Phage-Assisted Evolution of Allosteric Protein Switches

Authors: Nicholas T. Southern[1], Anna von Bachmann[1,2], Alisa Hovsepyan[1], Marie-Louise Griebl[1], Benedict Wolf[1], Nina Lemmen[1], Ann-Sophie Kroell[1], Simon Westermann[1], Jan Mathony[1]*, Dominik Niopek[1]*

[1] Institute of Pharmacy and Molecular Biotechnology (IPMB), Faculty of Engineering Sciences, Heidelberg University, Heidelberg, 69120, Germany.
[2] Zuse School ELIZA, Darmstadt, 64289, Germany.

* To whom correspondence should be addressed:
dominik.niopek@uni-heidelberg.de;
jan.mathony@uni-heidelberg.de

ABSTRACT

Allostery, the transmission of locally induced conformational changes to distant functional sites, is a key mechanism for protein regulation. Artificial allosteric effectors enable remote manipulation of cell function; their engineering, however, is hampered by our limited understanding of allosteric residue networks. Here, we introduce a phage-assisted evolution platform for *in vivo* optimization of allosteric proteins. It applies opposing selection pressures to enhance activity and switchability of phage-encoded effectors and leverages retron-based recombineering to broadly explore fitness landscapes, covering point mutations, insertions and deletions. Applying our pipeline to the transcription factor AraC yielded optogenetic variants with light-controlled activity spanning ~1000-fold dynamic range. Long-read sequencing across selection cycles revealed adaptive trajectories and corresponding allosteric interactions. Our work facilitates phage-assisted evolution of allosteric proteins for programmable cellular control.

INTRODUCTION

Proteins are the principal engines of cellular function, executing dynamic tasks such as catalysis, signaling, and structural organization. A defining feature of many proteins is *allostery*, i.e. the ability to transmit local conformational changes across distant sites to modulate protein activity in response to environmental or cellular cues. This regulatory principle, termed the 'second secret of life' by Monod (*1, 2*), underpins diverse biological processes, from gene regulation to metabolism and sensory perception, and serves as a fundamental mechanism for information processing and signal transmission in cells.

Technologies based on synthetic allosteric proteins seek to harness this regulatory mechanism for remote control of biological functions (*3-6*). Optogenetics, for instance, repurposes naturally light-sensitive proteins to regulate processes such as neuronal action potential firing with precise spatial and temporal resolution (*7*). A more modular strategy involves inserting light-sensitive domains - such as light-oxygen-voltage (LOV) domains - into allosteric surface sites of unrelated effectors to render their activity input-dependent (*5*). While the transformative potential of “synthetic allostery” is widely appreciated, engineering allosteric proteins is challenging. The key difficulty lies in establishing reliable communication between sensor and effector domains so that input-dependent structural changes elicit the desired functional output. Advances in molecular modeling and machine learning have improved predictions of permissive sensor-effector fusion sites (*8*) but fall short of capturing the energetic couplings and conformational dynamics that govern the complex allosteric networks required for effective switching. These limitations highlight a core challenge of synthetic allostery: its governing principles are often emergent, encoded not in isolated residues but in distributed, nonlinear interactions.

Rather than solving this complexity through rational design, nature navigates it through variation and selection, thereby gradually tuning allosteric behavior by sampling and amplifying useful configurations across generations. When selection pressures change frequently as it is the case for environmental cues, cellular states, or signaling rhythms, the optimal activity of a protein depends on its current environmental context. Over evolutionary timescales, proteins respond to such pressures by acquiring new domains, accumulating mutations, and reorganizing structural elements to enable context-specific control (Fig. 1A). The evolutionary success of an organism can therefore hinge on the emergence of proteins that toggle between functional states in response to certain inputs. A canonical example is the circadian clock, which evolved independently across life forms to track the day-night cycle (*9*), often through light-dependent allostery (Fig. 1A and B). Similar regulatory architectures emerge in systems responsive to ligands (*10, 11*), redox state (*12*), or mechanical force (*13*), allowing proteins to adapt their activity to diverse and dynamic conditions. In such systems, evolution effectively see-saws proteins across distinct fitness landscapes, enriching mutations that optimize function under alternating conditions (Fig. 1B).

Taking inspiration from natural systems, we hypothesized that directed evolution under dynamic selection regimes could provide a powerful means to explore the complex fitness landscapes of allosteric protein behavior and uncover the context-dependent sequence-function relationships that lie beyond the reach of rational design (*14, 15*). To test this, we turned to phage-assisted evolution, a platform that links effector protein function to M13 phage propagation in *Escherichia coli* (*E. coli*) (*15, 16*). By pairing the rapid generation time of M13 phage with an immensely scalable, *in vivo* selection architecture, phage-assisted evolution has proven effective for solving a wide range of challenging protein engineering tasks (*17-21*). However, existing phage-assisted evolution systems focus on optimizing a selected activity under more or less static conditions, rendering them ill-suited for evolving proteins that must switch between distinct functional states in response a defined input.

Here we present POGO-PANCE (Protein-switch On/Off Gene Optimization using Phage-Assisted Non-Continuous Evolution), a directed evolution strategy inspired by the emergence of allosteric regulation in nature (Fig. 1A). POGO-PANCE cyclically "bounces" evolving proteins between alternating selection steps corresponding to the presence and absence of a defined input signal, thereby co-navigating two alternating sequence fitness landscapes simultaneously (Fig. 1B). This design enforces the emergence of effector protein variants that are both highly active and responsive to a chosen input.

A critical determinant of directed evolution success is the diversity of the genetic pool subjected to selection. In standard phage-assisted evolution systems, mutagenesis plasmids (MPs) increase point mutation rates by perturbing DNA replication and repair (*22*). However, they introduce strong mutational biases and cannot generate insertions or deletions (InDels), significantly limiting the accessible protein sequence space. We therefore developed RAMPhaGE (*Recombitron-Assisted Multiplexed Phage Gene Evolution*), a retron-based *in vivo* diversification method capable of introducing point mutations and InDels at tunable frequencies, thus unlocking regions of the protein sequence space previously inaccessible to phage assisted evolution.

To demonstrate our novel concepts, we applied POGO-PANCE and RAMPhaGE to the arabinose-responsive transcription factor AraC, a well-characterized bacterial effector protein. Through iterative evolution cycles, we obtained light-responsive AraC variants with almost binary switch-like behavior and dynamic control spanning over three orders of magnitude. Tracking evolving phage pools via short- and long-read sequencing revealed key adaptive trajectories and pinpointed structural features and epistatic interactions that underlie effective allosteric switching. Together, POGO-PANCE and RAMPhaGE not only enable evolution of programmable molecular switches, but also open a window into the emergence and architecture of allosteric networks.

RESULTS

Conceptualization of POGO-PANCE to evolve allosteric switches

To mimic natural dynamic selection regimes as required to evolve switchable proteins, POGO-PANCE alternates between temporally separated positive and negative selection steps, each aligned to the presence or absence of a defined input, such as light (Fig. 1C). In positive selection, as in conventional Phage-Assisted Non-Continuous Evolution (PANCE), effector activity promotes expression of M13 gene III (gIII), encoding an essential phage coat protein (pIII), from an accessory plasmid (AP) in trans, thereby enabling phage propagation (*15*). In contrast, during negative selection, the same regulatory circuit is inverted such that effector protein activity drives expression of a dominant-negative gIII variant (gIII-Neg), which is incorporated into assembling phage particles and blocks their release from the host (*23*). To enable selective propagation of inactive variants during this step, wild-type (WT) gIII is provided at low levels from a separate AP. Because the input states and circuit logic are mutually exclusive, selections are performed in batch culture sequentially, where environmental conditions can be precisely toggled. As input, POGO-PANCE requires phages encoding an effector with at least minimal stimulus responsiveness, e.g. from rational design or screening. For evolution of trigger-activated effectors, the input is present during positive selection and absent during negative selection; for trigger-inhibited effectors, the inverse applies. By iteratively applying diversification (e.g. via mutagenesis plasmids), negative and positive selection, each time toggling the input accordingly, POGO-PANCE recapitulates dynamic evolutionary pressures analogous to those found in nature, such as the day-night cycles that shaped circadian clocks (Fig. 1A and D).

Evolution of highly active AraC variants that are independent of arabinose

Having conceptualized the POGO-PANCE approach, we applied it to the evolution of optogenetic effector proteins. Light, as an input, offers superior spatiotemporal precision compared to chemical control, and it can be readily added or removed in a PANCE setup. As candidate effector, we selected the *E. coli* transcription factor AraC, a canonical genetic switch widely used in microbiology for L-arabinose inducible gene expression via the pBAD promoter. AraC has a well-characterized dual role as a genetic regulator, switching from a repressive to an activating state in response to its native ligand, L-arabinose (fig S1A), and is known for its high evolvability, capable of adapting to novel inducers or acquiring constitutive activity (*24, 25*). Given that previous studies have demonstrated its ability to accommodate light-responsive domain insertions at specific sites (*26, 27*), AraC appears as an ideal candidate to explore the POGO-PANCE concept. However, because its activity depends on the presence of arabinose, we first sought to eliminate this dependency in order to reprogram its regulatory logic for non-native input.

To this end, we designed a PANCE circuit that separates evolution into distinct mutagenesis and selection phases. For mutagenesis, we adapted the previously published drift plasmid 6 (DP6; (*23*)) for IPTG-inducible mutagenesis (*28*) and anhydrotetracycline-dependent gIII expression

(Fig. 2A). For positive selection, we constructed an AP encoding gIII under the control of the pBAD promoter and confirmed that it enables propagation of active AraC variant–encoding phages, supporting approximately 20-fold higher propagation in the presence versus absence of arabinose (Fig. 2B and fig. S1B).

