two alleles, chosen at random, are different in state -- within and between replicates from the same and different starting genotypes. To estimate the effects of chance, we compared the genetic variance between replicates initiated from the same starting genotype (V_g) to the within-replicate genetic variance (V_r). We found that V_g was on average 1.3-fold greater than V_r, indicating that chance causes evolution to produce substantially divergent genetic outcomes (Figure 5A). We estimated the effects of contingency by comparing the genetic variance among replicates from different starting genotypes (V_t) to the genetic variance among replicates from the same starting genotype. Contingency had an even larger effect than chance, increasing V_t by an average of 1.8 fold compared to V_g. Together, chance and contingency had a multiplicative effect, increasing the genetic variance among replicates from different starting genotypes (V_t) by an average of 2.4-fold compared to V_r.

To determine the evolutionary tempo and mode of chance and contingency, we evaluated how chance and contingency changed as the phylogenetic distance between starting genotypes increased. Across all pairs of starting genotypes, the combined effects of chance and contingency increased significantly with phylogenetic distance (slope = 0.208, $P = 1\text{x}10^{-4}$) (Figures 5B and S5A); this relationship remained statistically significant after accounting for phylogenetic non-independence by comparing only successive ancestors along phylogenetic lineages ($P = 0.03$). Across the metazoan evolutionary timescale of our experiments, contingency and chance together increased genetic variance more than three-fold. This relationship was driven almost entirely by an increase in the effects of contingency with phylogenetic distance (slope = 0.128, $P = 5\text{x}10^{-4}$), with the effects of chance displaying a very weak increase (slope=0.019, p=0.002). The combined effect of chance and contingency increased faster with phylogenetic distance than contingency alone did. Chance and contingency therefore magnify each other's effects, increasing the unpredictability, and reducing necessity, in evolutionary outcomes as sequences diverge through history.

Sources of contingency

Contingency arises when sequence outcomes differ between genetic starting points; this conditionality arises when epistasis exists – that is, when the effects of mutations depend on the sequence background into which they are introduced. Epistasis could produce contingency if mutations that confer selected phenotypes in some backgrounds fail to confer that phenotype in others or are not tolerated at all; alternatively, epistasis might merely change the probability that selection favors one mutation over another by altering the relative magnitude of their effect on the phenotype. To experimentally distinguish between these possibilities, we transferred sets of mutations that arose repeatedly during experimental evolution into other starting genotypes and then measured their effects on BID and NOXA binding (Figures 6A and 6B). Eleven of the 12 "swaps" failed to confer the PPI specificity on other genetic backgrounds that they did on their evolved genetic background. Of the six swaps of mutation sets that arose when MCL-1-like proteins were evolved to lose NOXA binding and preserve BID binding, five compromised BID binding and three had no effect on NOXA binding when introduced into other backgrounds. Of the six swaps of mutation sets that caused BCL-2-like proteins to gain NOXA binding, four failed to gain any detectable NOXA binding and one compromised BID binding in other backgrounds. The only case in which mutations that conferred the target phenotype during experimental evolution had the same effect in another background was the swap of mutations evolved on AncB5 into AncB4, and these two genotypes are more similar to each other than any other pair. Thus, contingency arose because of strong epistatic interactions between the mutations that confer new specificities in our experiments and the historical substitutions that occurred during the intervals between ancestral proteins; these substitutions transiently opened and blocked routes to adaptive phenotypes by making specificity-changing mutations subsequently—and previously—deleterious or invisible to selection.

To identify when these epistatic effects arose, we mapped these incompatibilities onto the phylogeny (Figure 6C). If mutations that arise in PACE using some ancestral protein as a starting point decrease BID binding when introduced into one of that protein's descendants, then restrictive substitutions must have occurred historically on the branches between the two proteins. Conversely, if PACE-derived mutations compromise or abolish BID activity when introduced into a protein's ancestor, then historical permissive substitutions occurred between them. Similarly, if PACE-derived mutations confer the selected change in NOXA activity in one protein but fail to have that effect in its ancestors, then potentiating substitutions must have occurred on the phylogeny. Finally, if PACE-derived mutations fail to have the effect on NOXA activity in a protein's descendants, then depotentiating substitutions occurred during history.

We found that all four types of epistatic effects occurred, with multiple types of contingency-inducing substitutions present on most branches. The only exception—the branch from AncB4 to AncB5, on which only depotentiating substitutions occurred—is the branch immediately after NOXA function changed during history, the shortest of the branches examined, and the one with the smallest effect of contingency on genetic variance (Figure 5A). Even across this branch, the PACE mutations that restore the ancestral PPI specificity in AncB4 can no longer do so in AncB5. These results indicate that the paths through sequence space that allow new NOXA functions to evolve repeatedly changed during the BCL-2 family's history, even during intervals when the proteins' PPI binding profiles did not change.

Consistent with this conclusion, the mutations that altered NOXA binding during PACE almost never recapitulated or reversed substitutions that occurred during the historical interval in which NOXA binding changed (Figures 5C and S5B). The sole exceptions, f160L and y259F, arose when AncB4 was evolved to regain the ancestral NOXA binding and are reversals to ancestral amino acids observed in AncB1, prior to the historic loss of NOXA binding. No PACE trajectories for the loss of NOXA binding recapitulated substitutions that occurred during this interval. These observations suggest that the substitutions that changed PPI specificity during historical evolution had the capacity to confer this function—and, if reverted, to restore the ancestral function—only during a very limited temporal window.

Sources of chance and determinism

For chance to strongly influence the outcomes of adaptive evolution, multiple paths to a selected phenotype must be accessible with reasonably similar probabilities of being taken; such a situation could arise if several different mutations (or sets of mutations) can confer a new function, or if there are mutations that have no effect on function that can accompany function-changing mutations by chance. To distinguish between these possibilities, we measured the functional effects of sets of mutations that arose in different replicates when hsMCL-1 was evolved to lose NOXA binding (Figures 6D and S6A). One mutation (v189G) was found at high frequency in all four replicates, but it was always accompanied by other mutations, which varied among trajectories. We found that the apparently deterministic mutation v189G was a major contributor to the loss of NOXA binding, but it had this effect only in the presence of the other mutations, which did not decrease NOXA binding on their own. v189G therefore required permissive mutations, and there are multiple sets of mutations that can exert that effect, so which ones occur in any replicate is a matter of chance. All permissive mutations were located near the NOXA binding cleft (Figure 6E). Other starting genotypes showed a similar pattern of multiple key mutations capable of conferring the selected function (Figures S6B and S6C).

To better understand the genetic causes of chance's effects on BCL-2 family protein evolution, we performed PACE experiments in which we evolved hsBCL-2 to retain its BID binding, without selection for or against NOXA binding, and then screened for variants that fortuitously gained NOXA binding by activity-dependent plaque assay (Figures S7A and S7B). This strategy allowed us to distinguish between the number of mutations that can confer this phenotype and the influence of selection in favoring a subset of mutations that most rapidly increase NOXA binding. All four replicate populations produced clones that gained NOXA binding at a frequency of about ~0.1% to 1% – lower than when NOXA binding was selected for, but five orders of magnitude higher than when NOXA binding was selected against (Figure 7A). From each replicate, we sequenced three NOXA-binding clones and found that all but one of them contained mutation r165L (Figure 7B), a mutation that also occurred at high frequency in hsBCL-2 trajectories when NOXA binding was selected for. We introduced r165L into hsBCL-2 and found that it conferred significant NOXA binding with little effect on BID binding (Figure S7C). Several other mutations also appeared repeatedly in clones that fortuitously acquired NOXA binding, which were also acquired under selection for NOXA binding. A similar pattern of common mutations was observed in AncB4 and AncB5 clones that fortuitously or selectively evolved NOXA binding (Figures S7F-J). These observations indicates that the determinism

observed in these experiments arises because there are few genotypes that can increase NOXA binding while retaining BID binding, rather than because there are many such genotypes, but under strong selection a few are strongly favored over others.

Chance and contingency alter accessibility of new functions

Although we found a strong influence of chance and contingency in the sequence outcomes of evolution across ancestral starting points, they had little influence on the acquisition of the historical functions per se, because all replicates from all starting points acquired the selected BCL-2- or MCL-1-like specificity. However, changes in function that did not evolve during history might be subject to more chance or contingency. Using PACE and several proteins that can bind both BID and NOXA as starting points, we selected for variants that could bind NOXA but not BID, a phenotype not known in nature. hsMCL-1 readily evolved the selected phenotype, but two experimentally-evolved variants of hsBCL-2 that had acquired NOXA binding went extinct under the same selection conditions (Figures 7C and S7K-M). The inability of the derived hsBCL-2 genotypes to acquire NOXA specificity was not attributable to a general lack of evolvability, because these same genotypes successfully evolved in a separate PACE experiment to lose their NOXA binding but retain BID binding (Figure S7N). These results illustrate that contingency can influence the ability to evolve new functions.

DISCUSSION

Without chance, contingency in history would be inconsequential, because all phylogenetic lineages launched from a common ancestor would always lead to the same intermediate steps and thus the same ultimate outcomes. On the other hand, without contingency, chance events would be inconsequential, because the mutations that happen to occur in any time interval would not affect the next set of available steps or ultimate outcomes; every path that was ever open would remain forever so. Our work shows that, across the timescale of BCL-2 sequence evolution, the interaction of chance and contingency eliminates virtually all traces of necessity and repeatability in sequence evolution under strong selection (except for sites that are constrained and never vary).

To systematically assess chance and contingency in BCL-2 evolution, we developed an efficient new method to simultaneously select for particular interactions and against others, which allowed us to drive the evolution of new PPI specificities in multiple replicates without severe bottlenecks in just days. By applying this technology to reconstructed ancestral proteins, our experiments directly illuminate how the sequence divergence that occurred during BCL-2 historical evolution generates chance and contingency in the experimental evolution of the same functional specificities that evolved during the family's history. Our design does not directly reveal chance and contingency during the historical evolution of these proteins, because the selection pressures, environmental conditions, and population parameters that pertained in the deep past are unknown. It is likely, however, that chance and contingency were at least as significant during history as in our experiments for several reasons. First, our experiments likely favor determinism in the evolutionary process due to the very large population sizes, extremely strong selection pressures, and high mutation rates, all of which were directed at a single gene; if, as seems likely, BCL-2 historical evolution involved smaller

populations, weaker selection, lower mutation rates, and a larger genetic "target size" for adaptation, then chance would have played an even larger role during history than in our experiments. Second, chance and contingency both arise from the relationship between genotype and function—chance from the number of accessible genotypes that confer a selected-for function, and contingency from epistasis-induced differences in the accessibility of paths among starting genotypes. The fold and structural basis for binding of coregulatory proteins is conserved across extant BCL-2 family members, so it is unlikely that these factors, and the ways that they lead to chance and contingency, were dramatically different between the evolutionary processes underlying changes in BCL-2 specificity during history and in our experiments.

Epistasis is a common feature of protein structure and function, so we expect that the accumulating effect of contingency that we observed across phylogenetic time among BCL-2 family members will be a general feature of protein evolution, although its rate and extent is likely to vary among folds and functions (Chandler et al., 2013; Harms and Thornton, 2013; Shah et al., 2015; Storz, 2018). The influence of chance is likely to depend on the particular function that is evolving: more determinism is expected for functions with very narrow sequence-structure-function-constraints (e.g. (Hawkins et al., 2018; Karageorgi et al., 2019; Menendez-Arias, 2010; Meyer et al., 2012; Salverda et al., 2011; Storz, 2016), than those for which sequence requirements are less strict (e.g. (Blount et al., 2012; Starr et al., 2017; Yokoyama et al., 2008; Zheng et al., 2019). In the extreme, when diffuse selection pressures are imposed on whole organisms, virtually no repeatability has been observed among replicates, because loci across the entire genome are potential sources of adaptive mutations (Kryazhimskiy et al., 2014; Wunsche et al., 2017).

