

EVOLUTION

Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks

Jia Zheng^{1,2}, Joshua L. Payne^{2,3}, Andreas Wagner^{1,2,4*}

Cryptic genetic variation can facilitate adaptation in evolving populations. To elucidate the underlying genetic mechanisms, we used directed evolution in *Escherichia coli* to accumulate variation in populations of yellow fluorescent proteins and then evolved these proteins toward the new phenotype of green fluorescence. Populations with cryptic variation evolved adaptive genotypes with greater diversity and higher fitness than populations without cryptic variation, which converged on similar genotypes. Populations with cryptic variation accumulated neutral or deleterious mutations that break the constraints on the order in which adaptive mutations arise. In doing so, cryptic variation opens paths to adaptive genotypes, creates historical contingency, and reduces the predictability of evolution by allowing different replicate populations to climb different adaptive peaks and explore otherwise-inaccessible regions of an adaptive landscape.

Cryptic genetic variation is standing genetic variation that does not normally contribute to heritable phenotypic variation in a population but that can bring forth phenotypic variation after environmental change or genetic perturbation (1, 2). Cryptic variation exists because phenotypes are to some extent robust to genetic change (3–6). Because of its potential role in adaptive evolution, cryptic variation has attracted widespread interest (7–17), but supporting experimental evidence is limited (1, 17–19). One distinguishing feature of cryptic variation is that the conditions inducing its phenotypic effects are rare or absent in a population's history. In consequence, it can be protected from selection until a new environment arises in which cryptic variation may give rise to new and potentially beneficial phenotypes (1, 2). The molecular mechanisms of adaptation under cryptic variation are difficult to study for complex phenotypes of whole organisms, because their genetic basis often remains elusive (17, 20). Such mechanisms are better studied with simple and tractable systems, such as evolving proteins. Many mutations in proteins interact epistatically (i.e., nonadditively), which can render adaptive landscapes rugged and multipeaked (21–26). An evolving population's location on a rugged adaptive landscape influences which of these peaks are accessible (26–28). These observations hint that cryptic variation may help populations of evolving proteins enter re-

gions of an adaptive landscape that would otherwise remain inaccessible.

Results

To create cryptic genetic variation, we subjected each of four replicate populations of yellow fluorescent protein (YFP; populations V^C, with C for cryptic) to four rounds ("generations") of directed evolution subject to stringent stabilizing selection to maintain yellow fluorescence (phase I; Fig. 1 and fig. S1). Specifically, we allowed $\sim 5 \times 10^6$ YFP variants in each generation and in each replicate population to evolve, and we subjected these variants to PCR mutagenesis (0.84 amino acid-changing mutations per YFP molecule per generation; tables S1 and S2). In every generation of phase I, we allowed only those 20% of cells of evolving populations to survive whose yellow fluorescence intensity lay in a narrow interval around the median of ancestral YFP (Fig. 1) (29). Such stringent stabilizing selection allows the accumulation of cryptic variation, because only the mutations (or their combinations) that have little effect on yellow fluorescence can persist. We then initiated phase II, in which we subjected the same populations to four generations of stringent directed evolution toward green fluorescence (Fig. 1). As controls, we also subjected four populations (called V⁰, for zero initial cryptic variation) that started from identical ancestral YFP molecules to four generations of evolution toward green fluorescence. We then compared the change in green fluorescence intensity during phase II in populations V^C with that of the control populations V⁰. Populations V^C reached significantly higher green fluorescence during three of the four generations of evolution in phase II (Fig. 2A), and they adapted approximately three times faster during the first generation of phase II

(fig. S2A). In addition, populations V^C more rapidly evolved a green (512-nm) emission peak than populations V⁰ (Fig. 2B). At the evolutionary end point, three of four V^C populations showed significantly greater green fluorescence than the four V⁰ populations (two-way analysis of variance [$F_{7,16} = 46.5$, $P = 1.99 \times 10^{-9}$], post hoc Tukey's test [$P < 0.05$ for V^C replicates 1, 2, and 4 relative to the four V⁰ populations]) (Fig. 2C and table S3). In sum, the genetic variation accumulated in phase I facilitated the evolution of green fluorescence during phase II.

To study why this genetic variation facilitated adaptive evolution, we used single-molecule real-time sequencing (SMRT) to genotype ~ 500 to 1000 evolved variants for each replicate population and for each generation (table S4). We first noticed that V^C populations were more diverse than V⁰ populations throughout phase II. Specifically, they harbored on average more mutations per individual molecule (Fig. 3A). They also showed a broader distribution of mutations per individual molecule (fig. S2B), as well as greater overall genetic diversity (fig. S2C) (29). Additionally, the four V^C populations diverged to a much greater extent from each other (Fig. 3B and fig. S2D).

We then studied the dynamics of polymorphisms in each replicate population during phase II (fig. S3A) (29) and observed that two mutations (G66S and Y204C; see the legend for Fig. 3D for a full list of amino acid abbreviations and an explanation of mutation descriptions) swept through all replicate V⁰ and V^C populations, with two other mutations (F65L or F47L) achieving high or medium frequency (>10%) in two or more V⁰ populations and in one or two V^C populations. Because of their ubiquity, we refer to these four mutations as typical mutations (fig. S3B). At the evolutionary end point, most of these mutations co-occurred in three similar and high-frequency genotypes that share the two mutations G66S and Y204C and that harbor one additional mutation each, i.e., F65L, F47L, or L43M. We refer to these genotypes as T1, T2, and T3, or typical genotypes (Fig. 3, C and D), and to the combination of the G66S and Y204C mutations as genotype T.