Next, to evolve constitutive AraC, we applied two distinct selection regimes (R) across three independent evolution pools each to promote diverse adaptive trajectories. In one condition set, phages were selected entirely in the absence of arabinose (R1-R3), while in the other, arabinose levels were incrementally reduced over five passages (100 µM to 0 µM, R4-R6) (Fig. 2C and Table S2). To maintain selective pressure without population saturation, we imposed sequential bottlenecks between mutagenesis and selection by controlling input phage levels (see Methods).

Over five rounds of mutagenesis and selection, phage output titers remained high (fig. S1C) throughout the experiment, with progressively increased propagation observed after each round (fig. S1D). This suggested the emergence of strong constitutive activity. Phage pool sequencing, i.e. sanger-sequencing of the AraC PCR amplicons derived from the evolved phage pool, indicated diverse evolutionary lineages (Fig. 2D). We selected one representative variant from each of the six pools (R1-R6) for further testing using a plasmid-based reporter assay *in E. coli*, in which AraC activates a pBAD-driven red fluorescent protein (Fig. 2D, fig. S1E). To our surprise, overnight cultures not induced with arabinose or IPTG (which controls AraC expression via a LacI-regulated promoter) appeared in dark red for five of the six AraC variants (Fig. 2E). This striking phenotype indicated high constitutive activity, even at low expression levels driven solely by leaky transcription from the lac promoter.

To facilitate quantitative comparisons between evolved variants, we introduced a substantially weaker ribosome binding site (RBS) upstream of AraC (*29*). Under these conditions, all six evolved variants induced higher reporter levels than WT AraC in the presence of arabinose (Fig. 2F). In line with the observed red culture coloring (Fig. 2E), five of the six variants exhibited constitutive activity in the absence of arabinose (Fig. 2F). Moreover, both R2 and R5 even surpassed the output of fully induced WT AraC, indicating that accumulated mutations conferred potent ligand-independent function (Fig. 2F, G and fig. S2).

To understand the molecular basis of the observed activity gains, we examined mutations in the evolved AraC variants based on its structure (*30*). Recurrent substitutions at A152 and R251, jointly present in the potent R2 variant (A152T, R251L (*24*)), have been previously linked to enhanced transcriptional activity, likely by destabilizing inhibitory interdomain contacts and shifting the conformational equilibrium toward the active state. Several mutations also mapped to the N-terminal arm of AraC (e.g., A2E, E3K, A4T, S14I), a region known to enforce repression in the absence of arabinose (*25*). Disruption of this arm has similarly been associated with constitutive AraC activity, suggesting that both mutation classes act through complementary mechanisms that weaken repressive intramolecular interactions. In R5, the L156I and R251H

combination appeared to form an alternative strong activator; although L156 was not recurrent across variants, it has previously been linked to altered ligand specificity, possibly tuning allosteric responsiveness (*31*). Together, these mutational patterns highlight diverse, mechanistically convergent routes to constitutive activation and underscore the power of experimental evolution to uncover functionally optimized protein states.

POGO-PANCE yields potent optogenetic switches and probes allosteric networks

Having created particularly strong, constitutive AraC variants as evolutionary stepping stones, we next assembled circuits to implement the full POGO-PANCE pipeline for evolving light-switchable AraC variants (Fig. 1C and D). We constructed an accessory plasmid (AP_{Neg}) in which a dominant-negative gIII variant (gIII-neg) is expressed under control of the pBAD promoter (Fig. 2H), and co-transformed this plasmid with AP_{Prop} (Accessory Plasmid for Propagation), a LacI-IPTG controlled circuit for tunable gIII expression, into S2060 (fig. S3A) (*14, 32*). This configuration enables tunable negative selection against AraC activity, with IPTG induction modulating basal phage propagation rates.

To validate the negative selection host cells, we ran experiments in which phages encoding WT AraC or the evolved, hyperactive R2 or R5 variants were propagated under mild IPTG induction (25 µM) and in absence or presence of arabinose (Fig. 2H, fig. S3B). As expected, WT AraC phages showed a clear arabinose-mediated propagation deficit, as did the R5 variant. R2-encoding phages showed minimal propagation on negative selection hosts (fig. S3B), even in the absence of arabinose, consistent with the very strong constitutive activity conferred by this AraC variant (Fig. 2F). Finally, mixing WT AraC encoding phage with R2 or R5 phage in equal quantities resulted in arabinose-dependent propagation again (fig. S3B). Importantly, sequencing five single phage plagues from the mixed WT AraC:R5 group without arabinose revealed that all were of WT AraC background (Fig. 2I, fig. S3C), indicating strong selection against constitutively active variants. These results confirm that the negative selection regime effectively enriches for phages encoding inactive AraC.

To explore whether evolved AraC variants could serve as a foundation for light-switchable gene expression control, we inserted the blue light-responsive LOV2 domain from *Avena sativa* phototropin-1 (*5*) into each of the six variants at a previously established allosteric site immediately following residue S170 (*26*), flanked by SG and GS linkers (fig. S4A). Light-induced unfolding of LOV2 terminal α-helices is a well-established mechanism, commonly harnessed in synthetic allostery for generating local disorder and conveying an allosteric signal (*33*). However, existing AraC-LOV hybrids have shown limited dynamic range and compromised activity (*26, 27*). We first evaluated the newly generated R2-LOV and R5-LOV constructs by quantifying phage propagation under dark and blue light conditions (fig. S4B-E). However, these phages showed minimal or no light-dependent propagation and markedly reduced propagation overall, suggesting that LOV2 insertion disrupted AraC function or failed to confer robust photoswitching under the selection conditions.

We therefore returned to assessing all R1-R6 LOV hybrid variants in a plasmid-based mRFP reporter assay using the same expression conditions as the original constructs. Under uninduced conditions (no IPTG), none of the AraC-LOV variants produced measurable reporter output in either light or dark, indicating that LOV2 insertion severely impaired Arabinose-independent transcriptional activation from the pBAD promoter (compare Fig. 2F and fig. S4F). Heavily overexpressing the constructs with 400 μM IPTG, however, restored detectable reporter activity in the dark for four of the six variants (R1, R2, R4 and R5 LOV hybrids), with a consistent reduction in activity upon blue light exposure (fig. S4G). While the switching dynamic range was limited, these results indicated that residual light-dependency could be uncovered at elevated expression levels. Given that the R2-LOV and R5-LOV variants have non-overlapping AraC mutational backgrounds and were, in principle, capable of propagating in the phage context, we selected these variants as starting points for optimization using the POGO-PANCE approach.

We next evolved both variants using three distinct selection strategies, each run in duplicate pools (Fig. 3A and fig. S5). The first POGO-PANCE regime (P1) consisted of three cycles of the following steps: (1) mutagenesis, (2) negative selection under blue light to eliminate leaky or constitutive variants, enriching those with low activity in the presence of the trigger, and (3) positive selection to favor variants that exhibit potent activity in the dark, thus promoting switch-like behavior. The second regime (P2) began with two cycles of (1) mutagenesis and (2) positive selection to maximize activity, followed by two cycles of (1) mutagenesis, (2) negative selection in presence of light and (3) positive selection in absence of light. The third regime (P3) was identical to the first, but reversed the light conditions during selection and served as a control with non-aligned input. Between steps, phages were diluted ~1:100, to preserve gene pool diversity (see Methods). Of note, a recombination event overtook one R2 pool in condition P3 and one R5 pool in conditions P2 and P3, eliminating these three pools from further consideration (see Discussion).

Upon completion of the evolution regimes, we assessed the mutational landscape at the end-point for the nine remaining pools (see Methods for details; note that for LOV2, residue numbers correspond to the full-length phototropin-1). While mutations were rather broadly distributed across the AraC-LOV fusion protein (fig. S6), indicating the LOV2 and AraC domain parts co-evolved, AraC residues E144, R146, V223 and S225 as well as LOV2 residues, D501, E541 and A543 emerged as recurrent mutational hot-spots (Fig. 3B), with specific combinations favorably co-occurring (fig. S7). Importantly, apart from mutations of residue 146 in R5-LOV, which was associated with increased constitutive activity in the constitutive R2 (compare Fig. 2D and 3B), none of the observed mutations in AraC-LOV variants had occurred in the prior PANCE evolution in any pool. This mutational divergence implies distinct, context-dependent mechanisms, potentially enhancing domain insertion compatibility or enabling light-dependent allosteric switching.

Regarding the newly mutated sites in AraC, E144 and R146 are positioned near the arabinose-binding pocket and are directly involved in ligand recognition (*30*), making them key nodes in AraC's allosteric network. V223 and S225 lie in the C-terminal DNA-binding domain, specifically within the second helix-turn-helix motif that recognizes the I1 and I2 half-sites of the pBAD promoter and were previously shown to lie directly adjacent to a highly dynamic AraC region (*34*). In LOV2 (*35-40*),. D501 is located near the FMN chromophore and participates in the hydrogen-bonding network (*35*) that stabilizes the chromophore binding pocket. E541 and A543 are situated in the C-terminal Jα helix, a key structural element that undocks from the LOV2 core upon light activation (*40*). Also, in conjunction with residue serine in the right linker (SRL2) at the AraC-LOV boundary, E541 and A543 are part of a key junction between the AraC and LOV protein parts.