Our results have implications for protein biochemistry, engineering, and evolution. First, we found no evidence that ancestral proteins were more "evolvable" than extant proteins: the selected-for phenotypes readily evolved from both extant and ancestral proteins; further, the effect of chance in these processes was virtually constant across about a billion years of evolution, indicating that the number of accessible mutations that can confer a selected-for function was no greater in the past than it is in the present. Second, the strong effect of contingency suggests that efforts to produce proteins with new functions by design or directed evolution will be most effective if they use multiple different protein sequences as starting points, ideally separated by long intervals of sequence evolution. Third, our finding that affinity for new partners can evolve fortuitously during selection to maintain existing binding indicates that new interactions can sometimes be acquired neutrally and then, if they become functionally significant, be amplified or preserved by selection; conversely, maintaining specificity during natural and directed evolution requires selection against off-target interactions (Levin et al., 2009).

Finally, our observations suggest that the sequence-structure-function associations apparent in sequence alignments are, to a significant degree, the result of shared but contingent constraints that were produced by chance events during history (Gong et al., 2013; Harms and Thornton, 2014; Starr et al., 2018; Starr et al., 2017). Present-day proteins are physical anecdotes of a particular history: they reflect the interaction of accumulated chance events during descent from common ancestors with necessity imposed by physics, chemistry and natural selection. Apparent "design principles" in extant or evolved proteins express not

how things must be—or even how they would be best—but rather the contingent legacy of the constraints and opportunities that those molecules just happen to have inherited.

ACKNOWLEDGEMENTS

We thank members of the Dickinson and Thornton groups for helpful comments on the manuscript, S. Ahmadiantehrani for editing, and R. Ranganathan for the use of the Illumina MiSeq instrument. This work was supported by CAREER Award 1749364 from the National Science Foundation (NSF) to B.C.D and National Institutes of Health grants R01GM131128 and R01GM121931 to J.W.T. V.C.X was supported by an NSF Graduate Research Fellowship (DGE-1746045). B.P.H.M. was supported by a NIH NRSA award (F32GM122251).

AUTHOR CONTRIBUTIONS

All authors contributed to conception of the project. BCD, VCX, and JP designed the PACE dual-selection system, which VCX and JP engineered, optimized, and implemented. VCX and JP performed PACE, biochemical assays, and sequencing experiments, with input from BPHM. BPHM and JWT developed and designed the evolutionary and genetic analyses. BPHM led and performed the phylogenetic, genetic, and evolutionary analyses, with input from VCX and JP. All authors contributed to writing the manuscript.

DECLARATION OF INTERESTS

J.P. and B.C.D. have a patent on the proximity-dependent split RNAP technology used in this work. The content is solely the responsibility of the authors and the funders had no input on the study design, analysis, or conclusions.

DATA AVAILABILITY

Raw high-throughput sequencing data are available on SRA (accession number PRJNA647218). Processed HTS data are available on Dryad (https://datadryad.org/stash/share/Ty32n-2c8nUZuiFegxDdy7wNeGDEsqlwtL8tOKz1RWU).

CODE AVAILABILITY

Code used for analysis and phylogenetic data are available at https://github.com/JoeThorntonLab/BCL2.ChanceAndContingency

REFERENCES

Aristotle (1938). Categories. On interpretation. Prior analytics. (Cambridge, MA: Harvard University Press).

Badran, A.H., Guzov, V.M., Huai, Q., Kemp, M.M., Vishwanath, P., Kain, W., Nance, A.M., Evdokimov, A., Moshiri, F., Turner, K.H., *et al.* (2016). Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. Nature *533*, 58-63.

Baier, F., Hong, N., Yang, G., Pabis, A., Miton, C.M., Barrozo, A., Carr, P.D., Kamerlin, S.C., Jackson, C.J., and Tokuriki, N. (2019). Cryptic genetic variation shapes the adaptive evolutionary potential of enzymes. Elife *8*.

Banjara, S., Suraweera, C.D., Hinds, M.G., and Kvansakul, M. (2020). The Bcl-2 Family: Ancient Origins, Conserved Structures, and Divergent Mechanisms. Biomolecules *10*.

Blount, Z.D., Barrick, J.E., Davidson, C.J., and Lenski, R.E. (2012). Genomic analysis of a key innovation in an experimental Escherichia coli population. Nature *489*, 513-518.

Blount, Z.D., Lenski, R.E., and Losos, J.B. (2018). Contingency and determinism in evolution: Replaying life's tape. Science *362*.

Bollback, J.P., and Huelsenbeck, J.P. (2009). Parallel genetic evolution within and between bacteriophage species of varying degrees of divergence. Genetics *181*, 225-234.

Carlson, J.C., Badran, A.H., Guggiana-Nilo, D.A., and Liu, D.R. (2014). Negative selection and stringency modulation in phage-assisted continuous evolution. Nat Chem Biol *10*, 216-222.

Certo, M., Del Gaizo Moore, V., Nishino, M., Wei, G., Korsmeyer, S., Armstrong, S.A., and Letai, A. (2006). Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members. Cancer Cell *9*, 351-365.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet *29*, 358-366.

Chen, L., Willis, S.N., Wei, A., Smith, B.J., Fletcher, J.I., Hinds, M.G., Colman, P.M., Day, C.L., Adams, J.M., and Huang, D.C. (2005). Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. Mol Cell *17*, 393-403.

Chipuk, J.E., Moldoveanu, T., Llambi, F., Parsons, M.J., and Green, D.R. (2010). The BCL-2 family reunion. Mol Cell *37*, 299-310.

Counago, R., Chen, S., and Shamoo, Y. (2006). In vivo molecular evolution reveals biophysical origins of organismal fitness. Mol Cell *22*, 441-449.

Danial, N.N., and Korsmeyer, S.J. (2004). Cell death: critical control points. Cell *116*, 205-219.

Dickinson, B.C., Leconte, A.M., Allen, B., Esvelt, K.M., and Liu, D.R. (2013). Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. Proc Natl Acad Sci U S A *110*, 9007-9012.

Dutta, S., Gulla, S., Chen, T.S., Fire, E., Grant, R.A., and Keating, A.E. (2010). Determinants of BH3 binding specificity for Mcl-1 versus Bcl-xL. J Mol Biol *398*, 747-762.

Esvelt, K.M., Carlson, J.C., and Liu, D.R. (2011). A system for the continuous directed evolution of biomolecules. Nature *472*, 499-503.

Gong, L.I., Suchard, M.A., and Bloom, J.D. (2013). Stability-mediated epistasis constrains the evolution of an influenza protein. Elife *2*, e00631.

Gould, S.J. (1989). Wonderful life : the Burgess Shale and the nature of history, 1st edn (New York: W.W. Norton).

Harms, M.J., and Thornton, J.W. (2013). Evolutionary biochemistry: revealing the historical and physical causes of protein properties. Nat Rev Genet *14*, 559-571.

Harms, M.J., and Thornton, J.W. (2014). Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature *512*, 203-207.

Hawkins, N.J., Bass, C., Dixon, A., and Neve, P. (2018). The evolutionary origins of pesticide resistance. Biol Rev Camb Philos Soc.

Jablonski, D. (2017). Approaches to Macroevolution: 1. General Concepts and Origin of Variation. Evol Biol *44*, 427-450.

Kacar, B., Ge, X., Sanyal, S., and Gaucher, E.A. (2017). Experimental Evolution of Escherichia coli Harboring an Ancient Translation Protein. J Mol Evol *84*, 69-84.

Kale, J., Osterlund, E.J., and Andrews, D.W. (2018). BCL-2 family proteins: changing partners in the dance towards death. Cell Death Differ *25*, 65-80.

Karageorgi, M., Groen, S.C., Sumbul, F., Pelaez, J.N., Verster, K.I., Aguilar, J.M., Hastings, A.P., Bernstein, S.L., Matsunaga, T., Astourian, M.*, et al.* (2019). Genome editing retraces the evolution of toxin resistance in the monarch butterfly. Nature *574*, 409-412.

Kryazhimskiy, S., Rice, D.P., Jerison, E.R., and Desai, M.M. (2014). Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science *344*, 1519-1522.

Lanave, C., Santamaria, M., and Saccone, C. (2004). Comparative genomics: the evolutionary history of the Bcl-2 family. Gene *333*, 71-79.

Levin, K.B., Dym, O., Albeck, S., Magdassi, S., Keeble, A.H., Kleanthous, C., and Tawfik, D.S. (2009). Following evolutionary paths to protein-protein interactions with high affinity and selectivity. Nat Struct Mol Biol *16*, 1049-1055.

Lobkovsky, A.E., and Koonin, E.V. (2012). Replaying the tape of life: quantification of the predictability of evolution. Front Genet *3*, 246.

Lomonosova, E., and Chinnadurai, G. (2008). BH3-only proteins in apoptosis and beyond: an overview. Oncogene *27 Suppl 1*, S2-19.

Menendez-Arias, L. (2010). Molecular basis of human immunodeficiency virus drug resistance: an update. Antiviral Res *85*, 210-231.

Meyer, J.R., Dobias, D.T., Weitz, J.S., Barrick, J.E., Quick, R.T., and Lenski, R.E. (2012). Repeatability and contingency in the evolution of a key innovation in phage lambda. Science *335*, 428-432.

Monod, J. (1972). Chance and necessity; an essay on the natural philosophy of modern biology (New York,: Vintage Books).

Orgogozo, V. (2015). Replaying the tape of life in the twenty-first century. Interface Focus *5*, 20150057.

Ortlund, E.A., Bridgham, J.T., Redinbo, M.R., and Thornton, J.W. (2007). Crystal structure of an ancient protein: evolution by conformational epistasis. Science *317*, 1544-1548.

Petros, A.M., Olejniczak, E.T., and Fesik, S.W. (2004). Structural biology of the Bcl-2 family of proteins. Biochim Biophys Acta *1644*, 83-94.

Pu, J., Zinkus-Boltz, J., and Dickinson, B.C. (2017). Evolution of a split RNA polymerase as a versatile biosensor platform. Nat Chem Biol *13*, 432-438.

Ramsey, G., and Pence, C.H. (2016). Chance in evolution (Chicago ; London: The University of Chicago Press).

Salverda, M.L., Dellus, E., Gorter, F.A., Debets, A.J., van der Oost, J., Hoekstra, R.F., Tawfik, D.S., and de Visser, J.A. (2011). Initial mutations direct alternative pathways of protein evolution. PLoS Genet *7*, e1001321.

Shah, P., McCandlish, D.M., and Plotkin, J.B. (2015). Contingency and entrenchment in protein evolution under purifying selection. Proc Natl Acad Sci U S A *112*, E3226-3235.

Starr, T.N., Flynn, J.M., Mishra, P., Bolon, D.N.A., and Thornton, J.W. (2018). Pervasive contingency and entrenchment in a billion years of Hsp90 evolution. Proc Natl Acad Sci U S A *115*, 4453-4458.

Starr, T.N., Picton, L.K., and Thornton, J.W. (2017). Alternative evolutionary histories in the sequence space of an ancient protein. Nature *549*, 409-413.

Storz, J.F. (2016). Causes of molecular convergence and parallelism in protein evolution. Nat Rev Genet *17*, 239-250.

Storz, J.F. (2018). Compensatory mutations and epistasis for protein function. Curr Opin Struct Biol *50*, 18-25.

Thornton, J.W. (2004). Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet *5*, 366-375.

Travisano, M., Mongold, J.A., Bennett, A.F., and Lenski, R.E. (1995). Experimental tests of the roles of adaptation, chance, and history in evolution. Science *267*, 87-90.

Wunsche, A., Dinh, D.M., Satterwhite, R.S., Arenas, C.D., Stoebel, D.M., and Cooper, T.F. (2017). Diminishing-returns epistasis decreases adaptability along an evolutionary trajectory. Nat Ecol Evol *1*, 61.

Yokoyama, S., Tada, T., Zhang, H., and Britt, L. (2008). Elucidation of phenotypic adaptations: Molecular analyses of dim-light vision proteins in vertebrates. Proc Natl Acad Sci U S A *105*, 13480-13485.

Zheng, J., Payne, J.L., and Wagner, A. (2019). Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. Science *365*, 347-353.