Populations V^C evolved differently from populations V⁰. First, 17 alternative mutations attained a frequency of more than 10% in V^C populations but in none of the V⁰ populations [except the mutation V164A, which reached a frequency of 10.9% in V⁰₄ (fig. S3A)]. Also, typical genotypes dominated only one replicate V^C population (number 2), in contrast to their importance in V⁰ populations. The remaining populations were dominated by one or two of four other alternative genotypes (A1 to A4), which contained some combination of 11 alternative mutations in the genetic background T (Fig. 3, C and D). We measured the green fluorescence intensity of the three typical genotypes as well as of the four alternative

¹Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland. ²Swiss Institute of Bioinformatics, Quartier Sorge-Batiment Genopode, Lausanne, Switzerland. ³Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland. ⁴The Santa Fe Institute, Santa Fe, NM, USA.

*Corresponding author. Email: andreas.wagner@ieu.uzh.ch

genotypes (29). Three of the alternative genotypes exhibited greater green fluorescence than all typical genotypes (Fig. 3E).

In sum, during directional selection for green fluorescence: (i) More diverse genotypes attain high frequency in populations V^C than in populations V^0 ; (ii) different alternative genotypes dominate each of three replicate populations V^C , and (iii) three of the four alternative genotypes show significantly higher green fluorescence than all three typical genotypes.

Because V^C populations evolved faster than V^0 populations in phase II (Fig. 2A), we suspected that some of their adaptive mutations or genotypes had already accumulated in phase I. We thus studied the phase I evolutionary dynamics of the 4 typical mutations and 11 alternative mutations (fig. S4A). All 15 mutations were already present above our phase I detection limit of 0.064 to 0.16% (29), and 11 of the 15 mutations reached frequencies between 0.5 and 2.5% in at least one of the V^C populations.

This demonstrates that the variants accumulated in phase I are relevant to the exploration of different high-fitness genotypes in phase II. We performed additional directed evolution experiments starting from the YFP ancestor but in the complete absence of selection, which allowed us to determine how fast individual variants would increase in frequency through mutation alone. High-throughput sequencing showed that the frequency of all but one (F47L) of the mutations had not increased significantly more than

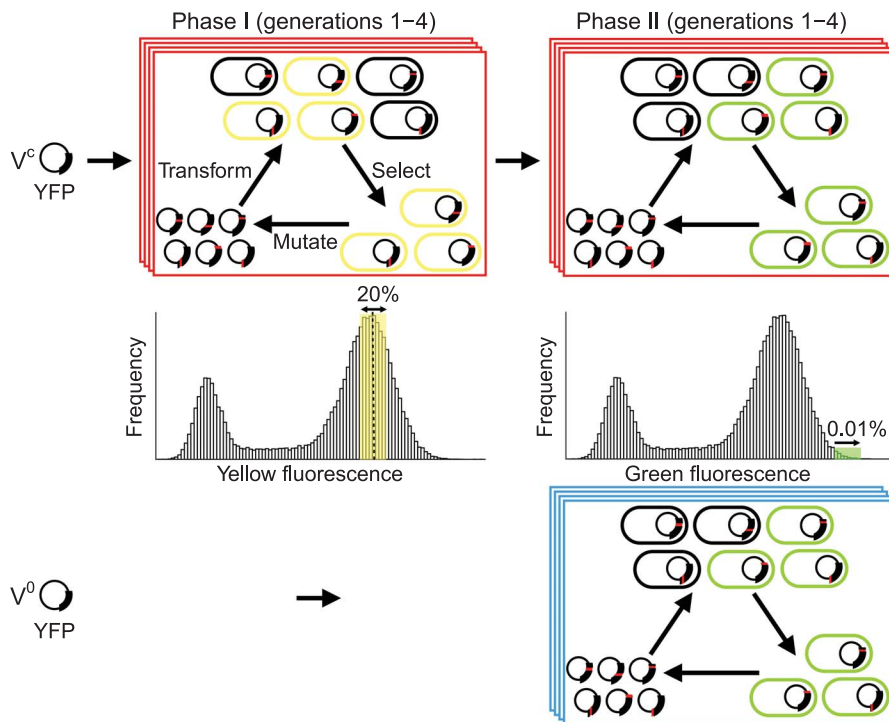


Fig. 1. Experimental evolution of YFP. In phase I, we subjected four replicate populations of YFP to four generations of directed evolution under stabilizing selection for the native yellow fluorescence, allowing only those ~20% of cells closest to the median (dashed vertical line) of ancestral yellow fluorescence [V^C , $\lambda_{ex} = 488$ nm and $\lambda_{em} = 530 \pm 15$ nm (29)]. In phase II, we subjected these populations to four further generations of strong directional selection for green fluorescence, allowing only 0.01% of cells to survive [$\lambda_{ex} = 405$ nm and $\lambda_{em} = 525 \pm 25$ nm (29)]. As controls, we subjected four populations (V^0) consisting of initially identical YFP molecules to the same stringent directed evolution for green fluorescence (29).