To dissect the evolutionary trajectories underlying switch optimization, we performed long-read sequencing (*41*) across all mutagenesis and selection steps of one evolution pool (P1.2, R2-LOV variant; Fig. 3C). Distinct mutational patterns emerged by the second cycle; Across cycles 2 and 3, mutations were differentially enriched or depleted depending on the selection step, underscoring the complementary roles of opposing selection pressures in shaping sequence space. Mapping mutation frequencies after each evolution cycle onto the AraC-LOV structure revealed a variety of mutational hotspots scattered across both domains and affecting all types of secondary structure elements (Fig 3D). Together, these observations suggest a distributed allosteric network, which coordinates signal transmission from the LOV2 photosensor to AraC's functional domains.

Next, we isolated four phages from each pool and cloned the corresponding AraC-LOV fusion variants into expression vectors. Sanger sequencing revealed a total of 28 genetically unique variants (Fig. 4A), with mutational patterns aligning well with the long-read sequencing data (compare Fig. 3B and 4A). Interestingly, several mutations preferentially co-occurred, albeit located far-apart in the AraC structure, e.g. E144K and V223M (Fig. 4A and 3D). Vice versa, several mutations were mutually exclusive, e.g. V223M and S225R, suggesting distinct sequence fitness maxima.

Subsequent plate reader assays of RFP plasmid reporter activity showed that the majority of R2-LOV and R5-LOV variants derived from P1 and P2 pools exhibited strong dark-state activity and, in several cases, highly potent response to blue light, far exceeding that of the non-evolved benchmarks (Fig. 4B). Closer investigation of two particularly promising R2-LOV variants, namely clone 3 from each, the P1.1 and P2.1 pools, by flow cytometry confirmed high dark-state activity and exceptionally strong photoswitching, with median fluorescence shifting by about three orders of magnitude upon illumination (Fig. 4C). This indicated that the mutations that emerged during POGO-PANCE substantially improved allosteric switching.
In contrast, the control evolution condition (P3) predominantly yielded either recombination events (cheaters, see above) or variants with abolished or constitutive activity, underscoring that

aligning the light input with the appropriate selection step, i.e. negative selection for off-switches, was critical for successful evolution (Fig. 4A,B).

Collectively, these data indicate that POGO-PANCE is a highly effective approach for evolving optogenetic proteins with exceptional performance. Beyond allosteric switch engineering, POGO-PANCE also provides a powerful lens to observe and interrogate allosteric networks by evolving them in real time, revealing how effective communication and signal relay emerges across protein domains from a network of seemingly unrelated residues.

RAMPhaGE: A highly effective, targeted gene diversification strategy for PANCE

Mutagenesis plasmids such as MP6 and DP6 are widely used to diversify transgenes and generate gene pools for phage-assisted evolution. However, this diversification strategy presents significant limitations that constrain access to broad protein sequence landscapes. First and foremost, mutagenesis plasmids are restricted to point mutations and cannot introduce InDels. However, InDels may be particularly important in the context of allosteric switch engineering, where fine-tuning lengths, relative orientation of, and distances between (communicating) structural elements is likely critical (*42*). Second, the mutations MPs/DPs generate are highly biased toward specific base changes (*22*), thus sampling only a narrow portion of the sequence landscape.

To overcome these limitations, we set out to develop a versatile *in vivo* transgene diversification strategy based on retron-mediated recombineering (*43*), a genome editing method capable of introducing targeted substitutions, insertions, and deletions (*44*). Retrons encode an RNA template and a reverse transcriptase that together produce intracellular single-stranded DNA (ssDNA), which serves as a recombination donor during replication (*43*). In contrast to conventional recombineering, retrons generate these donors autonomously and continuously from a plasmid, eliminating the need for external ssDNA supply. To turn retrons into a targeted *in vivo* diversification strategy of M13 phage carried transgenes, we used a Recombitron: a fully genetically encoded operon combining a retron, single-stranded annealing protein (CspRecT), and *mutL*, a dominant-negative variant of the *E. coli* mismatch repair suppression protein *mutS* (Fig. 5A).

While Recombitrons have been shown to mediate efficient genome editing in dsDNA phages, their compatibility with ssDNA phages such as M13 has not been demonstrated (*43*). Thus, to validate this concept, we created an IPTG-inducible Recombitron carrying a 90-nt Eco1 retron encoding a silent single-nucleotide substitution in the AraC-LOV transgene. We then expressed the Recombitron in *E. coli* S2060 and propagated AraC-LOV encoding M13 phages on these hosts overnight (Fig. 5A). Stunningly, pooled sanger sequencing of an AraC-LOV PCR amplicon derived from ~10,000 phages revealed precise, 100% editing of the target site (Fig. 5B and fig. S8). In contrast, no editing was observed in controls carrying either (i) a non-targeting retron, (ii) the Recombitron with the reverse transcriptase (RT) disabled by mutation or (iii) a

functional Recombitron with a retron encoding the WT variant (Fig. 5B and fig. S8), indicating that editing was dependent on the presence of all key Recombitron components. Of note, we observed that phage titers in the targeting retron group remained at input levels (fig. S8A). This effect was not seen in non-targeting retrons with active reverse transcriptase, indicating that neither ssDNA production nor metabolic burden alone explained the reduced phage propagation.

To assess compatibility of retron editing with library-scale evolution, we repeated the above experiment, this time using 10^{10} pfu/mL input phage titer, and found >70% editing efficiency at the targeted site as revealed by long-read sequencing (fig. S8C). These findings demonstrate that while Recombitron induction can transiently constrain phage amplification, this issue is easily overcome by increasing input phage titers.

To realize the potential of retron-based editing for phage assisted evolution workflows, we next adapted the Recombitron system for the generation of complex, targeted mutational libraries. We first engineered a Golden Gate-compatible recombineering plasmid (RP) featuring a ccdB dropout cassette, enabling efficient insertion of retron-encoding DNA libraries. To facilitate library design, we developed a custom software tool that automates the generation of barcoded retron templates, enabling direct oligo pool synthesis with precise control over mutation identity and distribution across M13-encoded transgenes (fig. S9 and Methods). Using this framework (Fig. 5C), we constructed two specialized libraries. The first, a deep mutational scanning (DMS) library, encoded all (2,820) possible single amino acid substitutions across the LOV2 domain in AraC-LOV using *E. coli* preferred codons for uniform representation. This provides a direct comparison to DP6-based “random” point mutagenesis. The second library mediated insertions, deletions and residue alterations at the sensor-effector junctions of AraC-LOV fusions (“linker library”), thus highly diversifying the elements particularly critical for allosteric signal relay.

To assess editing efficiency and mutational outcomes, we propagated AraC-LOV-encoding phages overnight on the DMS library hosts followed by short- and long-read deep sequencing. As comparison, we also propagated the same phage on pLlacO1-DP6-transformed cells. Retron-based editing produced high frequency of programmed mutations, with >40% of reads containing programmed DMS library mutations overall (Fig. 5D). Further analysis revealed that the average mutation rate per retron-targeted position was ~0.4% - more than triple of that achieved with pLlacO1-DP6 (fig. S10A).

In stark contrast to the highly biased mutational profile generated by pLlacO1-DP6, the DMS library yielded broad and relatively uniform coverage across the full spectrum of encoded amino acid and codon substitutions (Fig. 5E, fig. S10B and C). Variant frequencies in the DMS phage library closely matched those in the retron-encoding plasmid library (fig. S10D; Spearman r = 0.77; Fig. 5F, top). Because codon substitutions require between one and three mutations, we expected a dependence of editing efficiency on substitution size. As expected, single-base codon edits occurred at approximately twice the rate of double or triple substitutions, consistent with

known recombination biases favoring smaller InDels (*43*) (fig. S10E). Off-target mutation rates outside the retron-targeted LOV2 region were minimal, as confirmed by short-read sequencing at the AraC-LOV2 junction and long-read analysis of the full transgene (Fig. 5F and fig. S11).

Excitingly, the linker library introduced not only a diverse spectrum of amino acid substitutions but also a broad range of insertions and deletions that span up to five amino acids - demonstrating that even >15 bp changes within the 90 nt retron payload could be efficiently installed (fig. S12A). Variant frequencies in plasmid input and phage output libraries were again correlated (fig. S12B; Spearman r of 0.79 and 0.41 for the left and right AraC-LOV junction), and editing rates depended on the encoded InDel size as expected (fig. S12A). Interestingly, after a single round of overnight propagation on Recombitron-expressing hosts, mutations were typically confined to one junction (either left or right), while double-mutants were rare (fig. S12C). This is consistent with the biology of M13 infection (*45*); at the chosen input levels, host cells are likely infected immediately, rendering them immune to subsequent infection by newly edited phage.

That said, retron libraries enable sequential, layered mutagenesis on demand. To demonstrate this, we first introduced DMS edits within the LOV2 coding region by overnight phage propagation on host cells expressing the DMS retron library. The resulting phage progeny were then transferred to hosts carrying the InDel library to diversify the AraC-LOV junctions. This two-step protocol yielded cumulative editing, with widespread point mutations across LOV2 and substantial linker diversification at the insertion boundaries after propagation on both host libraries (fig. S12D and Fig. 5D).

We refer to this strategy as RAMPhaGE (Recombitron-Assisted Multiplexed Phage Gene Evolution), an M13-based transgene diversification platform that merges the precision of retron-guided editing with the scalability of *in vivo* phage-assisted evolution.