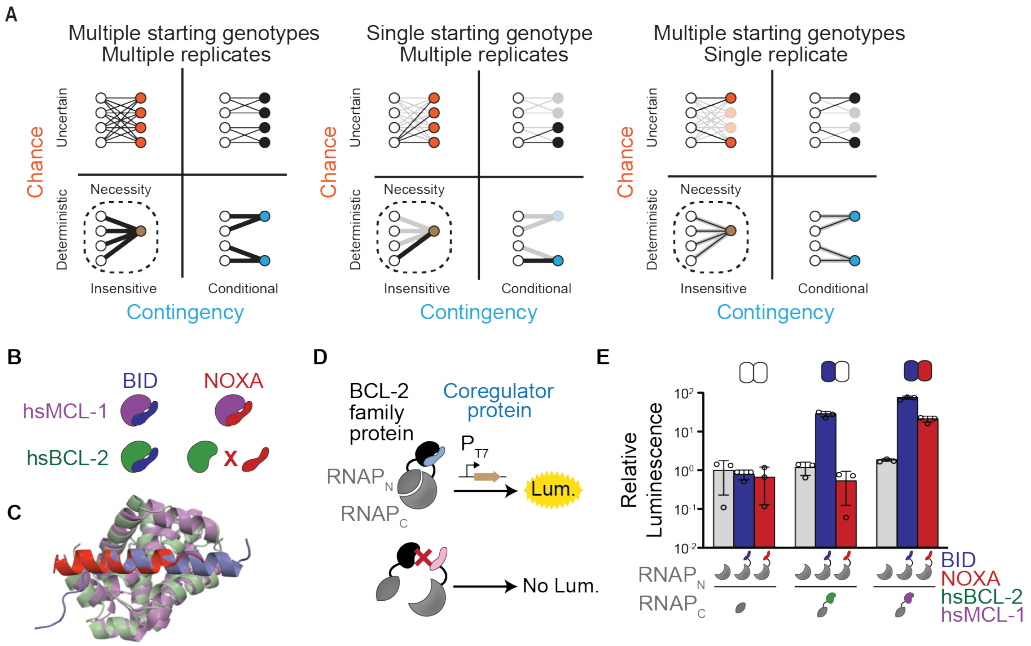


Figure 1. Assessing chance and contingency in the BCL-2 family of proteins

(A) The outcomes of evolution can be influenced by chance (y-axis) and/or contingency (x-axis). Black lines connect starting genotypes (white circles) to evolutionary outcomes; thickness is proportional to a path's probability. Each cluster depicts evolutionary outcomes under the influence of chance (orange), contingency (blue), or both (black); outcomes are necessary (brown, with dotted line) when neither is important. *Left:* evolution experiments in which replicate trajectories are initiated from multiple starting points can distinguish chance and contingency. *Middle*, multiple replicates from one starting point can detect chance but not contingency, because outcomes from other starting points are not observed (faded lines and circles). *Right*, initiating one replicate from multiple starting points underestimates chance.

(B) Protein binding specificities of extant BCL-2 family members. Human MCL-1 (hsMCL-1, purple) binds BID (blue) and NOXA (red), while human BCL-2 (hsBCL-2, green) binds BID but not NOXA.

(C) Crystal structures of human MCL-1 (purple) bound to NOXA (red, PDB 2nla), and BCL-xL (green, a closely related paralog of BCL-2) bound to BID (blue, PDB 4qve).

(D) Schematic of the luciferase reporter assay to assess PPIs. If a BCL-2 family protein (black) binds a coregulator protein (blue), the split T7 RNAP biosensor (gray) assembles and drives luciferase expression. If a coregulator (pink) is not bound, no luciferase is expressed.

(E) Interactions of human BCL-2 and MCL-1 with BID (blue bars) and NOXA (red) in the luciferase assay, compared to no-coregulator control (gray). Activity is scaled relative to no-coregulator control with no-BCL-2 protein. Columns and error bars, mean ± SD of three biological replicates (circles). Shaded boxes above show the same data in heatmap form: BID activity is normalized relative to hsBCL-2 with BID; NOXA activity is normalized to hsMCL-1 with NOXA.

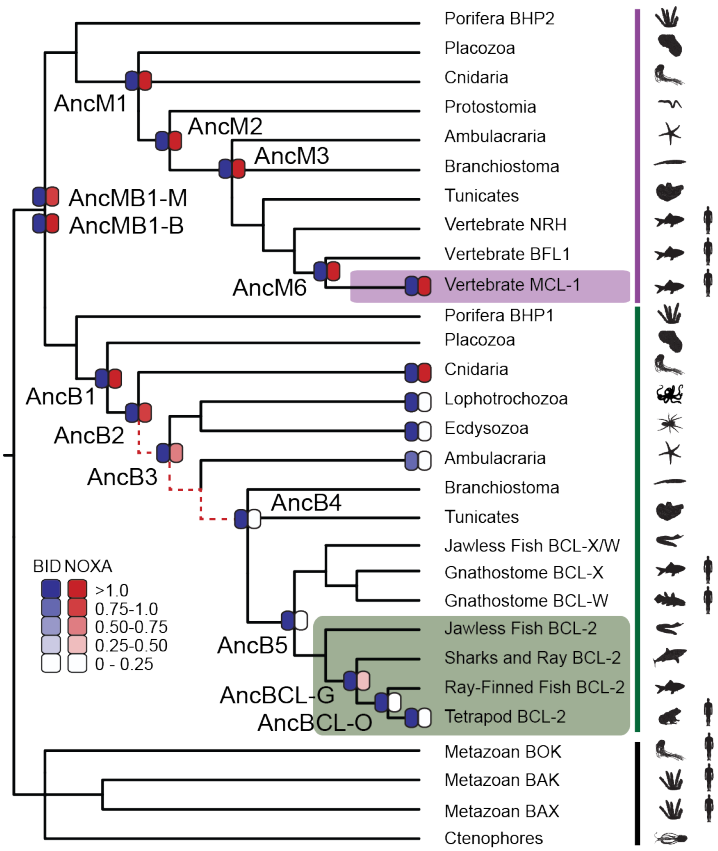
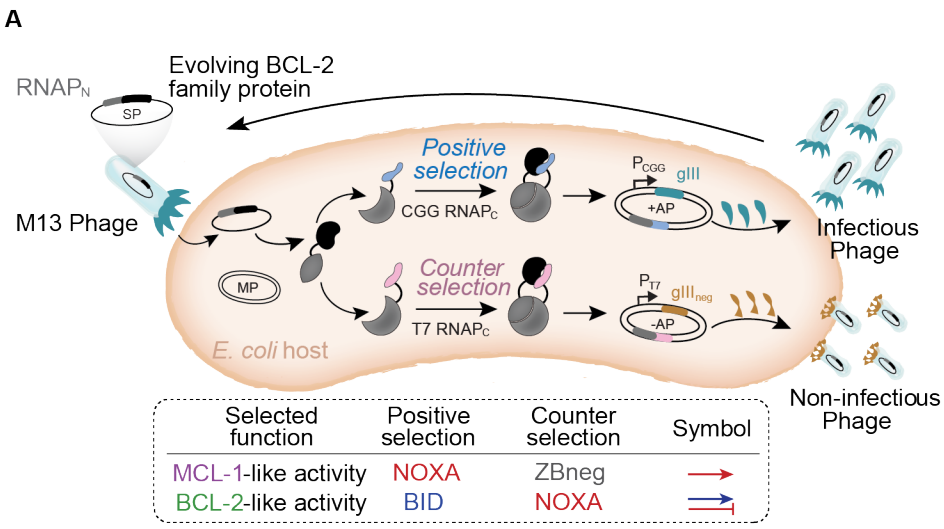
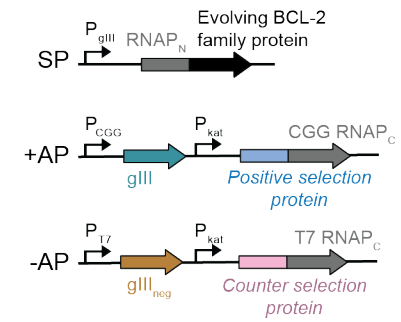
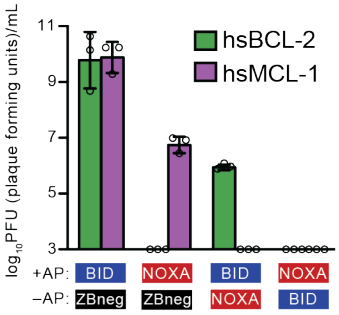


Figure 2. BID specificity was acquired during vertebrate BCL-2 evolution.
Reduced maximum likelihood phylogeny of BCL-2 family proteins. Purple bar, MCL-1 class; green bar, BCL-2 class. The phylogeny was rooted on outgroup using paralogs BOX, BAK, and BAX (black bar). Heatmaps indicate BID (blue) and NOXA (red) binding measured using the luciferase assay as in Figure 1. Each shaded box shows the normalized mean of three biological replicates. Red dotted lines, interval during which NOXA binding was lost, yielding BID specificity in the BCL-2 protein of vertebrates (green box). Purple box, vertebrate MCL-1. Silhouettes, representative species in each terminus. AncMB-M and -B are alternative reconstructions using different approaches to alignment ambiguity (see Methods).







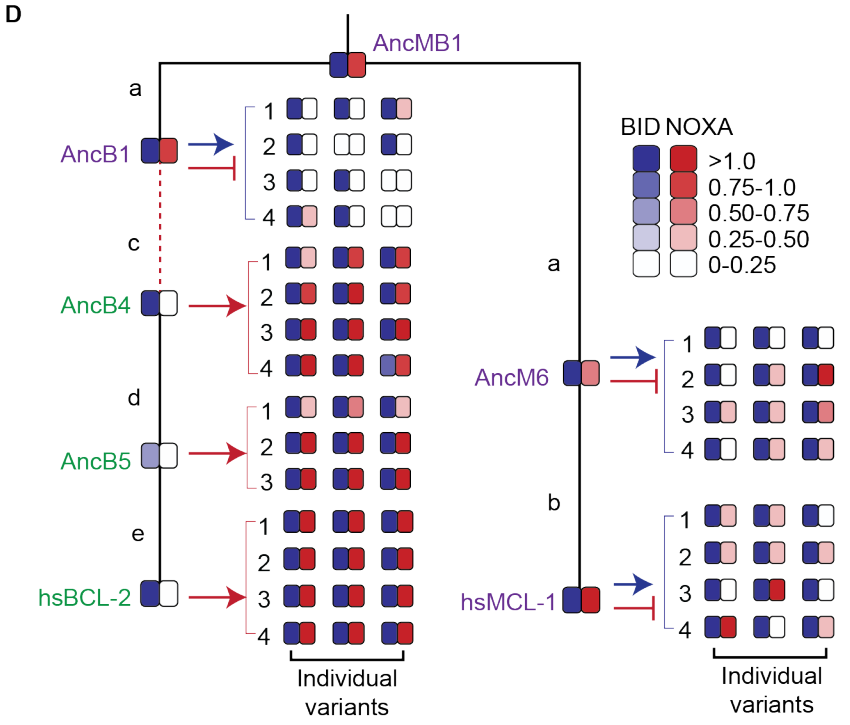


Figure 3. Continuous directed evolution of modern and ancestral BCL-2 family proteins

(A) PACE system for evolving specificity in PPIs. A BCL-2 family protein (black) is fused to the N-terminal portion of T7 RNA polymerase ($RNAP_N$, gray) and placed in the phage genome (SP). Host *E. coli* carry a positive selection accessory plasmid (+AP), containing gIII (teal) driven by a $T7_{CGG}$-promoter (P_{CGG}) and the C-terminal portion of $T7_{CGG}$ RNAP (CGG $RNAP_C$) fused to the target binding protein (light blue). Target binding reconstitutes $RNAP_{CGG}$ and drives expression of gIII, generating infectious phage. The counterselection accessory plasmid (-AP) contains dominant negative gIII ($gIII_{neg}$, gold) driven by a T7 promoter (P_{T7}) and the C-terminal portion of T7 RNA polymerase (T7 RNAP) fused to a counterselection protein. An arabinose-inducible mutagenesis plasmid (MP) increases the mutation rate. To evolve BCL-2 like specificity, positive selection to bind BID was imposed with counterselection to avoid binding NOXA (blue arrow and red bar). To evolve MCL-1 like activity, positive selection to bind NOXA (red arrow) was imposed with counterselection to avoid nonspecific binding using a control zipper peptide (ZBneg).

(B) Map of the phage selection plasmid (SP) and the positive and counterselection accessory plasmids (+AP and -AP). hsBCL-2 (green). hsMCL-1 (purple).

(C) Growth assays to assess selection and counterselection. PFU is shown after culturing 1000 phage containing hsBCL-2 (green) or hsMCL-1 (purple) on *E. coli* containing various APs. Detection limit 10^3 PFU/mL. Bars show mean ± SD of three replicates (circles).