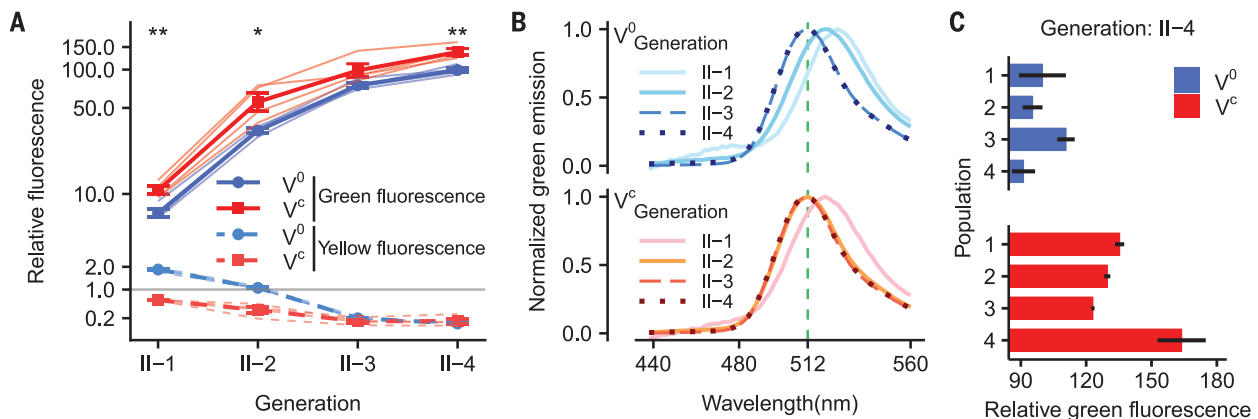


Fig. 2. Cryptic variation leads to faster color change and higher fluorescence. (A) Fold change of yellow and green fluorescence intensities relative to the ancestral YFP during phase II evolution (29). Error bars represent 1 SEM, from four replicate populations (thin lines). Note the logarithmic vertical scale. * $P < 0.05$; ** $P < 0.01$ (one-sided t tests with Holm adjustments). (B) Emission spectra (shown as mean values of four replicate populations) of evolving populations V^0 and V^C at

the new excitation wavelength (405 nm) in phase II (29). The vertical axes indicate the relative fluorescence intensity at a given emission wavelength (horizontal axis) relative to the maximal fluorescence intensity at the emission peak 512 nm (green vertical dashed line). (C) Fold change of green fluorescence intensity relative to the ancestral YFP for each replicate population at the evolutionary end point. Error bars denote SD ($n = 3$) (29).