Integrating RAMPhaGE and POGO-PANCE to evolve sensor-effector junctions for enhanced allosteric coupling

Sensor-effector junctions are key determinants of allosteric communication and switch fidelity in chimeric proteins (*46*). It is thus no surprise that these regions were also prime targets for evolution in POGO-PANCE (Fig. 3 and 4). Building on this observation, we integrated RAMPhaGE into the POGO-PANCE framework for targeted evolution of linker regions flanking the LOV2 insertion in AraC. As a starting point, we therefore returned to the unevolved R2-LOV variant (Fig. 3A and fig. S4F,G), providing a suitable effector background for probing how junction composition influences allosteric coupling. Using the retron linker library, we generated a diverse phage population encoding systematic modifications to the LOV2-flanking linkers, including flexible glycine-serine motifs, rigid proline-rich sequences, and all possible single-residue substitutions. Deep sequencing of the gene pool after RAMPhaGE mutagenesis

confirmed broad coverage of the linker library, with all designed variants represented, albeit at variable frequencies (fig. S12A,B).

Next, the resulting phage library was split into two replicate pools and subjected to one round of negative selection under blue light to remove leaky variants, followed by two rounds of positive selection in the dark to enrich for constructs with high activity in absence of the trigger (Fig. 5G and fig. S14A-B). Following POGO-PANCE selection, a distinct enrichment pattern emerged, indicating strong selective pressure favoring specific linker compositions - consistently observed across both replicate pools (Fig. 5G and fig. S13C-D). Over 90% of enriched variants carried mutations in at least one linker, with pronounced asymmetry between the two junctions. The N-terminal (left) linker exhibited a strong bias toward single-residue deletions, particularly replacement of the original GS motif with a single lysine, which accounted for over 40% of the selected variants. In contrast, C-terminal (right) linker mutations were less common, appearing in only ~17% of sequences (fig. 5G,H, fig. S13C-E).

Sequencing of individual plaques confirmed the enrichment trends observed by deep sequencing (fig. S13C), revealing a striking convergence on specific amino acid substitutions within the left linker. Arginine (R), lysine (K), glutamine (Q), and asparagine (N) consistently replaced the original serine-glycine motif, suggesting selective pressure for side chains with hydrogen-bonding or helix-stabilizing potential. Structural predictions using AlphaFold3 (*47*) suggested that the lysine substitution extends the left linker into a continuous α-helix that seamlessly connects LOV2 to the AraC effector domain (Fig. 5I). The observed helix extension likely enhances allosteric transmission in a particularly dynamic protein region (*34*, *35*) by minimizing structural discontinuities at the sensor-effector interface. Thus, the accompanying single-residue deletion is plausibly required to align helical registers, providing a mechanistic rationale for the strong enrichment of these left linker variants during selection. Finally, to directly assess the impact of linker optimization on allosteric coupling, we cloned five highly enriched variants into an IPTG-inducible plasmid with a weak ribosome binding site as done before. All variants showed very strong activity in the dark and practically digital switching upon illumination, consistent with improved transmission of conformational signals from the LOV2 domain to AraC (Fig. 5J).

In summary, by combining RAMPhaGE with POGO-PANCE, we establish a powerful evolutionary strategy to selectively optimize switchable proteins and uncover latent design principles that illuminate and advance the engineering of synthetic allostery.

DISCUSSION

Allosteric regulation enables proteins to dynamically couple spatially distant sites, including functionally distinct domains (*1*). However, the diffuse, epistatic nature of the underlying residue

networks (*48, 49*) makes rational engineering of allosteric behavior extremely challenging. Here, we introduced POGO-PANCE, a phage-assisted evolution platform that builds upon principles of natural protein switch evolution by applying alternating selection regimes aligned to a defined stimulus. By iteratively selecting for both high activity and tight regulation across two environmental contexts, POGO-PANCE effectively navigates dual fitness landscapes, a critical requirement for evolving switchable protein function. We demonstrate this by converting a poorly active, modestly photoswitchable AraC-LOV hybrid into a potent optogenetic transcription factor exhibiting light-regulated control over gene expression spanning more than three orders of magnitude. The success of this transformation underscores the necessity of aligning selection regimes with functional goals, which previously was observed to be highly challenging (*50-52*). Indeed, pools evolved under mismatched regimes failed to yield useful switches and were taken over by "cheater" phages, while those exposed to appropriately toggled selection pressures converged on high-performance variants.

By integrating deep sequencing into POGO-PANCE, as previously shown for evolution of static protein features (*18, 53*), we gained insight into how adaptive trajectories unfold under opposing selection pressures, offering a unique, empirical access to how allosteric communication networks are built and/or refined. Traditional approaches, including rational design and structure prediction tools, remain limited by static models and typically fail to capture the conformational ensembles or context-specific dependencies that underlie allosteric regulation (*47*). By contrast, POGO-PANCE evolves and observes allosteric systems as they adapt over time, under defined environmental constraints. The resulting observations highlight the role of epistasis and environmental context in shaping the allosteric landscape. For example, mutations at E144K and V223M emerged in POGO-PANCE performed on AraC-LOV fusions and not the prior conventional PANCE of AraC alone, suggesting context-dependent adaptation and underscoring that combinations of mutations that may not be immediately obvious are often required for optimal function.

Crucially, traditional mutagenesis approaches fall short in exploring the structural plasticity required for tuning allosteric coupling, particularly at sensor-effector junctions. To address this, we developed RAMPhaGE, a retron-based diversification platform that enables precise, programmable introduction of substitutions, insertions, and deletions. This capability is particularly important for engineering allosteric proteins, where signal transmission often hinges on subtle geometric features, such as precise length of structural elements, helix phasing, or domain orientation, that lie beyond the reach of simple point mutagenesis. In our evolved AraC-LOV constructs, for example, a consistently enriched single-residue deletion at the sensor-effector junction was predicted to extend a continuous α-helix across the interface, likely enhancing allosteric signal relay. These findings illustrate the utility of InDel mutagenesis in modulating geometric parameters that are otherwise inaccessible via point mutations alone. More broadly, RAMPhaGE opens access to regions of sequence space previously obscured in phage-assisted evolution.

Conceptually, POGO-PANCE and RAMPhaGE should be compatible for studying and evolving allostery for a broad range of effector domains as well as inputs, including chemicals, pH or redox state. This opens a path to wide future exploration: Evolving various effectors for trigger-responsive behavior could not only feed various applications in biotechnology and biomedicine, but eventually reveal fundamental principles of allostery as such. Finally, our platform may open a path toward the *de novo* evolution of switchable proteins. While this remains a formidable challenge, analogous to how nature repeatedly evolved photosensitive proteins from unrelated scaffolds, POGO-PANCE provides the evolutionary infrastructure for similar efforts in the lab.

In summary, the combined POGO-PANCE and RAMPhaGE framework enables not only the directed evolution of powerful protein switches, but allows us to spy into the evolutionary trajectories that yield well-orchestrated allosteric networks. Together, they provide a powerful foundation for both practical applications of and fundamental insights into allostery.

REFERENCES

1. J. Monod, J. Wyman, J. P. Changeux, On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol* **12**, 88-118 (1965).
2. J. Monod, *Chance and Necessity: Essay on the Natural Philosophy of Modern Biology*. (Penguin Books 1977).
3. K. A. Reynolds, R. N. McLaughlin, R. Ranganathan, Hot Spots for Allosteric Regulation on Protein Surfaces. *Cell* **147**, 1564-1575 (2011).
4. O. Dagliyan, N. V. Dokholyan, K. M. Hahn, Engineering proteins for allosteric control by light or ligands. *Nat Protoc* **14**, 1863-1883 (2019).
5. O. Dagliyan *et al.*, Engineering extrinsic disorder to control protein activity in living cells. *Science* **354**, 1441-1444 (2016).
6. A. B. Guo *et al.*, Deep learning-guided design of dynamic proteins. *Science* **388**, eadr7094 (2025).
7. E. S. Boyden, F. Zhang, E. Bamberg, G. Nagel, K. Deisseroth, Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci* **8**, 1263-1268 (2005).
8. B. Wolf *et al.*, Rational engineering of allosteric protein switches by in silico prediction of domain insertion sites. *bioRxiv*, 2024.2012.2004.626757 (2024).
9. M. W. Young, S. A. Kay, Time zones: A comparative genetics of circadian clocks. *Nat Rev Genet* **2**, 702-715 (2001).
10. M. T. Alam *et al.*, The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. *Nature Communications* **8**, (2017).
11. F. Jacob, J. Monod, Genetic Regulatory Mechanisms in Synthesis of Proteins. *Journal of Molecular Biology* **3**, 318-356 (1961).
12. A. S. Chaudhari *et al.*, Light-dependent flavin redox and adduct states control the conformation and DNA-binding activity of the transcription factor EL222. *Nucleic Acids Res* **53**, gkaf215 (2025).
13. W. R. Gordon *et al.*, Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch. *Dev Cell* **33**, 729-736 (2015).
14. S. M. Miller, T. Wang, D. R. Liu, Phage-assisted continuous and non-continuous evolution. *Nat Protoc* **15**, 4101-4127 (2020).