(D) Phenotypic outcomes of directed evolution experiments. Each starting protein was evolved in multiple independent replicate trajectories (indexed by number) under the specified selection conditions (arrows, selection for binding to BID (blue) or NOXA (red); blunt bars, selection against binding). Green, proteins selected to gain NOXA binding; purple, proteins selected to lose NOXA binding. Heatmaps, luciferase assay activity (mean of three biological replicates) of starting proteins and three randomly isolated clones from each replicate evolved population. Phylogenetic relationships of extant and ancestral genotypes used as PACE starting points are shown. Red dashed line, interval during which NOXA binding was historically lost, yielding BID specificity in the BCL-2 clade. Letters, index of phylogenetic intervals between ancestral proteins referred to in Fig. 5.

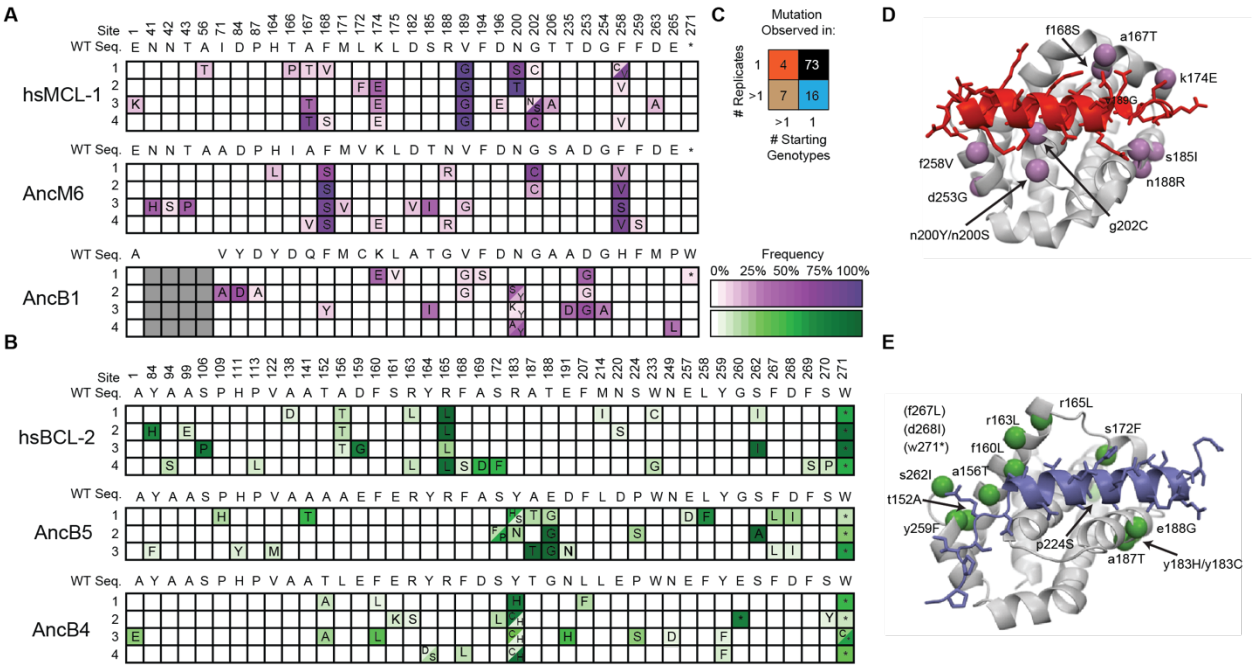


Figure 4. Chance and contingency shape evolutionary outcomes

(A) Frequency of acquired states in PACE experiments when proteins with MCL-1-like specificity were selected to maintain BID and lose NOXA binding. Rows, outcomes of each replicate trajectory. Columns, sequence sites that acquired one or more non-wildtype amino acids (letters in cells) at frequency >5%; color saturation shows the frequency of the acquired state. Site numbers and wildtype (WT) states are listed. Gray, sites that do not exist in AncB1.

(B) Frequency of acquired states when BCL-2-like proteins were selected to gain NOXA binding.

(C) Repeatability of acquired states across replicates. The 100 non-WT states acquired in all experiments were categorized as occurring in 1 or >1 replicate trajectory from 1 or >1 unique starting genotype, with the number in each category shown. The vast majority of states evolved in just one replicate from one starting point (black).

(D) Location of repeated mutations when hsMCL-1, AncM6, and AncB1 were selected to lose NOXA binding (purple spheres), represented on the structure of MCL-1 (gray) bound to NOXA (red, PDB 2nla).

(E) Location of repeated mutations when hsBCL-2, AncB5, and AncB4 were selected to gain NOXA binding (green spheres), on the structure of hsBCL-xL (gray) bound to BID (blue, PDB 4qve).

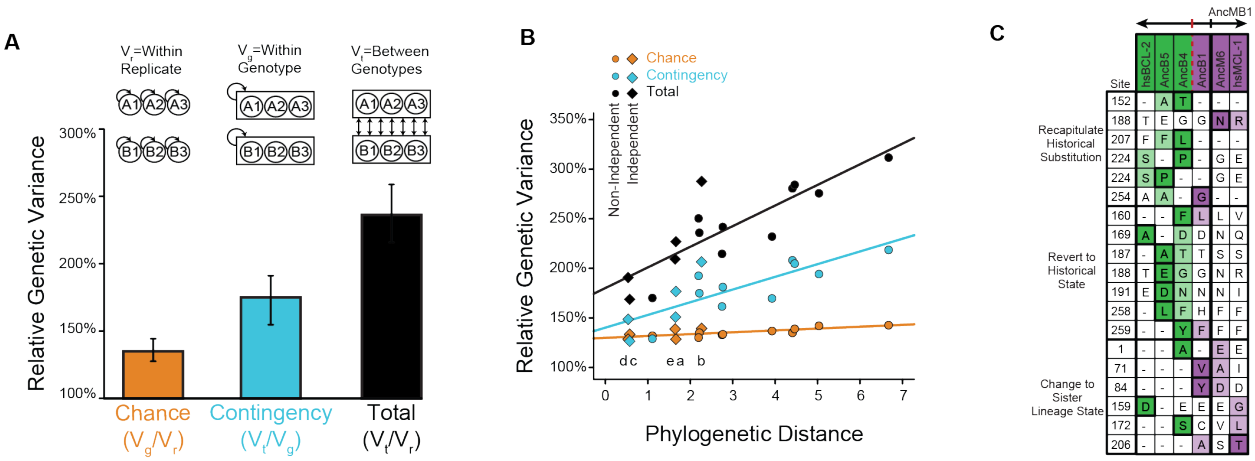


Figure 5. Effect of chance and contingency

(A) Variation in evolutionary sequence outcomes caused by chance (orange), contingency (teal), and both (black). Inset: schematic for estimating the effects of chance and contingency. Chance was estimated as the average genetic variance among replicates from the same starting genotype (V_g) divided by the within-replicate genetic variance (V_r). Contingency was estimated as the average genetic variance among replicates from different starting genotypes (V_t) divided by the average genetic variance among replicates from the same starting genotype (V_g). Combined effects of chance and contingency were estimated as the average genetic variance among replicates from different starting genotypes (V_t) compared to the within replicate genetic variance (V_r). Error bars, 95% confidence intervals on the mean as estimated by bootstrapping PACE replicates. Variance is the probability that two randomly drawn alleles are different in state.

(B) Change in the effects of chance and contingency over phylogenetic time. Each point represents a pair of starting proteins used for PACE, plotted by phylogenetic distance (the total length of branches separating them, in substitutions per site) versus the effects of chance, contingency, or both when PACE outcomes are compared between them, as defined in panel A. Solid lines, best-fit linear regression. Diamonds, points that are phylogenetically independent (i.e. branches connecting direct ancestors and descendants). Circles, non-independent points. Letters indicate the phylogenetic branch indexed in Figure 3D.

(C) States acquired during PACE that also occurred during evolutionary history. Rows, sites at which a state was acquired at >5% frequency in a PACE experiment that is also found in one or more of the ancestral or extant proteins (columns). Dark shaded cells, wild-type state in the protein used as the starting point for the PACE trajectory. Light shaded cells, acquired PACE state in the column for the historical protein in which the state occurs. Unshaded cells with dashes have the same state as the PACE starting point. Purple, proteins with PPI specificity like MCL-1; green, like BCL-2. PACE mutations are grouped into those that i) recapitulate historical substitutions by acquiring a derived state that evolved in history in a descendant of the ancestor used as the PACE starting point; ii) reverse historical substitutions by acquiring a state from an ancestor earlier than the PACE starting point; or iii) confer a state acquired in history along a different lineage.

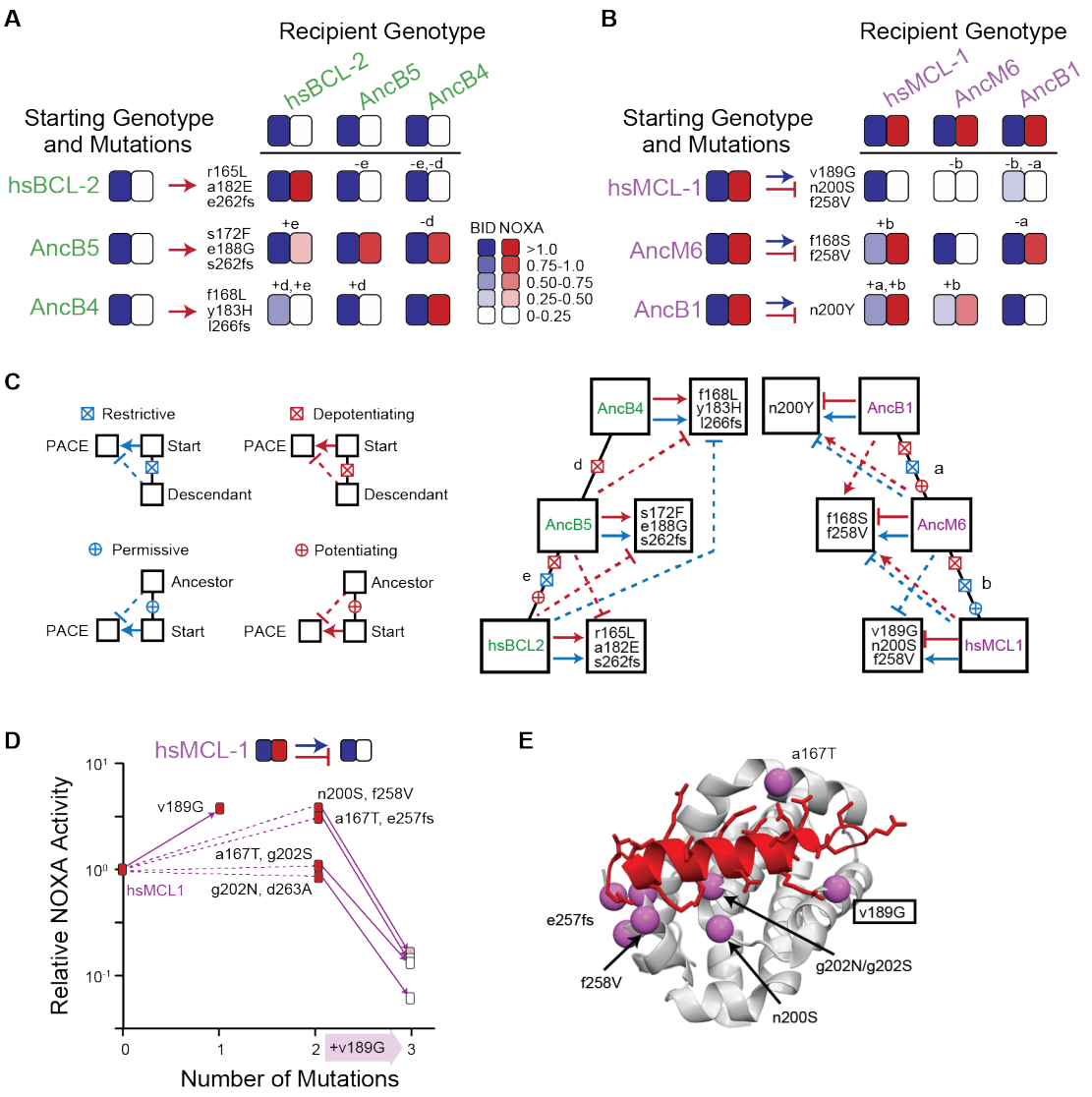


Figure 6. Sources of chance and contingency

(A) Epistatic incompatibility of PACE mutations in other historical proteins. Effects on activity are shown when amino acids acquired in PACE under selection to acquire NOXA binding (red arrows) are introduced into ancestral and extant proteins. The listed mutations that occurred during PACE launched from each starting point (rows) were introduced as a group into the protein listed for each column. Observed BID (blue) and NOXA (red) activity in the luciferase assay for each mutant protein are shown as heatmaps (normalized mean of three biological replicates). Letters above each cell indicate the phylogenetic branch in Figure 3D that connects the PACE starting genotype to the recipient genotype for that experiments. Plus and minus signs indicate whether mutations were introduced into a descendant or more ancestral sequence, respectively.