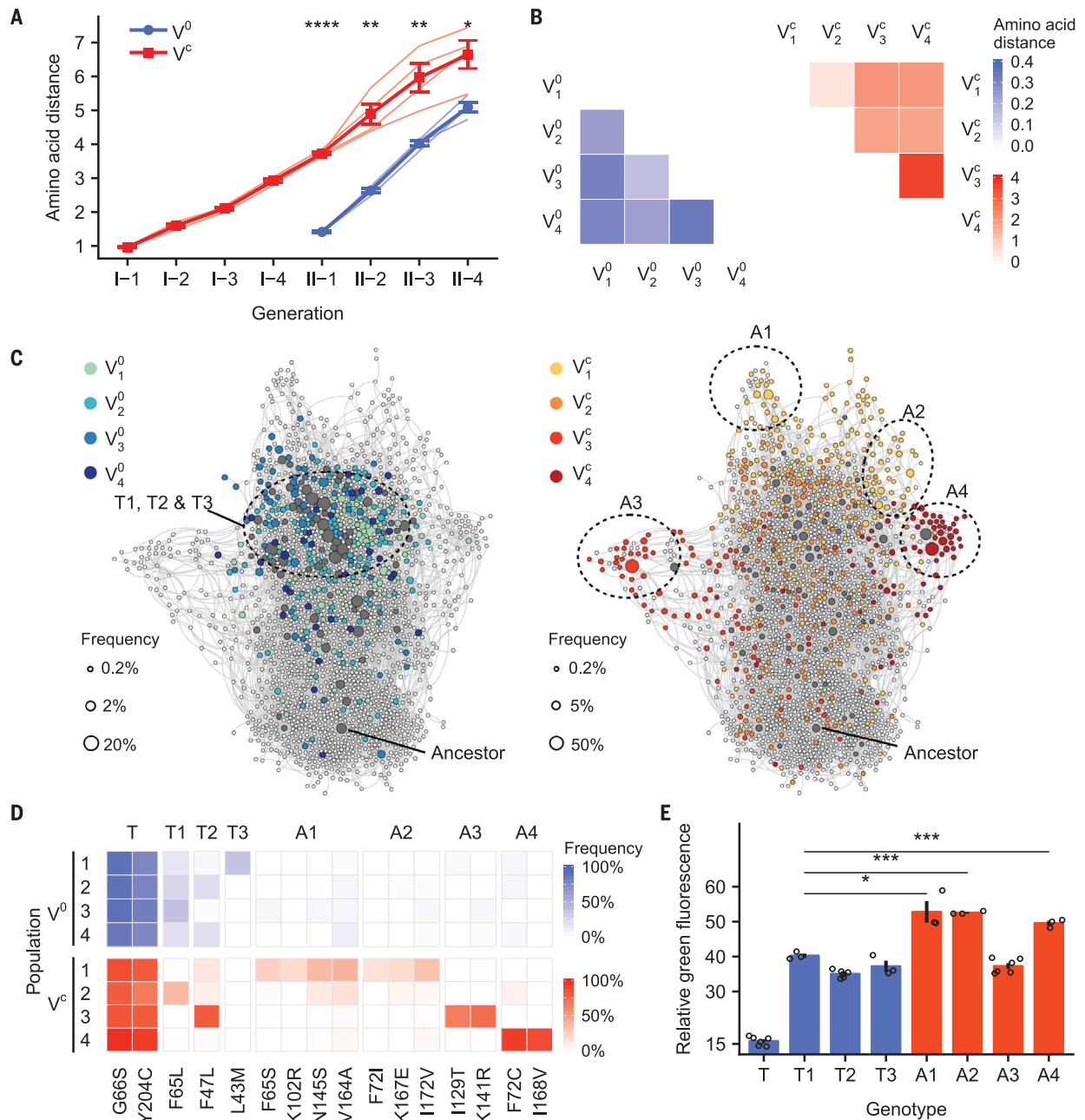


Fig. 3. Cryptic variation helps explore diverse high-fluorescence genotypes. (A) Number of amino acid changes per protein sequence based on genotyping hundreds of evolved variants in each population using SMRT sequencing. * $P < 0.05$; ** $P < 0.01$; **** $P < 0.0001$ (one-sided t tests). Thick lines indicate means for populations V^0 or V^c over four replicate populations (thin lines), and error bars denote SEM. (B) Average number of amino acid differences (at the evolutionary end point) between all protein sequences in the labeled populations (29). (C) Cryptic variation helps explore diverse genotypes. Each circle (node) represents a genotype that has been observed during evolution. An edge connects two genotypes if they differ in a single amino acid. Colored circles represent genotypes that exclusively occur in a single replicate population, where circle area (logarithmic scale) corresponds to genotype frequency. White and gray circles indicate genotypes that were not observed in populations at the end point or that were observed in at least two replicate populations at the end point, respectively. Sizes of gray circles correspond to the highest frequencies of the corresponding genotypes in those replicate populations. Dashed ovals

circumscribe each labeled high-fluorescence genotype, together with the genotypes composed of subsets of its constituent mutations. (D) The frequency of constituent mutations of typical and alternative genotypes in each replicate population at the evolutionary end point. The alternative genotypes A1, A2, A3, and A4 comprise the unique mutation combinations F65S/K102R/N145S/V164A, F72I/K167E/I172V, I129T/K141R, and F72C/I168V, respectively. In addition, each of these genotypes also harbors the mutations of T (G66S+Y204C), and genotype A3 also harbors the mutation F47L. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; E, Glu; F, Phe; G, Gly; I, Ile; K, Lys; L, Leu; N, Asn; R, Arg; S, Ser; T, Thr; and V, Val. Mutations (e.g., G66S) indicate the original residue (G66) and the residue created by the mutation (S). (E) Fold change in green fluorescence intensity of each typical (blue) or alternative (red) genotype relative to the ancestral YFP. [Note that here A1 does not contain the mutation K102R, because K102R does not significantly improve green fluorescence (table S5).] Error bars denote SD ($n = 3$ or 6). * $P < 0.05$; *** $P < 0.001$ (one-sided t tests with Holm adjustments).

expected with mutation pressure alone during phase I (two-way analysis of covariance with Holm adjustment [$P = 9.11 \times 10^{-5}$]) (fig. S4B). Specifically, 93.3% (14 of 15) of the genetic variants that were involved in adaptive evolution during phase II were not subject to positive selection in phase I. These observations demonstrated that most genetic variation that

was adaptive in phase II accumulated cryptically during phase I.

Because the typical and alternative genotypes were also the genotypes with the highest green fluorescence in each replicate population at the evolutionary end point (Fig. 3E, figs. S5 and S6, and table S5), we wanted to identify the accessible evolutionary paths to

these genotypes (fig. S5) (29). Each step on such a path involves a single point mutation, and we distinguish two kinds of steps, an accessible mutational step that increases green fluorescence significantly and an inaccessible step that does not. We call a path inaccessible if it contains at least one inaccessible step. We first engineered all mutations leading to