15. K. M. Esvelt, J. C. Carlson, D. R. Liu, A system for the continuous directed evolution of biomolecules. *Nature* **472**, 499-503 (2011).
16. T. Suzuki *et al.*, Crystal structures reveal an elusive functional domain of pyrrolysyl- tRNA synthetase. *Nat Chem Biol* **13**, 1261-1266 (2017).
17. B. C. Dickinson, M. S. Packer, A. H. Badran, D. R. Liu, A system for the continuous directed evolution of proteases rapidly reveals drug-resistance mutations. *Nat Commun* **5**, 5352 (2014).
18. B. C. Dickinson, A. M. Leconte, B. Allen, K. M. Esvelt, D. R. Liu, Experimental interrogation of the path dependence and stochasticity of protein evolution using phage-assisted continuous evolution. *Proc Natl Acad Sci USA* **110**, 9007-9012 (2013).
19. A. H. Badran *et al.*, Continuous evolution of Bacillus thuringiensis toxins overcomes insect resistance. *Nature* **533**, 58-63 (2016).
20. M. F. Richter *et al.*, Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol* **38**, 883-891 (2020).
21. Systematic molecular evolution enables robust biomolecule discovery. *Nat Methods* **19**, 55-64 (2022).
22. A. H. Badran, D. R. Liu, Development of potent in vivo mutagenesis plasmids with broad mutational spectra. *Nat Commun* **6**, 8425 (2015).
23. J. C. Carlson, A. H. Badran, D. A. Guggiana-Nilo, D. R. Liu, Negative selection and stringency modulation in phage-assisted continuous evolution. *Nat Chem Biol* **10**, 216-222 (2014).
24. D. Belin, J. Costafrolaz, F. Silva, AraC Functional Suppressors of Mutations in the C-Terminal Domain of the RpoA subunit of the RNA Polymerase. *Microorganisms* **12**, (2024).
25. M. Lowe *et al.*, Computational and experimental investigation of constitutive behavior in AraC. *Proteins* **82**, 3385-3396 (2014).
26. J. Mathony, S. Aschenbrenner, P. Becker, D. Niopek, Dissecting the Determinants of Domain Insertion Tolerance and Allostery in Proteins. *Adv Sci (Weinh)* **10**, e2303496 (2023).
27. E. Romano *et al.*, Engineering AraC to make it responsive to light instead of arabinose. *Nature Chemical Biology* **17**, 817-827 (2021).
28. A. K. Brödel, R. Rodrigues, A. Jaramillo, M. Isalan, Accelerated evolution of a minimal 63-amino acid dual transcription factor. *Sci Adv* **6**, eaba2728 (2020).
29. S. S. Ringquist S, Barrick D, Green L, Binkley J, Stormo GD, Gold L., Translation initiation in Escherichia coli: sequences within the ribosome-binding site. *Mol Microbiol.* **6**, 1219-1229 (1992).
30. S. M. Soisson, B. MacDougall-Shackleton, R. Schleif, C. Wolberger, Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**, 421-425 (1997).
31. X. Zhu *et al.*, New xylose transporters support the simultaneous consumption of glucose and xylose in Escherichia coli. *mLife* **1**, 156-170 (2022).
32. B. P. Hubbard *et al.*, Continuous directed evolution of DNA -binding proteins to improve TALEN specificity. *Nat Methods* **12**, 939-942 (2015).
33. M. D. Hoffmann, F. Bubeck, R. Eils, D. Niopek, Controlling Cells with Light and LOV. *Adv Biosyst* **2**, (2018).
34. A. Tischer, M. J. Brown, R. F. Schleif, M. Auton, Arabinose Alters Both Local and Distal H-D Exchange Rates in the AraC Transcriptional Regulator. *Biochemistry* **58**, 2875-2882 (2019).
35. A. S. Halavaty, K. Moffat, N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from. *Biochemistry* **46**, 14001-14009 (2007).
36. J. P. Zayner, C. Antoniou, A. R. French, R. J. Hause, Jr., T. R. Sosnick, Investigating models of protein function and allostery with a widespread mutational analysis of a light-activated protein. *Biophys J* **105**, 1027-1036 (2013).
37. S. M. Harper, L. C. Neil, K. H. Gardner, Structural basis of a phototropin light switch. *Science* **301**, 1541-1544 (2003).

38. S. Crosson, K. Moffat, Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. *Plant Cell* **14**, 1067-1075 (2002).
39. M. Salomon, J. M. Christie, E. Knieb, U. Lempert, W. R. Briggs, Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin. *Biochemistry* **39**, 9401-9410 (2000).
40. S. M. Harper, J. M. Christie, K. H. Gardner, Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. *Biochemistry* **43**, 16184-16192 (2004).
41. M. Jain, H. E. Olsen, B. Paten, M. Akeson, The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol* **17**, 239 (2016).
42. C. B. Macdonald *et al.*, DIMPLE: deep insertion, deletion, and missense mutation libraries for exploring protein variation in evolution, disease, and biology. *Genome Biol* **24**, 36 (2023).
43. C. B. Fishman *et al.*, Continuous multiplexed phage genome editing using recombitrons. *Nat Biotechnol*, (2024).
44. T. M. Wannier *et al.*, Recombineering and MAGE. *Nat Rev Methods Primers* **1**, (2021).
45. R. B. R. O'Callaghan, W. Paranchych, The effect of M13 phage infection upon the F pili of E. coli. *1* **54**, 220-229 (1973).
46. F. Bubeck *et al.*, Engineered anti-CRISPR proteins for optogenetic control of CRISPR-Cas9. *Nat Methods* **15**, 924-927 (2018).
47. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
48. D. Listov, C. A. Goverde, B. E. Correia, S. J. Fleishman, Opportunities and challenges in design and optimization of protein function. *Nat Rev Mol Cell Biol* **25**, 639-653 (2024).
49. T. N. Starr, J. W. Thornton, Epistasis in protein evolution. *Protein Sci* **25**, 1204-1218 (2016).
50. Z. M. Shaver *et al.*, Attempted use of PACE for riboswitch discovery generates three new translational theophylline riboswitch side products. *BMC Res Notes* **11**, 861 (2018).
51. E. Goicoechea Serrano, C. Blazquez-Bondia, A. Jaramillo, T7 phage-assisted evolution of riboswitches using error-prone replication and dual selection. *Sci Rep* **14**, 2377 (2024).
52. T. B. Roth, B. M. Woolston, G. Stephanopoulos, D. R. Liu, Phage-Assisted Evolution of Bacillus methanolicus Methanol Dehydrogenase 2. *ACS Synth Biol* **8**, 796-806 (2019).
53. L. Schmidheini *et al.*, Continuous directed evolution of a compact CjCas9 variant with broad PAM compatibility. *Nat Chem Biol* **20**, 333-343 (2024).

Acknowledgments: We thank the Niopek and Mathony labs, as well as the thesis advisory committee to N.S., Seth Shipman, UCSF and Kerstin Göpfrich, Heidelberg University for helpful discussions. We thank Konrad Herbst for advice on the Nanopore sequencing; Andreas K. Brödel and Mark Isalan for sharing information regarding mutagenesis plasmid; Kevin Esvelt for information regarding media for PANCE; Tobias Stadelmann for support in establishing the PANCE-method in our group. ChatGPT, model 4o was used for language editing.

Funding: This work was supported by the European Union (ERC, DaVinci-Switches, project number 101041570). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. D.N. is also grateful for funding by the Aventis Foundation. Anna von Bachmann is supported by the Konrad Zuse School of Excellence in Learning and Intelligent Systems (ELIZA) through the DAAD programme Konrad Zuse Schools of Excellence in Artificial Intelligence, sponsored by the Federal Ministry of Education and Research.

Author contributions: DN conceived the study and refined it with NTS and JM. NTS and DN designed experiments. NS performed experiments with contributions from MLG, NL, ASK and SW. NS, AvB and AH analyzed data. NS, JM and DN interpreted data. BW created the software tool for RAMPhaGE oligo pool design based on concepts by N.S.. DN supervised the work with support by JM. NS and DN wrote the manuscript draft. All authors reviewed and edited the manuscript.

Competing interests: The authors report no competing interests.

Data and materials availability: Supplementary Tables 1-6 are available as Supplementary Data S1; Table S1 contains all used strains and plasmid sequences. Raw measurement data is available as Data S2. Scripts used for data analysis, RFP measurement data, as well as the tool for RAMPhaGE oligo pool design, are available on GitHub: https://github.com/Niopek-Lab/POGO_PANCE.git. NGS raw data is available on Zenodo (doi:10.5281/zenodo.15650047). Relevant plasmids will be made available via Addgene. Other information can be obtained from the corresponding authors upon reasonable request.