(B) Effects on activity when amino acids acquired in PACE under selection to lose NOXA binding and acquire BID binding are introduced into different ancestral and extant proteins, represented as in panel A.

(C) Epistatic interactions between historical substitutions and PACE mutations. *Left*: schematic of types of interactions. Restrictive historical substitutions cause mutations that maintain BID

activity in PACE to abolish BID activity when introduced into later historical proteins. Permissive substitutions cause PACE mutations to abolish BID activity when introduced into earlier but not later historical proteins. Potentiating (or depotentiating) substitutions cause PACE mutations to confer NOXA activity only on later (or earlier) historical proteins. Arrow, gain or maintenance of binding. Blunt bar, loss of binding. Solid arrows, functional changes under PACE selection. Dashed lines, functional effects different from those selected for. Red, NOXA activity. Blue, BID activity. *Right:* inference of epistatic effects of historical substitutions from experiments in panels A and B. Mutations that confer selected functions in PACE are shown in the boxes at the end of solid arrows or bars. Dotted lines, introducing PACE mutations into other backgrounds causes different functional effects.

(D) Dissecting the effects of sets of mutations (white boxes) that caused hsMCL-1 to lose NOXA binding during four PACE trajectories. Filled boxes show the effect of introducing a subset of mutations into hsMCL-1 (normalized mean relative from three biological replicates). Solid lines show the effect of introducing v189G, which was found in all four sets. Dotted lines, effects of the other mutations in each set.

(E) Structural location of mutations in panel D. Alpha-carbon atom of mutated residues are shown as purple spheres on the structure of MCL-1 (light gray) bound to NOXA (red, PDB 2nla).

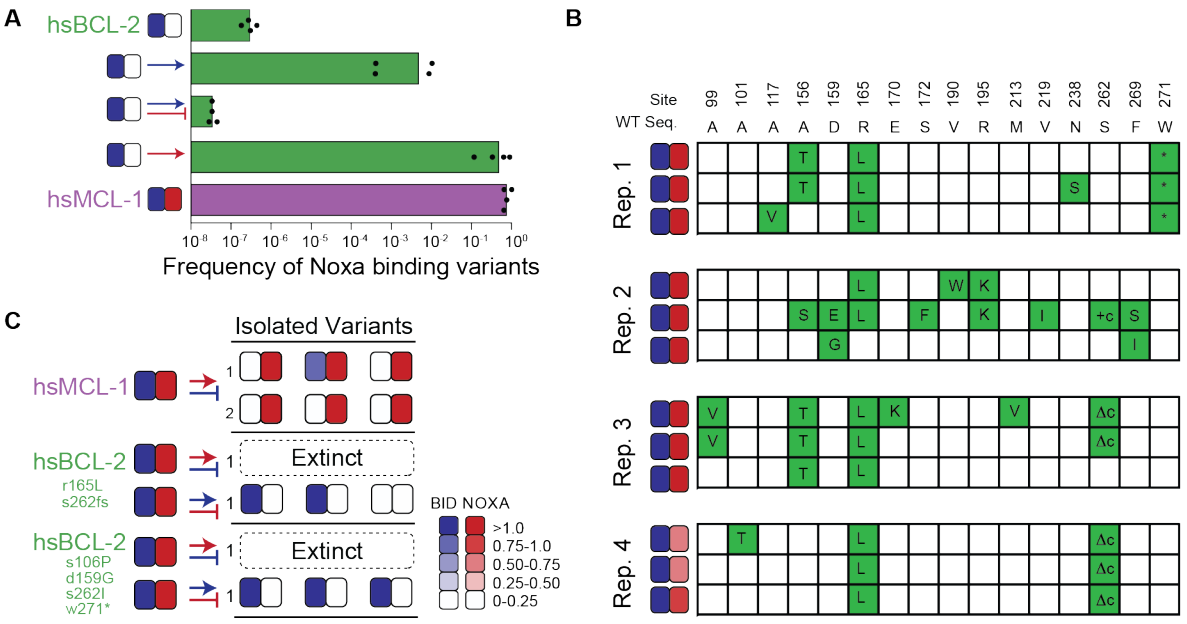


Figure 7. Evolutionary consequences of chance and contingency

(A) Evolution NOXA-binding phage from various selection regimes. Frequency was calculated as the ratio of plaque forming units (PFU) per milliliter on *E. coli* cells that require NOXA binding to the PFU on cells that require BID binding to form plaques. Wildtype hsBCL-2 (green) and hsMCL-1 (purple) are shown as controls. Arrow, positive selection for function. Bar, counter selection against function. Blue, BID. Red, NOXA. Columns show the mean of four trajectories under each condition (points).

(B) Phenotypes and genotypes of hsBCL-2 variants that evolved NOXA binding under selection only to maintain BID binding. Sites and WT amino state are indicated at top. For each variant, non-WT states acquired are shown in green. Heatmaps show binding to BID and NOXA in the luciferase assay for each variant (normalized mean of three biological replicates).

(C) Phenotypic outcomes when selecting for non-historical functions in PACE. Heatmaps show binding to BID and NOXA in the luciferase assay for each starting genotype (on the left) and for three individual variants picked from the trajectory's outcome. Each shaded box shows the normalized mean of three biological replicates.

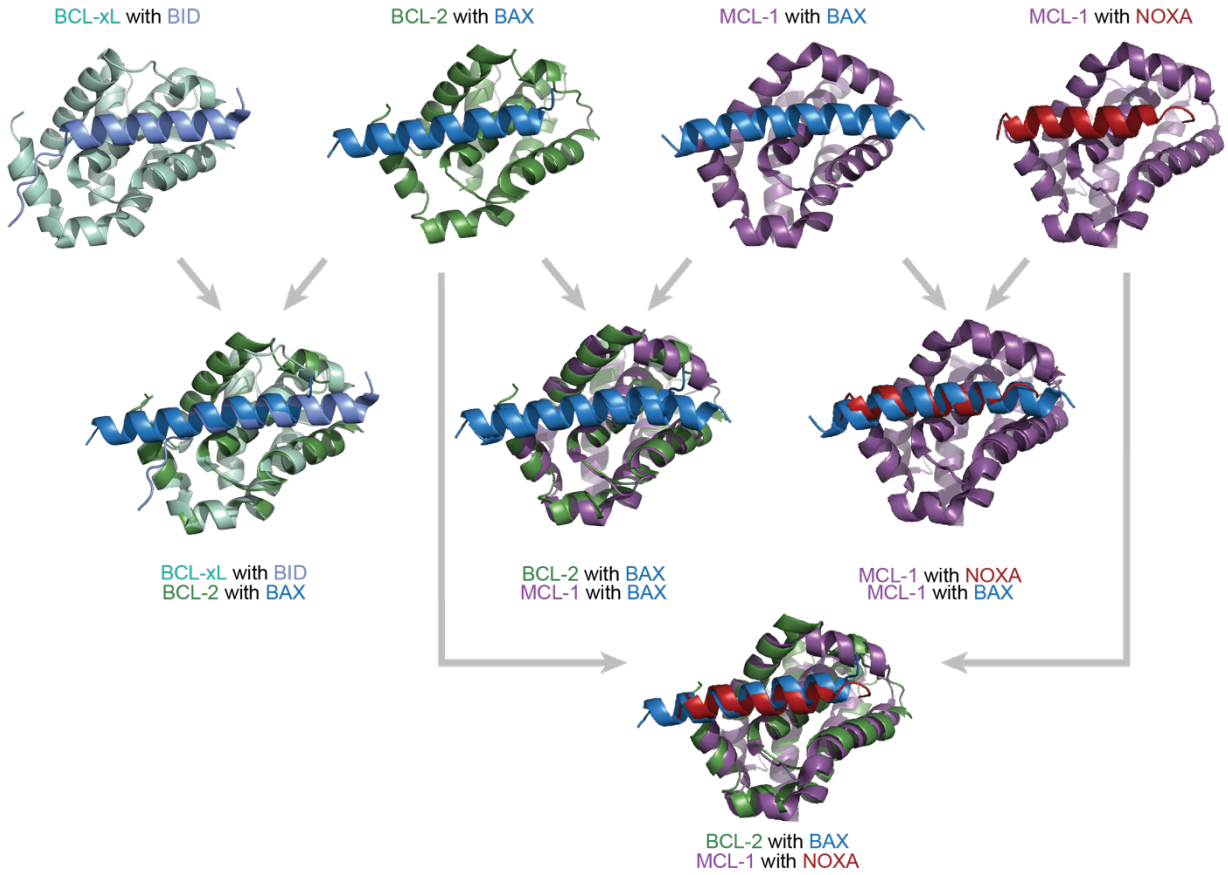


Figure S1. BCL-2 family proteins are structurally similar but have different binding profiles, Related to Figure 1

Crystal structures and overlays of BCL-xL (a vertebrate paralog of BCL-2, light green) bound to BID (light blue; PDB: 4qve); BCL-2 (green) bound to BAX (a protein with a BID-like binding profile, blue; PDB: 2xa0); MCL-1 (purple) bound to BAX (blue; PDB: 3pk1); and MCL-1 bound to NOXA (red; PDB: 2nla). The BCL-2 family proteins bind the coregulator proteins at the same interface.

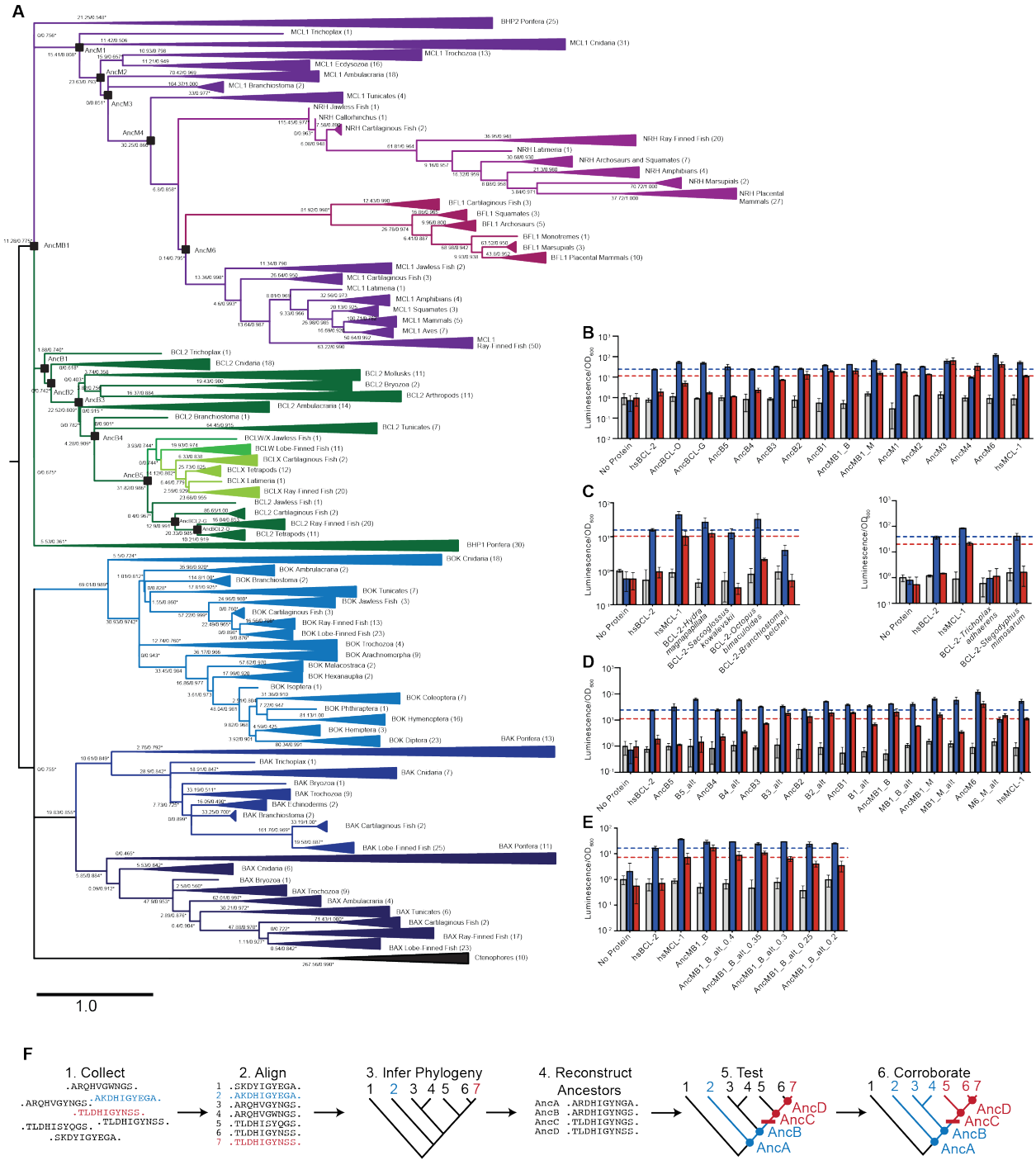


Figure S2. Phylogenetic reconstruction of the BCL-2 family proteins, Related to Figure 2