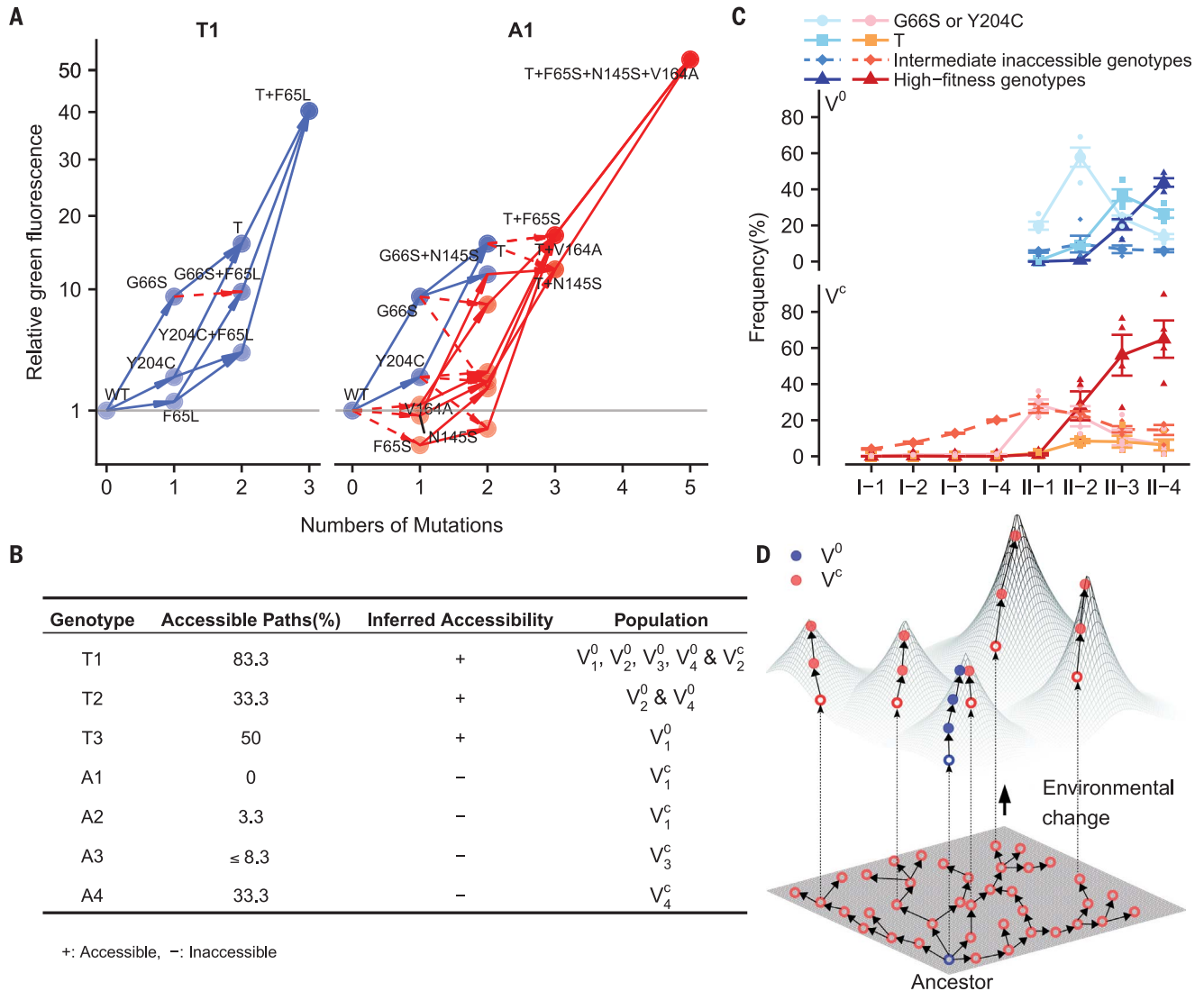


Fig. 4. Cryptic variation enables the exploration of alternative high-fluorescence genotypes. (A) Accessibility of mutational paths to two representative genotypes, the typical genotype T1 and the alternative genotype A1. [Note that the mutation K102R is not shown because it does not significantly improve A1's green fluorescence (table S5).] Blue solid lines indicate an accessible mutational step, which increases green fluorescence significantly, and dashed lines indicate an inaccessible step, which does not increase green fluorescence significantly. Solid red lines indicate a conditionally accessible step that significantly increases green fluorescence in the genetic background where it occurs but where the ancestral YFP must first experience one or more inaccessible steps to create this kind of genetic background. We call a path inaccessible if it contains at least one inaccessible step, and we consider a difference in green fluorescence between genotypes significant

if $P < 0.05$ (two-sided t test with Holm adjustment). (B) Percentages of accessible mutational paths to typical genotypes and to alternative genotypes, as well as accessibility inferred from mutation rates and genotype frequencies (29). The right-most entry indicates which populations harbored the genotype. (C) Evolutionary trajectories as indicated by frequency changes of mutants G66S and Y204C, genotype T, of all high-fitness genotypes that had significantly higher green fluorescence than genotype T (Fig. 3E and fig. S5), as well as those of all intermediate genotypes (averaged) leading to these high-fitness genotypes that are inaccessible through selection for green fluorescence alone. Error bars denote SEM ($n = 4$). Each circle indicates data from one replicate population. (D) An illustration of how cryptic genetic variation can accelerate adaptation and provide access to diverse adaptive peaks (see text for details).

each of the typical genotypes (T1 to T3) into the ancestor and measured their green fluorescence to determine path accessibility. No less than one-third of paths to the typical genotypes are accessible (Fig. 4, A and B).

We then engineered and analyzed the mutations leading to the alternative genotypes A1 to A4 and found that these genotypes are much less accessible (Fig. 4, A and B, and figs. S5 and S6) (29). For example, genotype A2, which had the highest green fluorescence among all typical and alternative genotypes, can be accessed by only 3.3% of all mutational paths (Fig. 4B and fig. S6) (29). The reason is that two mutations in this genotype (F72I and I172V) enhance green fluorescence only after the arrival of two other mutations (G66S and Y204C), and the remaining constituent mutation (K167E) only becomes beneficial once the four other mutations have arrived. An even more extreme example is the alternative genotype A1, because no path to it is accessible. Four of its six constituent mutations do not increase green fluorescence in the wild-type background or in the presence of the remaining two mutations, which suffices to block each path (Fig. 4A and fig. S5) (29).