FIGURES AND LEGENDS

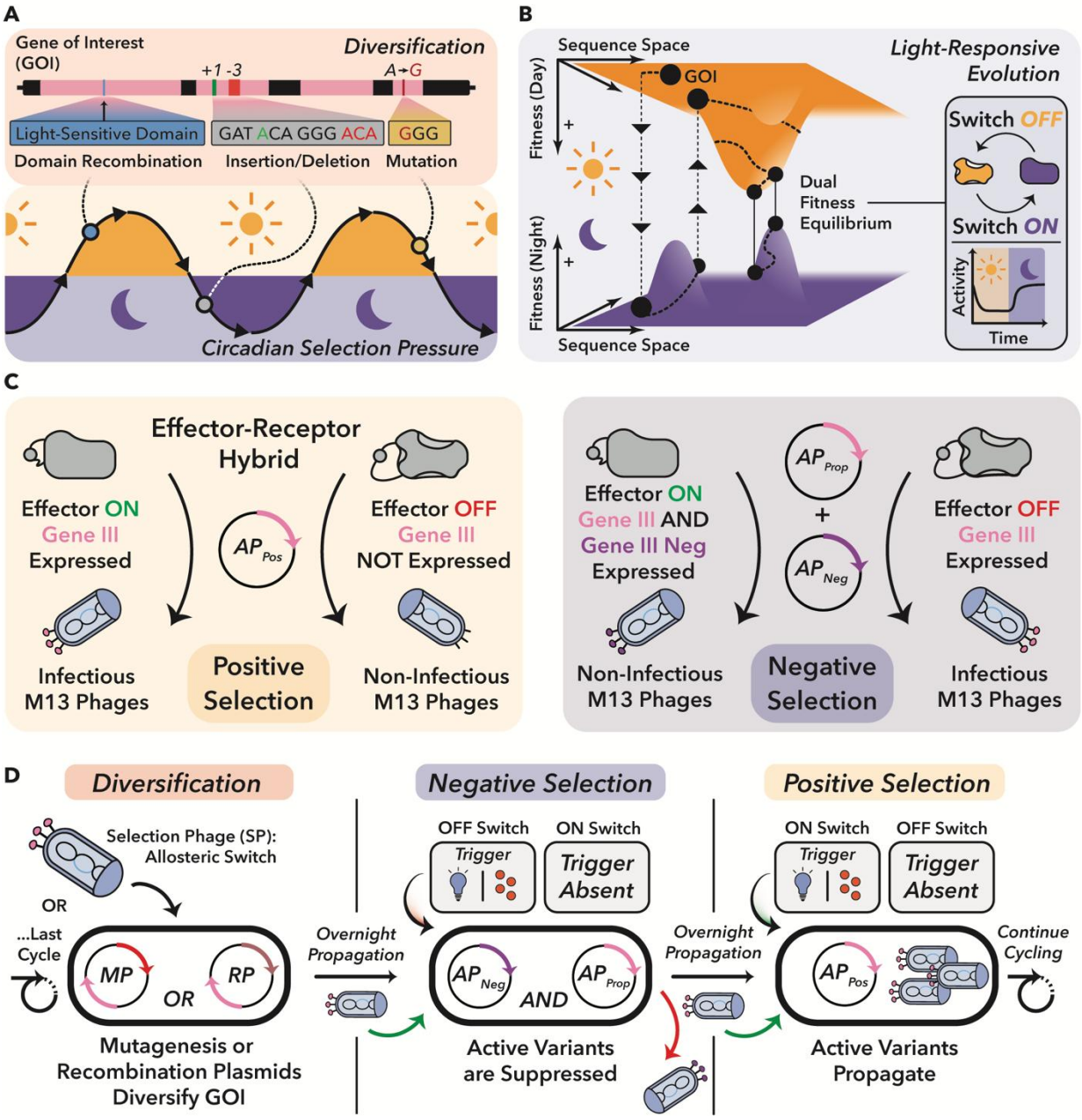

Fig. 1. Conceptual framework and workflow for the phage-assisted evolution of stimulus-responsive allosteric protein switches with POGO-PANCE. (A) Schematic of natural evolution of light-responsive proteins through domain recombination and mutation under alternating selective pressures such as day–night cycles. Such environmental cues drive the emergence of allosteric regulation by favoring variants with input-responsive activity. **(B)** Circadian selection imposes alternating fitness constraints on a dual-state protein, shaping evolution across fluctuating environmental conditions. **(C)** Engineered phage-assisted selection system towards dual-state regulation. A selection phage (SP) encodes an allosteric switch that

regulates gIII expression. In the positive selection circuit, switch activation induces gIII expression, enabling phage propagation. In the negative selection circuit, inappropriate activation triggers expression of a dominant-negative gIII variant, suppressing phage infectivity. **(D)** Overview of staged POGO-PANCE regime: a selection phage (SP) library is diversified, then iteratively subjected to alternating negative and positive selection under opposing environmental states, enriching for protein variants with precise, input-dependent regulation.

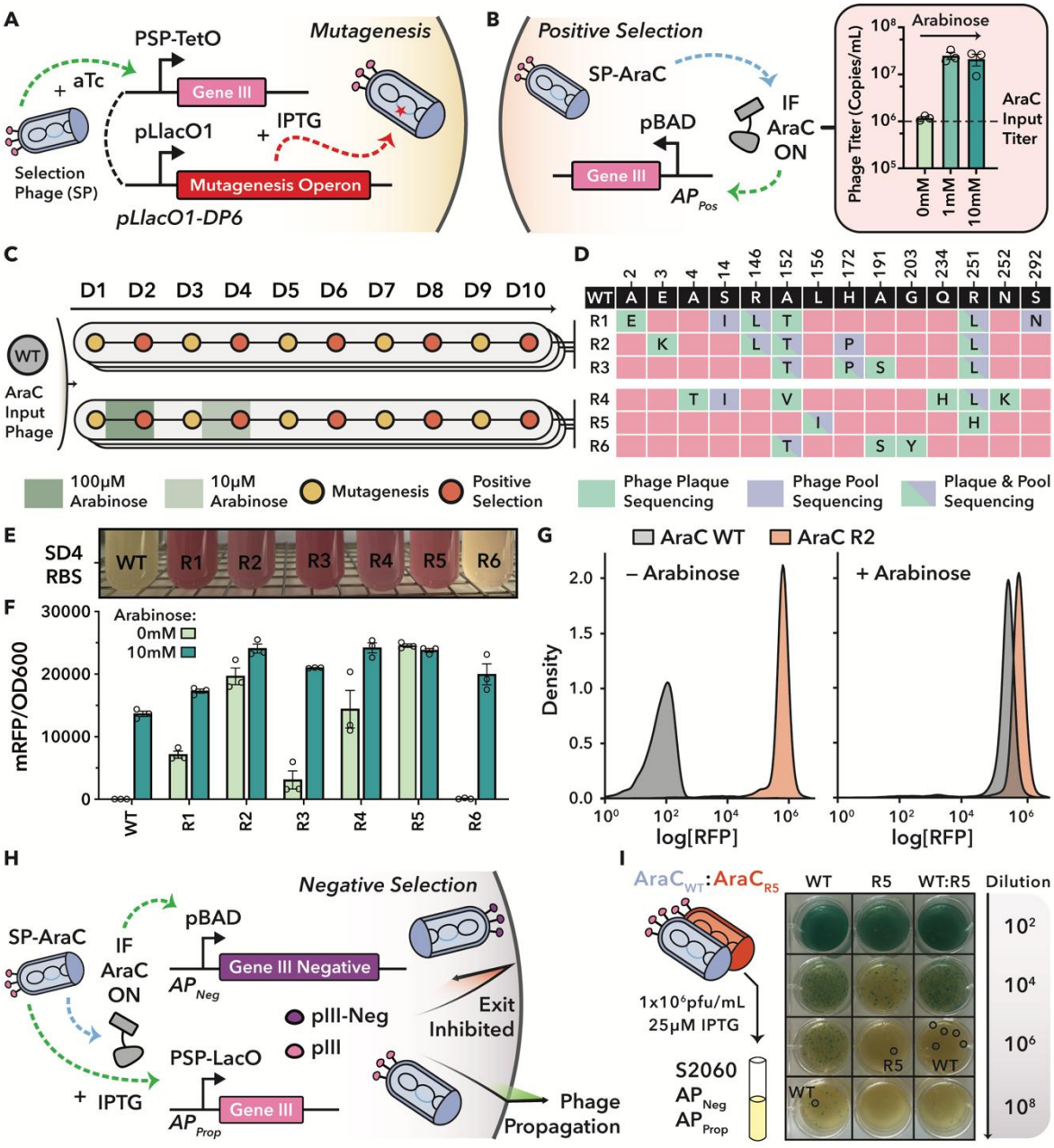

Fig. 2. Phage-assisted non-continuous evolution of hyperactive AraC variants. (A) Schematic of the used mutagenesis system. In the plLacO1-DP6 circuit, infection with a selection phage (SP) and addition of anhydrotetracycline (aTc) induces gIII expression and enable phage propagation, while IPTG addition activates a mutagenesis operon that diversifies the SP genome. **(B)** Positive selection circuit. SPs encoding active AraC variants induce gIII expression via the pBAD promoter. Right: Arabinose-dependent propagation of an SP encoding WT AraC, demonstrating induction-dependent propagation. Bars represent mean ± Standard Error of the Mean (S.E.M.) of three biological replicates, as quantified by qPCR. Corresponding phage plaque assay shown in Fig. S1B. **(C)** Summary of the AraC PANCE campaign. Phage titers from each day (D1-10) are shown in Fig. S1C, D. **(D)** Genotypic analysis of evolved AraC

variants. Pooled phage populations and one individual plaque from the final selection day (D10) were sequenced, revealing six distinct variants (R1–R6). **(E)** WT AraC and evolved variants (R1–R6) were expressed from a LacI-dependent expression cassette containing the SD4 *(29)* RBS and co-transformed with a pBAD regulated mRFP reporter plasmid (Fig. S1E). Cultures were grown without IPTG or Arabinose. **(F)** Quantification of AraC variant activity. AraC variants were expressed as in **E,** but using the weaker sd2 *(29)* RBS. RFP fluorescence normalized to OD_{600} was measured in the presence or absence of 10 mM L-arabinose in a plate reader. Data shown as mean ± S.E.M. from n = 3 biological replicates. **(G)** Flow cytometry analysis of WT AraC and R2 variant. Reporter fluorescence was quantified in presence or absence of 16 mM L-arabinose. Gating strategy described in fig. S2. **(H)** Schematic of the negative selection circuit. SPs encoding active AraC variants induce expression of a dominant-negative form of gIII (pIII-Neg) from the pBAD promoter, preventing phage host cell escape. Concurrently, IPTG induction drives WT gIII expression to support propagation. **(I)** Head-to-head competition assay between WT AraC and R5 under negative selection. Host cells harboring AP_{Prop} and AP_{Neg} were infected with (i) 1×10^6 pfu/mL of either phage individually or (ii) with 5×10^5 pfu/mL of each phage combined. Cultures were supplemented with 25 μM IPTG (0 mM Arabinose) at the time of phage infection, followed by plague assay. Resulting plaques were sequenced (circled in black) to determine dominant variants. Expanded results in Fig. S3B, with sequence confirmation in Fig. S3C.