(A) Inferred phylogenetic relationships among BCL-2 family proteins. Green, BCL-2 class; purple, MCL-1 class; black, ctenophore sequences. Blue, pro-apoptotic paralogs. Shades within each color group indicate paralogs. Parentheses, number of sequences in each clade. Black squares, ancestral sequences reconstructed and tested. Node labels, approximate LRT statistics and transfer bootstrap values. Nodes with an asterisk were constrained to be congruent with known taxonomic relationships.

(B) Interactions of ancestral reconstructed proteins with BID (blue) and NOXA (red) in the luciferase assay, compared to no-coregulator control (gray). Activity is scaled relative to no-coregulator control with no-BCL-2 family protein. Columns and error bars, mean ± SD of three biological replicates. hsBCL-2 with BID (dashed blue line). hsMCL-1 with (dashed red line).

(C) Same as B, but for extant species *Hydra magnapapillata* (Cnidaria), *Octopus bimaculoides* (Lophotrochozoa), *Saccoglossus kowalevskii* (Hemichordata), *Branchiostoma belcheri* (Cephalochordata), *Trichoplax adhaerens* (Placozoa) and *Stegodyphus mimosarum* (Ecdysozoa).

(D) Same as B, but contains alternative reconstructions (Alt) for each ancestor which combine all plausible alternative amino acid states (PP>0.2) in a single “worst-case” alternative reconstruction.

(E) Same as B, but contains multiple alternative reconstructions for AncMB1_B. In each case, all plausible alternative amino acid states greater than the listed value are included in a single “worst-case” alternative reconstruction.

(F) Ancestral state reconstruction. 1) Sequences conferring different functions (red v. blue) are collected, as well as related sequences whose function may be unknown (black). These sequences can be orthologs from a variety a species, paralogs from gene duplication events, or a combination of orthologs and paralogs. 2) Sequences are aligned to identify homologous sites and poorly aligned regions are removed. 3) A phylogeny is inferred. Additional sequences are added to fill in missing taxonomic groups and to help break apart long branches. If clear discrepancies between known species relationships and the inferred phylogeny remain, the phylogeny may be constrained to minimize the inferred number of gene duplication or loss events. The last common ancestor of sequences conferring different functions cannot be the root of the phylogeny. Instead, an outgroup is required (e.g. sequence 1) based on prior information. 4) Using the inferred phylogeny, the aligned sequences, and a model of sequence evolution, the most likely state at each ancestral node is determined. 5) Ancestral sequences between the sequences with known functional differences are synthesized and tested for function. 6) The function of ancestral proteins proposes an evolutionary hypothesis about the branch in which function was altered during history (red bar). Corroboration for this hypothesis can be gained by determining the function of extant sequences which bracket the specific branch in question (e.g. sequences 3, 4, 5, and 6).

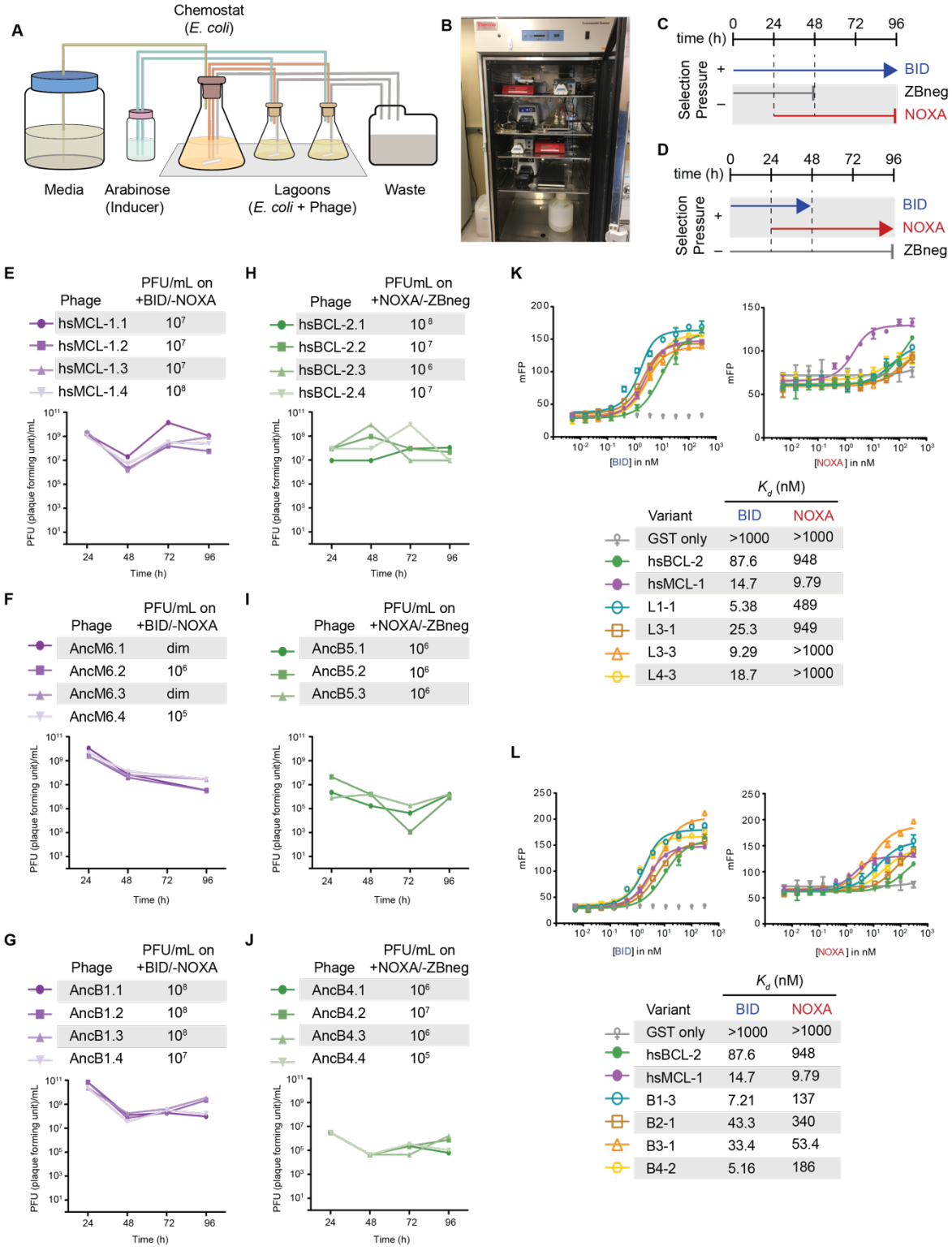


Figure S3. Using PACE to evolve target PPI binding specificity of BCL-2 family proteins, Related to Figure 3

(A) Schematic of a PACE experiment. Davis Rich carboy media flows into the chemostat, which contains *E. coli* with the positive selection (+AP), counter selection (-AP), and mutagenesis

plasmids (MP). The cells then flow into the lagoons, which contain phage with the evolving BCL-2 family protein. Arabinose is pumped into the lagoons to induce the mutagenesis plasmid in the *E. coli*. Both chemostats and lagoons are connected to the waste to maintain proper volume, cell density, and flow rate.

(B) Picture of representative PACE experiment from this work.

(C) Timeline of PACE experiments where hsMCL-1, AncM6, and AncB1 were evolved to lose NOXA binding. ZBneg is a control zipper peptide.

(D) Timeline of PACE experiments where hsBCL-2, AncB5, and AncB4 were evolved to gain NOXA binding.

(E) Phage titers (PFU/mL) over time (bottom) and activity-dependent phage titers at the end of the PACE experiments (top) where hsMCL-1 was evolved to lose NOXA binding. Activity-dependent plaque assays used plasmids 28-46 and Jin 487.

(F) Same as (E) for AncM6. “dim” means plaques were visible but weak, and therefore not quantifiable.

(G) Same as (E) for AncB1.

(H) Phage titers (PFU/mL) over time (bottom) and activity-dependent phage titers at the end of the PACE experiments (top) when hsBCL-2 was evolved to gain NOXA binding. Activity-dependent plaque assays used plasmids 28-48 and 29-39.

(I) Same as (H) for AncB5.

(J) Same as (H) for AncB4.

(K) Fluorescence polarization of hsMCL-1 variants evolved to lose NOXA binding. Bars are the mean of three replicates; error bars, SD. mFP, normalized measured fluorescent polarization.

(L) Fluorescence polarization of hsBCL-2 variants evolved to gain NOXA binding, same as (K).

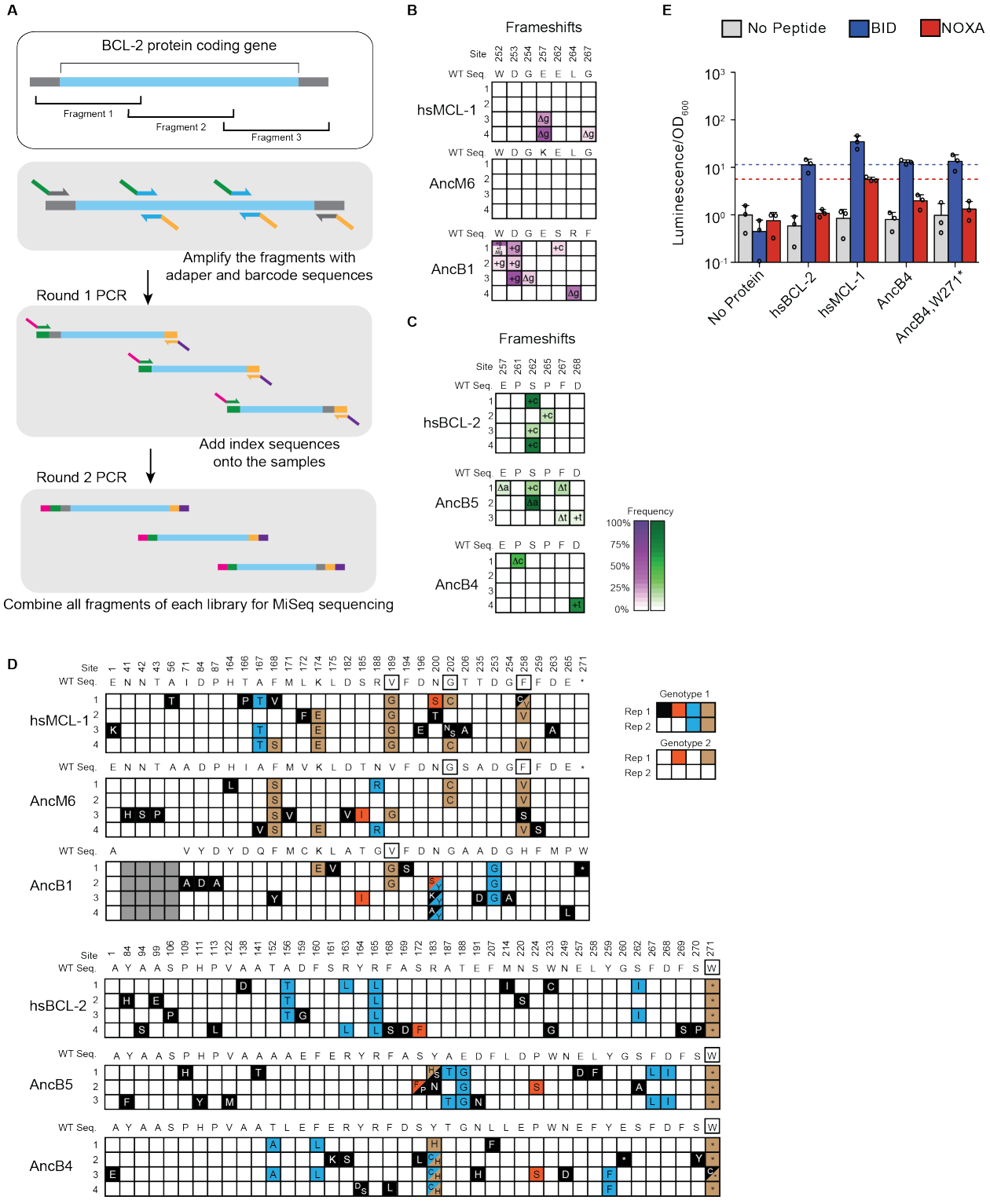


Figure S4. Allele frequency and mutational analysis of evolutionary outcomes, relate to Figure 4