We next examined our sequence data to study the order of mutations by which evolving populations approach those high-fitness genotypes that have the highest frequency in any one generation and population (Fig. 4C). All four V^0 populations followed similar mutational paths to each of the three typical genotypes, T1 to T3 (Fig. 4C). They first acquired either mutation G66S or Y204C, which arose to an average frequency of 20.1% after generation 1 of phase II (II-1). Next evolved the genotype T (G66S+Y204C), which reached a frequency of 9.2% one generation later. After that arose genotypes T1, T2, and T3, which incorporate the additional mutations F65L, F47L, and L43M, respectively. They show even higher green fluorescence (fig. S5) and reached a frequency of 18.2% in generation 3 (Fig. 4C). Inaccessible genotypes play no major role in these evolutionary dynamics, because their frequency remains low in populations V^0 (Fig. 4C).

These evolutionary dynamics differ from those observed in populations V^C with cryptic variation. Here, intermediate genotypes that would be inaccessible during selection for green fluorescence steadily increased in frequency before such selection started. At the end of phase I, the collection of all such genotypes had already reached a frequency of 16.9% in V^C populations (Fig. 4C). These otherwise-inaccessible intermediate genotypes served as stepping-stones toward high green fluorescence in phase II, as shown by a transition from inaccessible intermediate genotypes to high-fitness genotypes early in phase II (Fig. 4C). Specifically, inaccessible intermediate genotypes reached a frequency of 26.5% in V^C populations in the first generation of phase II, which enabled a rapid increase in

the frequency of high-fitness genotypes to 28.7% only one generation later.

We then studied the evolutionary dynamics leading to specific typical and alternative genotypes, which provided further support for our hypothesis that cryptic variation can help explore alternative trajectories and peaks (Fig. 4B and fig. S7) (29). One example involves the evolution of alternative genotype A4 in population $V_{4,g}^C$, where the occurrence of a crucial intermediate genotype (Y204C+F72C+I168V) that was inaccessible under selection for green fluorescence had been facilitated by phase I. Specifically, we detected all three constituent mutations of this intermediate genotype at the end of phase I (table S6). The intermediate genotype itself appeared in the first generation of phase II, and genotype A4, which harbors one additional mutation (G66S), had already attained a frequency of 27.9% one generation later (fig. S7) (29).

Discussion

Taken together, our observations indicate that cryptic variation not only helps populations traverse otherwise inaccessible trajectories to high-fitness genotypes but also helps them access diverse high-fitness genotypes (Fig. 4D). When populations are exposed to stabilizing selection while they diverge from an ancestral genotype (Fig. 4D, blue open circle), they may accumulate cryptic genetic variation (Fig. 4D, red open circles). An environmental change that alters selection pressure can alter the adaptive landscape on which such populations evolve and create new fitness peaks (Fig. 4D, upper panel). If a population without variation starts to adapt to a new environment, it may reach a nearby fitness peak (Fig. 4D, solid blue circles in upper panel), but it will not climb other, higher peaks if reaching such peaks requires traversing inaccessible low-fitness genotypes. In contrast, populations with cryptic variation may reach these peaks (Fig. 4D, solid red circles in upper panel) if the necessary genetic stepping-stones have arisen before the environmental change.

Sign epistasis, where a DNA mutation can change the sign of its effect on fitness from beneficial to detrimental in the presence of other mutations, is a source of complex, multi-peaked topographies in adaptive landscapes (23, 30, 31). Sign epistasis is widespread in proteins and RNA (26, 30, 32–35). It can create fitness plateaus or valleys, which constrain the order in which adaptive mutations occur, and slow down or prevent the ascent of peaks in an adaptive landscape (25, 30–32). How evolving populations can overcome such obstacles is a central question in evolutionary biology (1, 36, 37). Computational or phylogenetic studies suggest that fitness valley crossing may be a common phenomenon (38, 39). In addition, theory proposes that cryptic genetic variation can facilitate fitness valley crossing (40, 41), but experimental evidence

is still wanting. Our experiments demonstrate how cryptic variation can facilitate this process. During stabilizing selection on an ancestral phenotype, a population can accumulate not only neutral mutations but also mutations that would be deleterious when selection favors a new phenotype. Such mutations can become stepping-stones toward the new phenotype. For example, one of our populations with cryptic variation (and none of those without) reached the high-fluorescence genotype A1 because multiple stepping-stone variants had arisen during earlier stabilizing selection for yellow fluorescence. Furthermore, because different populations stochastically accumulate different cryptic variants, such cryptic variation creates stochasticity and historical contingency that not only reduces the predictability and reproducibility of evolution but also can uniquely solve evolution's problems (Fig. 3, C and D).

Evolving populations of whole organisms with different initial fitness sometimes converge on similar fitness (42, 43). The pervasive epistasis between amino acid-changing mutations makes such convergence less likely for our protein populations, because proteins with different, epistatically interacting mutations often evolve distinct genotypes and, in consequence, achieve quite different fitness (28, 44).