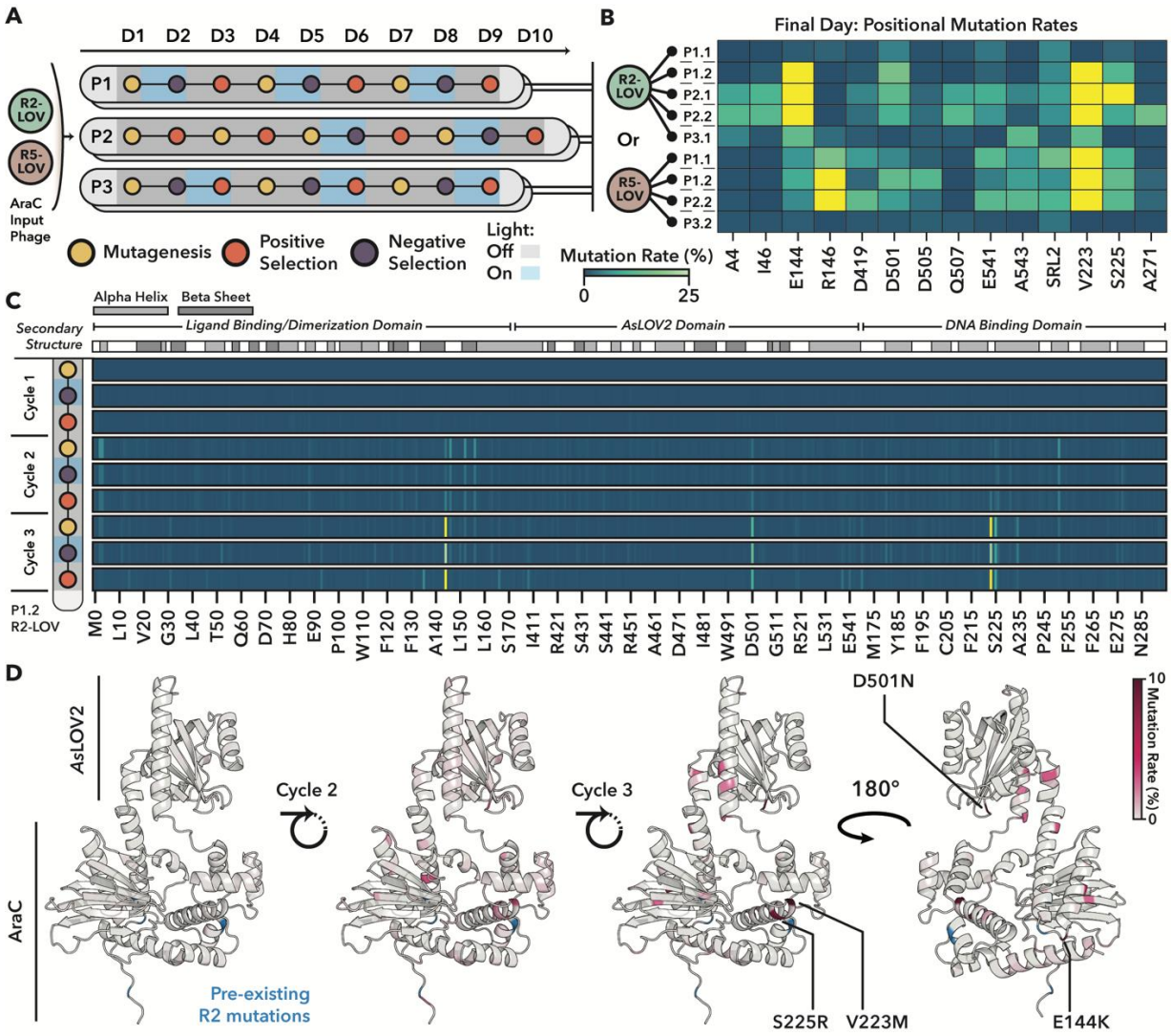

Fig. 3. Directed evolution of light-sensitive AraC S170 AsLOV2 variants using POGO-PANCE. (A) Summary of AraC-LOV POGO-PANCE campaigns. Mutagenized phage libraries encoding AraC-R2-LOV or AraC-R5-LOV were subjected to iterative rounds of positive and negative selection under alternating light and dark conditions (for variants see fig. S4D-G). Three selection strategies were tested, each performed in duplicate pools. The illumination set-up and phage titers are shown in fig. S4C and S5A,B, respectively. **(B)** Endpoint positional enrichment analysis of mutations across evolved variants quantified by Nanopore sequencing. Mutation rates at each residue within AraC-LOV are displayed as the percentage of total reads per position, with residues exceeding 25% shaded in yellow. Positions that have an enrichment >10% in at least one of the pools are shown, excluding position R38 which arose as a systematic artifact in Nanopore sequencing (fig. S14). **(C)** Temporal mutation enrichment during the R2-LOV P1.2 evolution campaign quantified by Nanopore sequencing. Mutation rates at each residue within AraC-LOV are displayed as the percentage of total reads per position, relative to enrichment observed after mutagenesis round 1 (see fig. S14). Residues exceeding 25% are shaded in yellow. Secondary structure annotations for AraC-LOV are shown above. (**B,C**) The shown "Mutation rate (%)" label applies to both panels **(D)** Mutation rates observed after

positive selection steps for the R2-LOV P1.2 evolution in **C** are mapped onto the Alphafold 3 predicted structure of AraC-LOV, with residues exceeding 10% shaded in burgundy. Pre-existing mutations in R2 (see Fig. 2) are marked in blue.

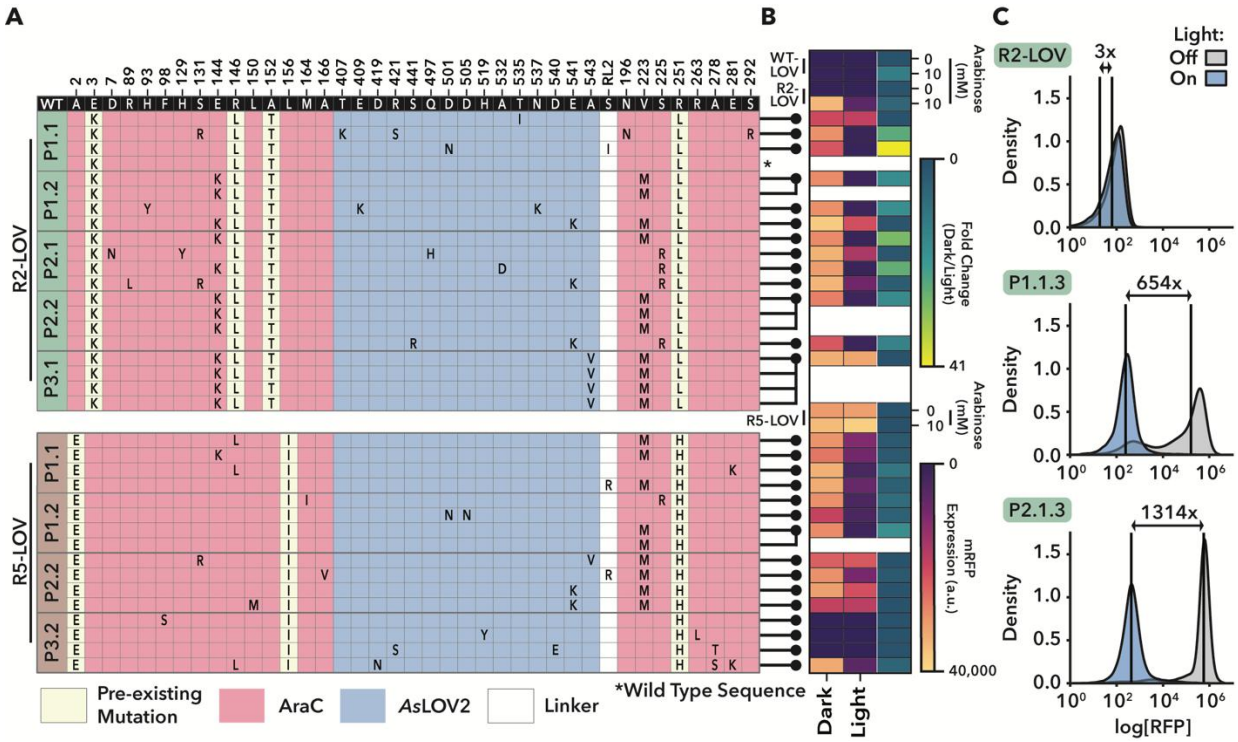

Fig. 4. Functional characterization of evolved AraC-LOV variants derived from POGO-PANCE. (A) Genotypes of individual endpoint variants. Four phage plaques from the final day of each evolution pool were isolated, cloned into an expression vector, and sequenced. Mutations across the LOV2-flanking linker, LOV2 domain, and AraC regions are indicated. **(B)** Heatmap of RFP reporter activity under light and dark conditions. Variants (linked to identities in panel A) were expressed in *E. coli* assessed for activity using the RFP reporter assay. Cultures were incubated in the light or dark for 18 hours, followed by measurement of mRFP fluorescence and OD_{600} in a plate reader. Variants were incubated with no Arabinose or IPTG added, and side-by-side compared to WT, R2-, and R5-LOV2 controls with or without L-Arabinose. Mean fluorescence values, normalized to OD_{600}, were obtained from the average of three biological replicates (first two columns); fold change values (dark/light) are shown in the third column. **(C)** Flow cytometry analysis of light-dependent RFP expression in top-performing variants. Cultures were prepared as described in **(B)**, without arabinose or IPTG, for the R2-LOV, R2-LOV P1.1.3, and R2-LOV P2.1.3 variants to assess switching behavior based on mRFP reporter fluorescence. Median mRFP fold changes between dark and light conditions are shown.