(A) MiSeq library preparation. After isolation of phage DNA, the coding region of the evolving BCL-2 family protein was amplified in three overlapping fragments, each of which was smaller than 300bp. The DNA fragments were then amplified using sequence-specific primers. MiSeq

sequence adapters were added in a second PCR step. These fragmented DNA libraries were combined and used for MiSeq high throughput sequencing. Blue, target gene coding region. Gray, adjacent vector sequence. Green, forward adapter and barcode sequence. Orange, reverse adapter and barcode sequence. Magenta, index 1 sequence. Purple, index 2 sequence.

(B) Allele frequency data of frameshifts from replicate PACE experiments started from hsMCL-1, AncM6, and AncB1 evolved to lose NOXA binding. Site numbers and wildtype (WT) amino acid states are listed above each sequence. Each row represents an independent replicate population. Non-wildtype insertions and deletions that reached > 5% in frequency are shown, with frequency proportional to color saturation. Split cells show populations with multiple non-WT states > 5%. Plus (+) indicates an addition of a nucleotide. Delta (Δ) indicates a deletion of a nucleotide.

(C) Same as (B) for replicate PACE experiments for hsBCL-2, AncB5, and AncB4 evolved to gain NOXA binding.

(D) Categories of the 100 non-WT states observed for each non-WT state. Black box with white letters, mutant states observed in only one replicate. Teal, mutant states observed in multiple replicates from the same starting genotype. Orange, mutant states observed in a single replicate from multiple different starting genotypes. Brown, mutant states observed in multiple replicates from the same starting genotype and in at least one other replicate from a different starting genotype. Black box outline, mutant states observed in multiple replicates from the same starting genotype and from multiple replicates from a different starting genotype. Gray boxes are sites that do not exist in a particular sequence.

(E) Luciferase assay of the w271* mutation in AncB4. Activity is scaled relative to the no BCL-2 family protein with no coregulator peptide control. Bars are the mean ± SD of three biological replicates (circles). Gray bar, no coregulator peptide. Blue bar, BID. Red bar, NOXA. Blue dotted lines mark the average signal of hsBCL-2 with BID, and red dotted lines mark the average signal of hsMCL-1 with NOXA.

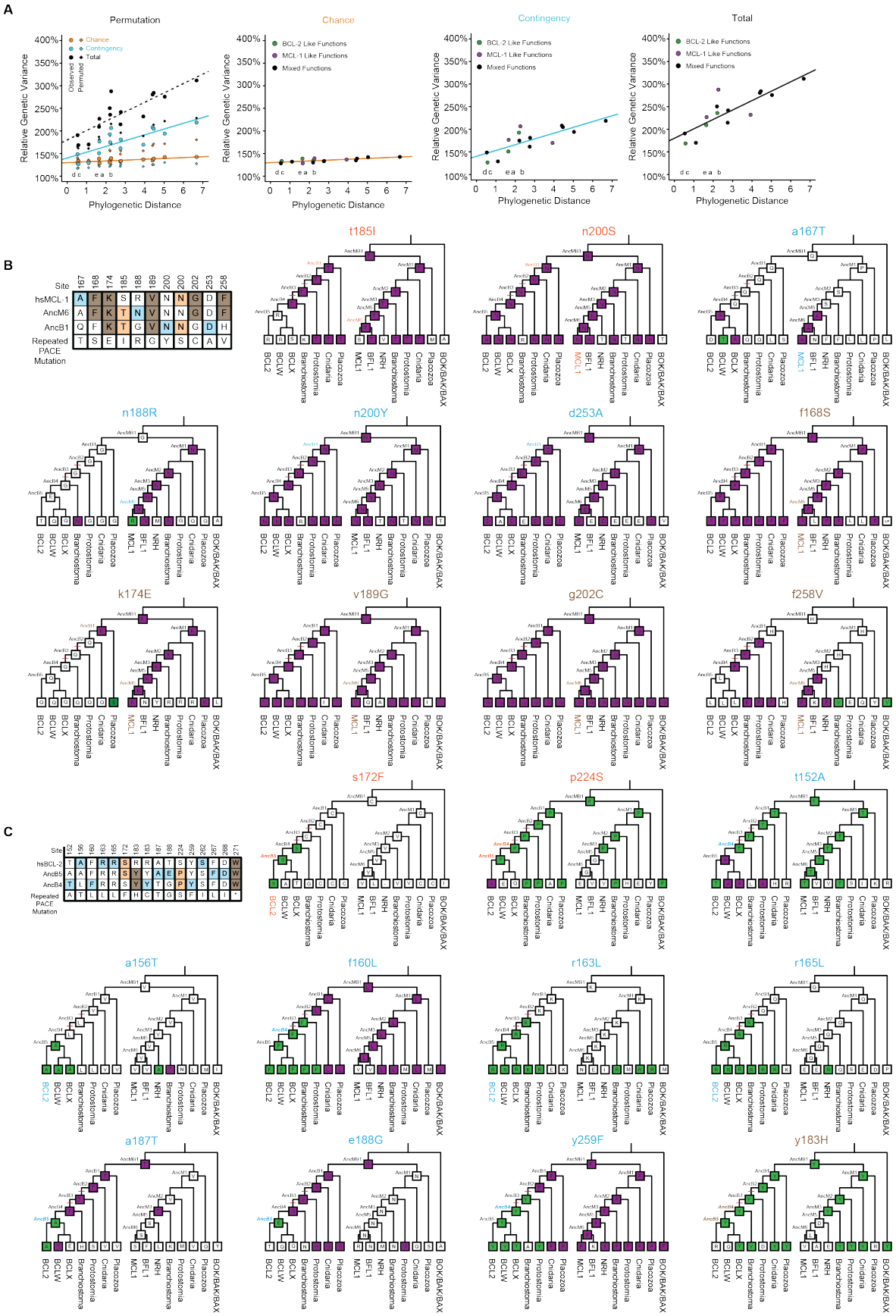


Figure S5. Chance and contingency in the repeated evolution of extant and ancestral BCL-2 family proteins, Related to Figure 5

(A) Change in chance and contingency over time. Relationship between phylogenetic distance between pairs of starting genotypes for experimental evolution (ancestral or extant proteins, as the total branch lengths separating them) and the effects of chance (orange), contingency (teal), or both (black) on the outcomes of evolution between them. Lines are best fits from linear models. Circles are observed values. Diamonds are averages of 1000 permutations of starting genotype labels. This shuffling of genotype labels results in more genetic variance among samples from the same 'starting genotype' than the observed data, and less genetic variance between samples from different 'starting genotypes' than the observed data. Letters indicate the specific branch from Figure 3D. Green, both starting genotypes had BCL-2 like function. Purple, both starting genotypes had MCL-1 like function. Black, starting genotypes differed in function.

(B) Phylogenetic distribution of repeated PACE mutations when selecting hsMCL-1, AncM6, and AncB1 against NOXA binding. Repeated mutation state and position are given above each cladogram. Lowercase letters, WT state. Uppercase letters, repeated mutant state. Top left**:** WT states for hsMCL-1, AncM6, and AncB1 at sites where repeated PACE mutations occurred. Repeated PACE mutation state is given in the last column. WT states are colored based on the pattern of repeated PACE mutations. Teal, sites that evolved the same mutation in multiple replicates from the same starting genotype. Orange, sites that evolved the same mutation in a single replicate from multiple starting genotypes. Brown, sites that evolved the same mutation in multiple replicates from the same starting genotype and from multiple different starting genotypes. Starting genotypes for PACE are colored similarly. Each cladogram shows the WT state for each node. Purple boxes; same WT state as the sequence in which the repeated PACE mutation emerged. Green boxes; same WT state as the repeated PACE mutation. Red bars indicate the interval in which NOXA binding was historically lost.

(C) Same as (B) when selecting hsBCL-2, AncB5, and AncB4 for NOXA binding. Green boxes; same WT state as the sequence in which the repeated PACE mutation emerged. Purple boxes; same WT state as the repeated PACE mutation.

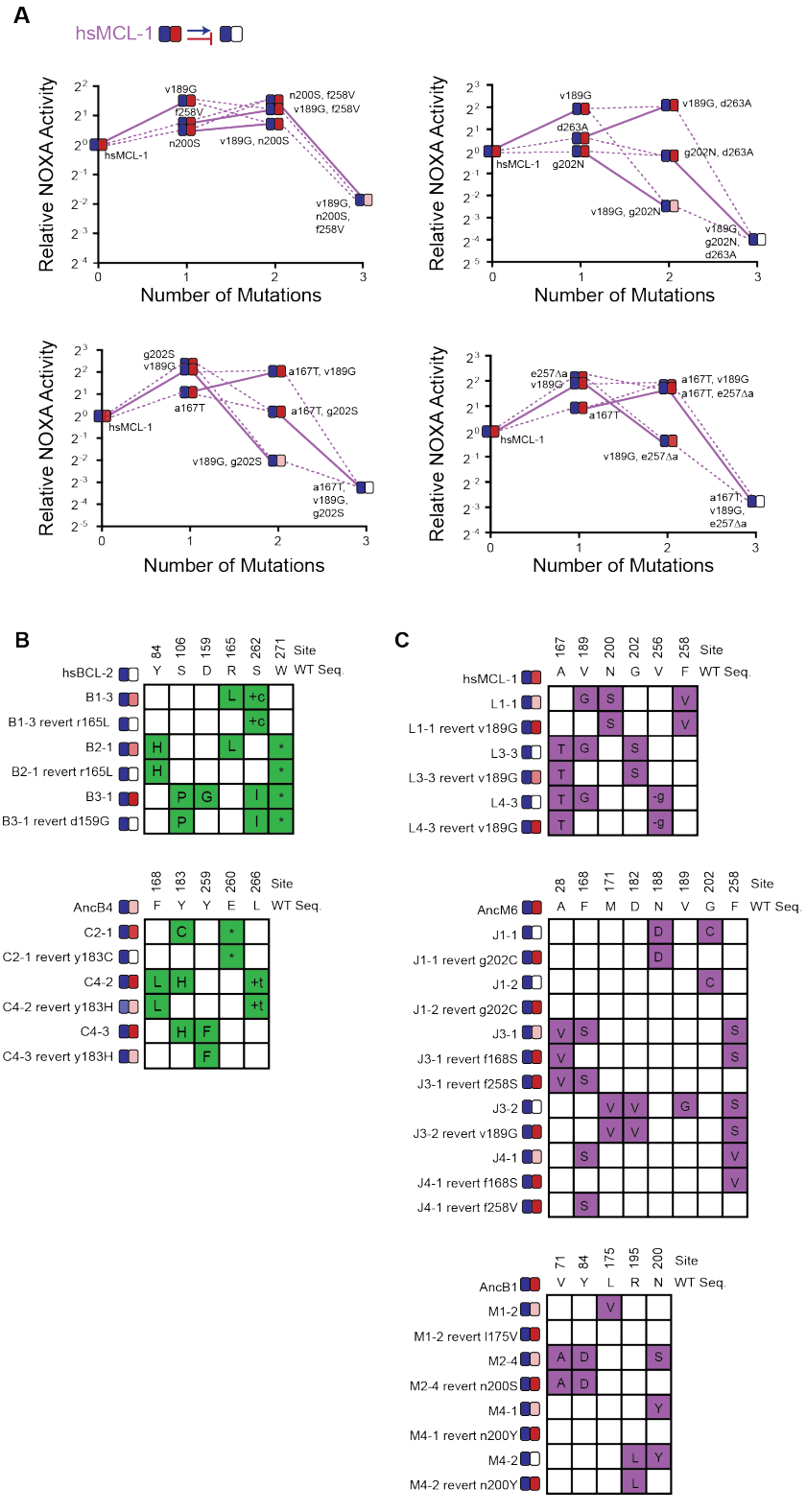


Figure S6. Effects of individual PACE-derived mutations, Related to Figure 6

(A) Effect on NOXA binding of mutations that occurred in PACE variants when hsMCL-1 was evolved to lose NOXA binding. Each panel shows NOXA binding (y-axis) for a unique variant as additional