When directed evolution relies on large population sizes (45), it can lead to repeatable evolutionary outcomes that cannot be further improved (46, 47). Small populations with strong neutral drift can be more effective (48, 49), but small populations will also accumulate limited diversity. In contrast, large populations subject to stabilizing selection not only will accumulate substantial cryptic variation but also may uncover different high-fitness phenotypes during subsequent directional selection on a new phenotype. Our observations call for experiments where many and large populations have evolved in parallel, first to accumulate cryptic variation during stabilizing selection on an existing phenotype and then to find different novel phenotypes during directional selection, such as for a previously unknown biomolecule. This approach may work when conventional directed evolution fails.

Directed evolution experiments require high mutation rates to allow observation of adaptive evolution on laboratory time scales. Such high mutation rates can generate multiple beneficial mutations that compete with each other through clonal interference. In consequence, only the most strongly beneficial mutations may survive, leading to “greedy” adaptation and repeatable outcomes (50–52), as in our V^0 populations (Fig. 3). In addition, high mutation rates can also increase the chance that deleterious mutations can hitchhike with beneficial mutations, which can facilitate fitness valley crossing (52). However, we did not observe such valley crossing

in our populations V^0 (fig. S7), perhaps because of the stringent selection in our experiment, where only 0.01% individuals survived every generation during phase II (Fig. 1). Such selection may purge deleterious mutations before compensatory mutations can arise (51). Consistent with this hypothesis, intermediate inaccessible genotypes stayed at a low frequency in V^0 populations, whereas they steadily decreased in V^C populations during phase II evolution (Fig. 4C).

In sum, our results illustrate why cryptic variation can help populations not only overcome obstacles to adaptive evolution but also find multiple routes around such obstacles. The sign epistasis that creates such obstacles is involved in processes as different as the evolution of sexual reproduction (53), the divergence and reproductive isolation of species (54), and the development of human diseases (55). By breaking its constraints, cryptic variation may thus have far-reaching effects on many biological processes.

REFERENCES AND NOTES

1. A. B. Paaby, M. V. Rockman, *Nat. Rev. Genet.* **15**, 247–258 (2014).
2. G. Gibson, I. Dworkin, *Nat. Rev. Genet.* **5**, 681–690 (2004).
3. K. McGuigan, C. M. Sgrò, *Trends Ecol. Evol.* **24**, 305–311 (2009).
4. N. Frankel *et al.*, *Nature* **466**, 490–493 (2010).
5. A. Burga, M. O. Casanueva, B. Lehner, *Nature* **480**, 250–253 (2011).
6. A. Bergman, M. L. Siegal, *Nature* **424**, 549–552 (2003).
7. C. H. Waddington, *Evolution* **7**, 118–126 (1953).
8. C. H. Waddington, *Evolution* **10**, 1–13 (1956).
9. C. Lai, R. F. Lyman, A. D. Long, C. H. Langley, T. F. Mackay, *Science* **266**, 1697–1702 (1994).
10. I. Dworkin, A. Palsson, K. Birdsall, G. Gibson, *Curr. Biol.* **13**, 1888–1893 (2003).
11. S. L. Rutherford, S. Lindquist, *Nature* **396**, 336–342 (1998).
12. S. L. Rutherford, *BioEssays* **22**, 1095–1105 (2000).
13. F. Duveau, M.-A. Félix, *PLOS Biol.* **10**, e1001230 (2012).
14. K. McGuigan, N. Nishimura, M. Currey, D. Hurwit, W. A. Cresko, *Evolution* **65**, 1203–1211 (2011).
15. I. Dworkin *et al.*, *Genetics* **181**, 1065–1076 (2009).
16. L. E. Cowen, S. Lindquist, *Science* **309**, 2185–2189 (2005).
17. E. Rigato, G. Fusco, *J. Exp. Zool. B Mol. Dev. Evol.* **326**, 31–37 (2016).
18. N. Tokuriki, D. S. Tawfik, *Nature* **459**, 668–673 (2009).
19. E. J. Hayden, E. Ferrada, A. Wagner, *Nature* **474**, 92–95 (2011).
20. N. Rohner *et al.*, *Science* **342**, 1372–1375 (2013).
21. C. Bank, S. Matuszewski, R. T. Hietpas, J. D. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 14085–14090 (2016).
22. C. A. Macken, A. S. Perelson, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6191–6195 (1989).
23. F. J. Poelwijk, D. J. Kiviet, D. M. Weinreich, S. J. Tans, *Nature* **445**, 383–386 (2007).
24. I. S. Povolotskaya, F. A. Kondrashov, *Nature* **465**, 922–926 (2010).
25. Z. R. Sailer, M. J. Harms, *PLOS Comput. Biol.* **13**, e1005541 (2017).
26. T. N. Starr, L. K. Pictou, J. W. Thornton, *Nature* **549**, 409–413 (2017).
27. E. R. Jerison *et al.*, *eLife* **6**, e27167 (2017).
28. F. Baier *et al.*, *eLife* **8**, e40789 (2019).
29. See supplementary materials.
30. J. Domingo, G. Diss, B. Lehner, *Nature* **558**, 117–121 (2018).
31. A. C. Palmer *et al.*, *Nat. Commun.* **6**, 7385 (2015).
32. T. N. Starr, J. M. Flynn, P. Mishra, D. N. A. Bolon, J. W. Thornton, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4453–4458 (2018).
33. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, *Science* **312**, 111–114 (2006).
34. C. Li, W. Qian, C. J. Maclean, J. Zhang, *Science* **352**, 837–840 (2016).
35. A. I. Podgornaia, M. T. Laub, *Science* **347**, 673–677 (2015).
36. D. M. Weinreich, L. Chao, *Evolution* **59**, 1175–1182 (2005).
37. D. B. Weissman, M. W. Feldman, D. S. Fisher, *Genetics* **186**, 1389–1410 (2010).
38. M. V. Meer, A. S. Kondrashov, Y. Artzy-Randrup, F. A. Kondrashov, *Nature* **464**, 279–282 (2010).
39. A. S. Kondrashov, S. Sunyaev, F. A. Kondrashov, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 14878–14883 (2002).
40. M. V. Trotter, D. B. Weissman, G. I. Peterson, K. M. Peck, J. Masel, *Evolution* **68**, 3357–3367 (2014).
41. J. Masel, *Genetics* **172**, 1985–1991 (2006).
42. S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, *Science* **344**, 1519–1522 (2014).
43. H. Teotónio, I. M. Chelo, M. Bradić, M. R. Rose, A. D. Long, *Nat. Genet.* **41**, 251–257 (2009).
44. M. L. M. Salverda *et al.*, *PLOS Genet.* **7**, e1001321 (2011).
45. M. S. Packer, D. R. Liu, *Nat. Rev. Genet.* **16**, 379–394 (2015).
46. P. A. Romero, F. H. Arnold, *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).
47. N. Tokuriki *et al.*, *Nat. Commun.* **3**, 1257 (2012).
48. R. D. Gupta, D. S. Tawfik, *Nat. Methods* **5**, 939–942 (2008).
49. W. S. Smith, J. R. Hale, C. Neylon, *BMC Res. Notes* **4**, 138 (2011).
50. M. L. M. Salverda, J. Koomen, B. Koopmanschap, M. P. Zwart, J. A. G. M. de Visser, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 12773–12778 (2017).
51. T. Vogwill, R. L. Phillips, D. R. Gifford, R. C. MacLean, *Proc. Biol. Sci.* **283**, (2016).
52. I. G. Szendro, J. Franke, J. A. G. M. de Visser, J. Krug, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 571–576 (2013).
53. J. A. G. M. de Visser, S. F. Elena, *Nat. Rev. Genet.* **8**, 139–149 (2007).
54. D. C. Presgraves, *Nat. Rev. Genet.* **11**, 175–180 (2010).
55. H. J. Cordell, *Nat. Rev. Genet.* **10**, 392–404 (2009).
56. J. Zheng, J. Payne, *Cryptic genetic variation*. GitHub (2019); <https://doi.org/10.5281/zenodo.3233964>.