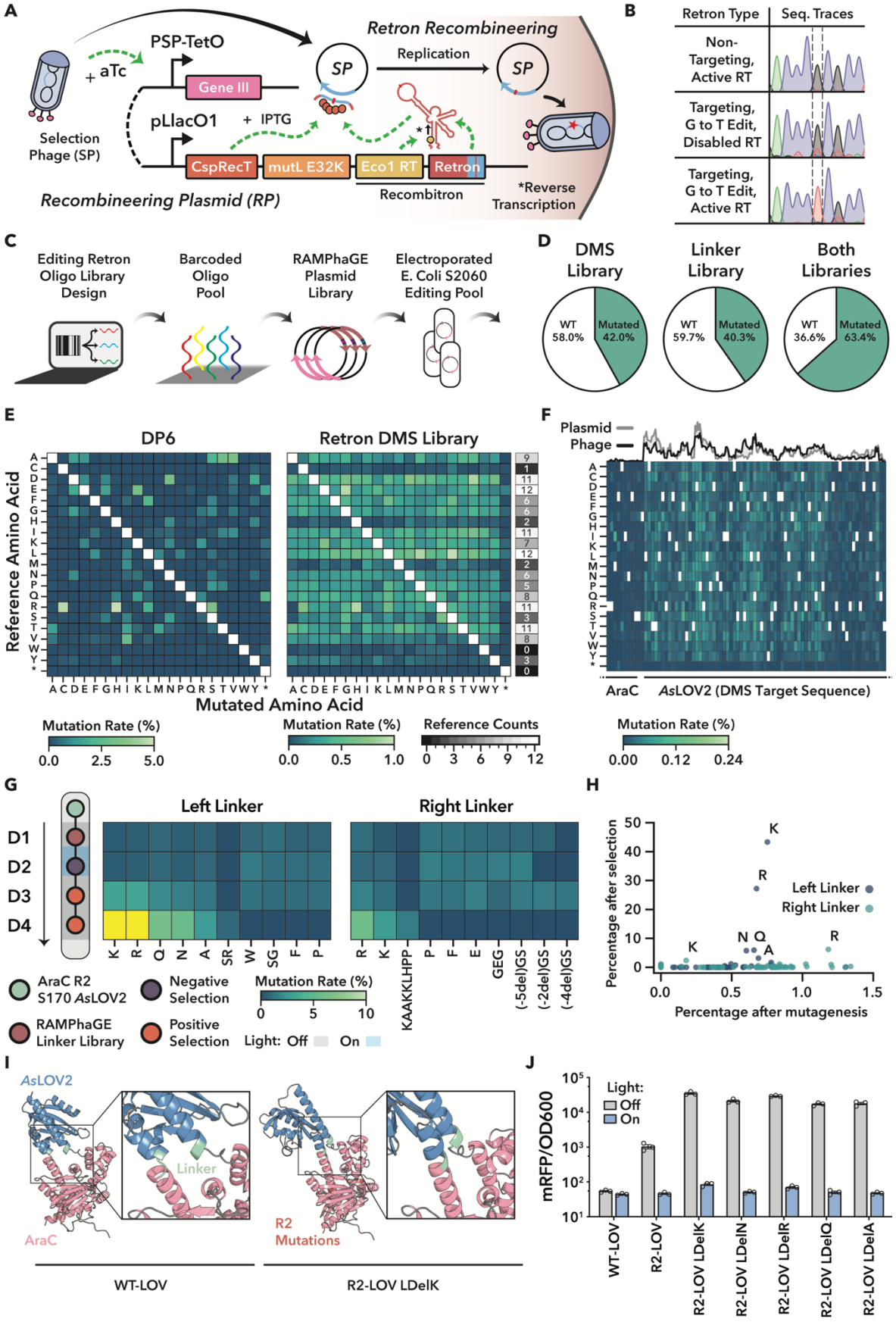
A
Retron Recombineering
PSP-TetO
+ aTc
Gene III
SP
Replication
SP
pLlacO1
+ IPTG
Selection Phage (SP)
CspRecT
mutL E32K
Eco1 RT
Retron
Recombitron
*Reverse Transcription
Recombineering Plasmid (RP)
B
Retron Type
Seq. Traces
Non-Targeting, Active RT
Targeting, G to T Edit, Disabled RT
Targeting, G to T Edit, Active RT
C
Editing Retron Oligo Library Design
Barcoded Oligo Pool
RAMPhaGE Plasmid Library
Electroporated E. Coli S2060 Editing Pool
D
DMS Library
WT 58.0%
Mutated 42.0%
Linker Library
WT 59.7%
Mutated 40.3%
Both Libraries
WT 36.6%
Mutated 63.4%
E
DP6
Retron DMS Library
Reference Amino Acid
Mutated Amino Acid
Mutation Rate (%)
0.0 2.5 5.0
Mutation Rate (%)
0.0 0.5 1.0
Reference Counts
0 3 6 9 12
F
Plasmid
Phage
AraC
AsLOV2 (DMS Target Sequence)
Mutation Rate (%)
0.00 0.12 0.24
G
D1
D2
D3
D4
Left Linker
Right Linker
K R Q N A SR W SG F P
R K KAAKKLHPP P F E GEG (-5del)GS (-2del)GS (-4del)GS
AraC R2 S170 AsLOV2
Negative Selection
RAMPhaGE Linker Library
Positive Selection
Mutation Rate (%)
0 5 10
Light: Off On
H
Percentage after selection
Percentage after mutagenesis
Left Linker
Right Linker
I
AsLOV2
Linker
AraC
R2 Mutations
WT-LOV
R2-LOV LDelK
J
mRFP/OD600
Light:
Off
On
WT-LOV
R2-LOV
R2-LOV LDelK
R2-LOV LDelN
R2-LOV LDelR
R2-LOV LDelQ
R2-LOV LDelA

Fig. 5. Directed evolution of AraC-LOV variants using Retron-Assisted, Multiplexed, Phage Gene Evolution (RAMPhaGE) and POGO-PANCE. (A) Overview of the retron recombineering system. Infection with a selection phage (SP) and addition of (aTc) induce gIII expression, while IPTG activates a recombitron operon encoding CspRecT, MutL E32K, and Eco1 reverse transcriptase (RT). Expressed retrons are reverse-transcribed into ssDNA donors containing a central predefined edit flanked by homology arms, enabling targeted editing of the SP genome during replication. Non-targeting relates to a control retron with no homology to the phage-encoded transgene. **(B)** Sanger sequencing chromatograms of phage pools analyzed after overnight editing with 90bp retrons under the indicated conditions. The G to T edit position is indicated with dashed lines. Expanded data in Fig. S8. **(C)** Schematic of RAMPhaGE editing library construction. Barcoded retron edit oligo pools were subcloned into a recombineering plasmid, and transformed into SP-infectable *E. coli* to enable pooled editing. Additional details in Fig. S9. **(D)** Editing outcomes showing the percentage of edited versus WT phages from retron libraries targeting either the LOV2 domain (DMS Library), the SG and GS linker regions (Linker Library), or both. **(E)** Comparison of amino acid-level mutation coverage following a single passage using either the plLacO1-DP6 mutagenesis system or the retron-encoded DMS library. In both cases, only the LOV2 region of the reference sequence was analyzed. The y-axis indicates the number of times each amino acid occurs in the LOV2 reference. **(F)** Amino acid-level mutation rates per position after a single phage passage on the DMS library carrying hosts. The heatmap shows observed mutation rate percentages at each amino acid position, with the reference sequence indicated by white boxes, for a region around the left AraC-LOV boundary. The top line plot compares the relative nucleotide mutation frequency encoded in the DMS library (gray line; refers to underlying retron-encoding plasmid pool) to those observed in the edited phage population (black line), with both distributions normalized to sum to 1. Expanded data are provided in fig. S11. **(G)** Enrichment analysis from a POGO-PANCE/RAMPhaGE evolution campaign using R2-LOV as input. Mutation rates observed in the left and right linkers of AraC-LOV were quantified by Nanopore sequencing. The top 10 linkers with the highest variance across all POGO-PANCE steps are shown at each site as the percentage of total reads, with residues exceeding 10% shaded in yellow. The SG variant of the left linker refers to a silent mutation. An expanded explanation of variant naming conventions are provided in Supplementary Table 6 additional linker library characterization and phage titer data are shown in Figs. S12 and S13. **(H)** Scatter plot comparing variant frequencies after mutagenesis (x-axis) versus final positive selection (y-axis) steps. Variants observed in the left and right linker are indicated. **(I)** AlphaFold3-predicted structure of WT AraC-LOV2 and comparison to that of R2-LOV2 in which the left SG linker is replaced with a single lysine. AraC, AsLOV2, and linker regions are colored pink, blue, and green, respectively. **(J)** Functional validation of enriched R2-LOV linker variants using the RFP reporter assay. Samples induced with 400μM IPTG were incubated in the dark or illuminated with blue light. RFP fluorescence normalized to OD600 was measured in a plate reader. Data shown as mean ± S.E.M. from n = 3 biological replicates.