mutations are added (x-axis). Values are the mean of three biological replicates. Heatmaps show the effects of each mutation on BID (blue) and NOXA (red) activity, and each shaded box represents the normalized mean of three biological replicates. Lines connect genotypes that differ by a single mutation. Solid lines show the effects of the v189G mutation. Dashed lines show the effects of all other mutations. Mutations come from variants L1-1 (top left), L3-1 (top right), L3-3 (bottom left), and L4-3 (bottom right).

(B) Phenotypic effects of reverting frequent PACE-derived mutations. Individual variants were isolated from PACE experiments that selected for the gain of NOXA binding in hsBCL-2 and AncB4. Sites and WT amino state are indicated at top. For each variant, non-WT states are shown in green. Heatmaps on the left show binding to BID and NOXA in the luciferase assay for each variant and their corresponding mutant without the key mutation. Each shaded box represents the normalized mean of three biological replicates.

(C) Same as (B), but for the loss of NOXA binding in hsMCL-1, AncM6, and AncB1. For each variant, non-WT states are shown in purple.

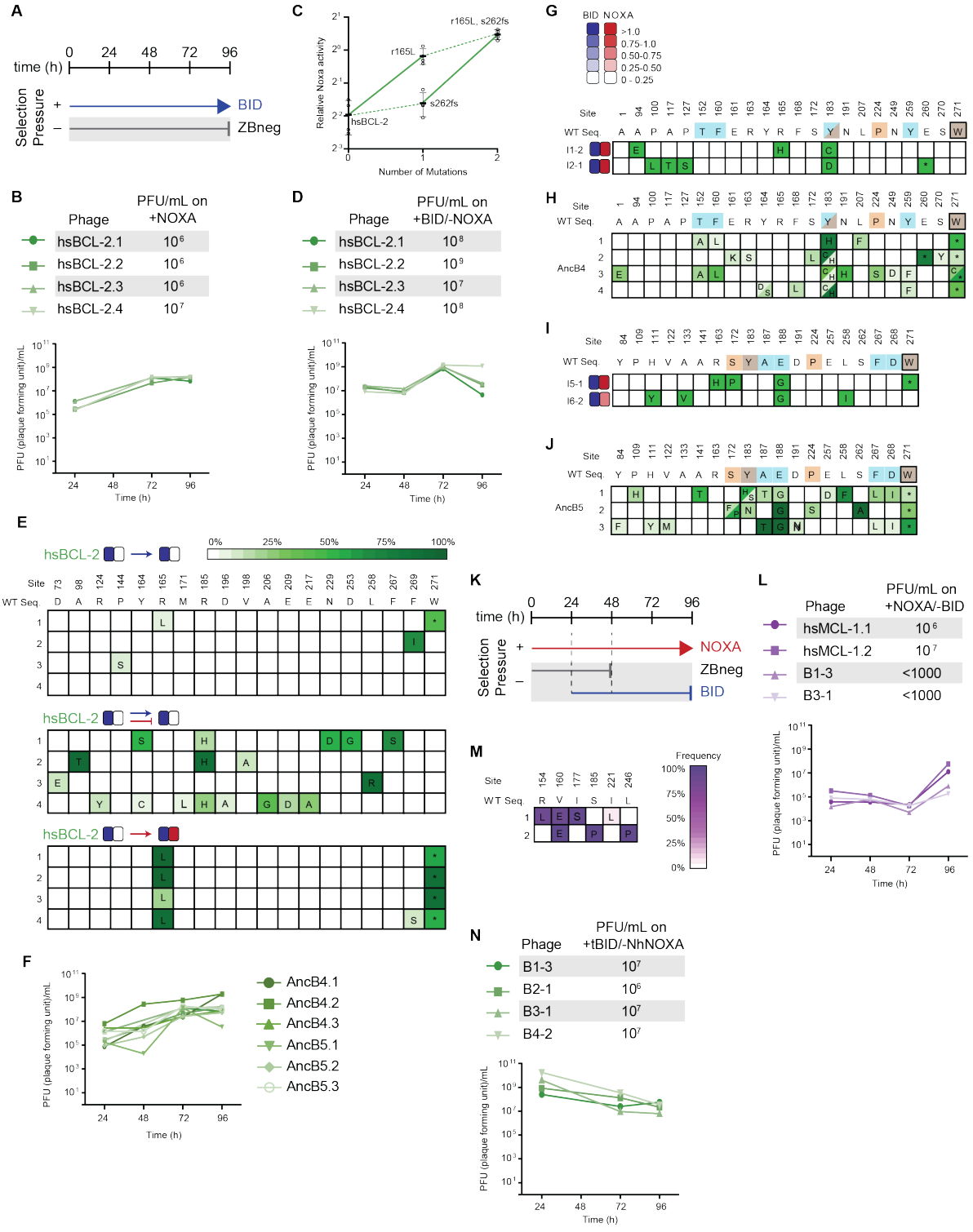


Figure S7. Spontaneous emergence of NOXA binding and selection for NOXA-specific binders, Related to Figure 7

(A) Timeline of PACE experiments where hsBCL-2, AncB5, and AncB4 were evolved with only positive selection to maintain BID binding. Selection conditions shown as arrows and blunt bars: arrow, selection for binding to BID (blue); blunt bar, selection against binding to ZBneg (gray).

(B) Phage titers (PFU/mL) over time (bottom) and activity-dependent phage titers at the end of the PACE experiment (top) where hsBCL-2 was evolved to maintain BID binding. Activity-dependent plaque assays used plasmid 28-48.
(C) Effect on NOXA binding of the key r165L mutation. Bars are the mean ± SD of three biological replicates (circles). Solid lines show the effects of the r165L mutation while dotted lines show the effect of a frameshift (fs) at site 262.
(D) Same as (B) where hsBCL-2 was evolved for binding and against NOXA binding. Activity-dependent plaque assays used plasmids 28-48 and Jin 487.
(E) Allele frequency of non-wildtype states when hsBCL-2 was evolved to maintain BID binding (top) or when hsBCL-2 was evolved to simultaneously maintain BID binding and lose NOXA binding (middle). For comparison, the same sites are also shown for when hsBCL-2 was evolved to gain NOXA binding (bottom). Site numbers and wildtype (WT) amino acid states are listed above each sequence. Each row represents an independent replicate population. Non-wildtype amino acids that reached > 5% in frequency are shown, with frequency proportional to color saturation.
(F) Phage titers (PFU/mL) over time from the PACE experiment where AncB4 and AncB5 were evolved to maintain BID binding.
(G) Phenotypes and genotypes of individual AncB4 variants that were isolated from PACE when selecting for BID binding and screened for the gain of NOXA binding. Site numbers and wildtype (WT) amino acid states are indicated at the top. Heatmaps on the left show binding to BID (blue) and NOXA (red) in the luciferase assay for each variant, and each shaded box represents the normalized mean of three biological replicates.
(H) Non-wildtype amino acid frameshifts that reached > 5% in frequency are shown for PACE where AncB4 was evolved to gain NOXA binding, for comparison with (F). Frequency is proportional to color saturation. Split cells show populations with multiple non-WT states > 5%. Each row represents an independent replicate lagoon. Color of WT state indicate if the mutation was seen among multiple replicates of the same starting genotype (teal), a single replicate from multiple starting genotypes (orange), or in multiple replicates and multiple starting genotypes (brown). Black box outline indicates mutant states observed in multiple replicates from the same starting genotype and from multiple replicates from a different starting genotype
(I) Same as (F) but for AncB5.
(J) Same as (G) but for AncB5 and for comparison with (I).
(K) Timeline of PACE experiments where hsMCL-1 and two previously-evolved NOXA-binding hsBCL-2 variants were evolved to maintain NOXA binding and lose BID binding. Selection conditions: arrow, selection for binding NOXA (red); blunt bar, selection against binding a specific peptide (BID (blue) or ZBneg (gray)).
(L) Phage titers (PFU/mL) over time (bottom) and activity-dependent phage titers at the end of the PACE experiment (top) where hsMCL-1 and NOXA-binding hsBCL-2 variants were evolved for binding NOXA and against BID. Activity-dependent plaque assays used plasmids 28-48 and Jin 518. Limit of detection = 10^3 PFU/mL.
(M) Allele frequency of non-wildtype states after hsMCL-1 was evolved to maintain NOXA binding and lose BID binding. Site numbers and wildtype (WT) amino acid states are listed above each sequence. Each row represents an independent replicate lagoon. Non-wildtype amino acid frameshifts that reached > 5% in frequency are shown, with frequency proportional to color saturation.
(N) Phage titers (PFU/mL) over time (bottom) and activity-dependent phage titers at the end of the PACE experiment (top) where NOXA-binding hsBCL-2 variants were evolved to lose NOXA binding. Activity-dependent plaque assays used plasmids 28-46 and Jin 487.

Table S1 (Excel File). Luciferase assay data for all experiments. Related to Figures 1, 2, 3, 6, 7, S4, S6, S7.

Table S2 (Excel File). Posterior probabilities for reconstructed ancestral sequences. Related to Figure 2. For each sequence, the site, maximum likelihood (ML) amino acid state, and posterior probability (PP) are given, along with the highest posterior probability alternative (ALT) state and posterior probability for this alternative state. Locations of paralog specific insertions are shown as gaps. For each reconstructed sequence, the average posterior probability for the maximum likelihood states and the alternative states is given, as are the number of sites where the posterior probability of a non-maximum likelihood state is greater than 0.2. Finally, the average, maximum, minimum, and variance among reconstructed ancestors is given for the average maximum likelihood posterior probability and the number of non-maximum likelihood states greater than 0.2 posterior probability.

Table S3 (Excel File). List of PACE experiments, amino acid alignments of hsBCL-2 and hsMCL-1 with their structural global alignment, and mutations found in individual variants isolated from PACE. fs is frameshift, aa is amino acid, co is codon change. Related to STAR Methods.

Table S4 (Excel file). PACE library and high-throughput sequencing (HTS) data. Related to STAR Methods. PACE experiments are listed in the tab "Library-info" which contains the name, purpose of the experiment, and HTS experiment numbers. The tab "Primers for HTS" lists all the primer sequences used for HTS library constructions. The tab "MiSeq reads number" include the read number of each library in this MiSeq run and the library sample information. The library samples are labeled as X*-end or X*-$$. "X" indicates the specific PACE experiment, "*" the experimental replicate, "end" means samples were collected after 96 hours when the experiment finished, and "$$" indicates the time point after removing chemostat A (e.g., "B2-24" is a sample from replicate 2 of evolution B and collected 24 hours after removing chemostat A, which is 72 hours from the start of PACE). The tab "genotype" includes the aligned protein sequences with corresponding residue numbers. The ‘Frequency’ tab contains the non-wildtype amino acid frequency of each sample for each site.

Table S5 (Excel file). Descriptions of plasmids and sequences used. Related to STAR Methods.