ACKNOWLEDGMENTS

We acknowledge the experimental support of the flow cytometry facility and the functional genomics center at the University of Zurich. We thank N. Guo and H. E. L. Lischer for help with data visualization and SMRT sequencing data analysis, respectively. We thank J. Duarte and Y. Schaerli for help with cell sorting and flow cytometric analysis. **Funding:** This project has received funding from the European Research Council under grant agreement 739874. We would also like to acknowledge support by Swiss National Science Foundation grant 31003A_172887 (A.W.) and PPOOP3_170604 (J.L.P.). **Author contributions:** J.Z. and A.W. designed the experiments. J.Z. performed the experiments. J.Z., J.L.P., and A.W. all contributed to data analysis. J.Z. and A.W. wrote the paper. All authors read and edited the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. SMRT sequencing data are available at DDBJ/EMBL/GenBank under the accession KCZY00000000. Custom code used in this study is available in a public GitHub repository (56).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/365/6451/347/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S24
References (57–66)

1 March 2019; accepted 6 June 2019
10.1126/science.aax1837

Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks

Jia Zheng, Joshua L. Payne and Andreas Wagner

Science **365** (6451), 347-353.
DOI: 10.1126/science.aax1837

Cryptic alleles make a bridge for adaptation

Protein function is generally constrained by selective parameters that can inhibit evolutionary potential. It has thus been difficult to determine how novelties arise. Zheng *et al.* allowed bacterial populations to accumulate mutations and then used directed evolution to evolve green fluorescent protein function from a gene that expressed yellow fluorescent protein (see the Perspective by Lee and Marx). Protein alternatives could evolve in cases where cryptic alleles—selectively neutral or mildly deleterious genetic variants with no apparent phenotypic differences—were present in the population. Thus, cryptic alleles provide an evolutionary bridge between diversity and selection and facilitate the generation of novel adaptations.

Science, this issue p. 347; see also p. 318

ARTICLE TOOLS

<http://science.sciencemag.org/content/365/6451/347>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2019/07/24/365.6451.347.DC1>

RELATED CONTENT

<http://science.sciencemag.org/content/sci/365/6451/318.full>

REFERENCES

This article cites 66 articles, 18 of which you can access for free
<http://science.sciencemag.org/content/365/6451/347#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)