Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. — **Source link** ⧉

Jia Zheng, Jia Zheng, Joshua L. Payne, Joshua L. Payne ...+3 more authors

Institutions: Swiss Institute of Bioinformatics, University of Zurich, ETH Zurich, Santa Fe Institute

Topics: Genetic variation and Adaptation

Related papers:

- Cryptic genetic variation: evolution's hidden substrate

- The causes of evolvability and their evolution.

- Darwinian evolution can follow only very few mutational paths to fitter proteins.

- Uncovering cryptic genetic variation.

- Robustness and evolvability: a paradox resolved.

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

Zheng, Jia ; Payne, Joshua L ; Wagner, Andreas

Abstract: Cryptic genetic variation can facilitate adaptation in evolving populations. To elucidate the underlying genetic mechanisms, we used directed evolution in <jats:italic>Escherichia coli</jats:italic> 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.

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

Authors: Jia Zheng[1,2], Joshua L. Payne[2,3] & Andreas Wagner[*1,2,4]

Affiliations:

[1]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.

[2]Swiss Institute of Bioinformatics, Quartier Sorge-Batiment Genopode, Lausanne, Switzerland.

[3] Institute of Integrative Biology, ETH, Zurich, Switzerland.

[4] The Santa Fe Institute, Santa Fe, New Mexico, USA.

*Correspondence to: Andreas Wagner (email: andreas.wagner@ieu.uzh.ch).

Abstract:

Cryptic genetic variation can facilitate adaptation in evolving populations. To elucidate the underlying genetic mechanisms, we used directed evolution in *E. coli* to accumulate variation in populations of yellow fluorescent proteins, and then evolved these proteins towards the phenotype of green fluorescence. Populations with cryptic variation evolved adaptive genotypes with greater diversity and higher fitness than populations without cryptic variation, which converge 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.

One Sentence Summary: Cryptic genetic variation opens paths to selectively inaccessible regions of protein sequence space.

Main Text:

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., non-additively), which can render adaptive landscapes rugged and multi-peaked (*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 regions 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, 'C' for cryptic) to four rounds ('generations') of directed evolution subject to stringent stabilizing selection to maintain yellow fluorescence

(phase I; Figs. 1, S1). Specifically, we evolved a population of ~5×10^6 YFP variants in each generation and in each replicate population, which we subjected to PCR mutagenesis (0.84 amino-acid changing mutations per YFP molecule per generation, Tables S1, 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 towards green fluorescence (Fig. 1). As controls, we also subjected four populations (called V^0, for zero initial cryptic variation) that started from identical ancestral YFP molecules to four generations of evolution towards green fluorescence. We then compared the change in green fluorescence intensity during phase II in populations V^C with that of the control populations V^0. Populations V^C reached significantly higher green fluorescence during three out 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 (512nm) emission peak than populations V^0 (Fig. 2B). At the evolutionary endpoint, three of four V^C populations showed significantly greater green fluorescence than the four V^0 populations (Two-way ANOVA: $F_{7,16}$=46.5, P=1.99×10^{-9}; post hoc Tukey's test, P<0.05 for V^C replicate 1, 2 and 4 relative to the four V^0 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 ~500-1000 evolved variants for each replicate population

and for each generation (Table S4). We first noticed that V^C populations were more diverse than V^0 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 (Figs. 3B, 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) swept through all replicate V^0 and V^C populations, with two other mutations (F65L or F47L) achieving high or medium frequency (>10%) in two or more V^0 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 endpoint, 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, and L43M. We refer to these genotypes as T1, T2, and T3 or *typical* genotypes (Fig. 3C, D), and to the combination of G66S and Y204C mutation as genotype T.

Populations V^C evolved differently from populations V^0. First, seventeen *alternative* mutations attained a frequency of more than 10% in V^C populations but in none of the V^0 populations (except the mutation V164A which reached a frequency of 10.9% in V_4^0; Fig. S3A). Also, typical genotypes dominated only one replicate V^C population (number 2), in contrast to their importance in V^0 populations. The remaining populations were dominated by one or two of four other, alternative genotypes (A1-A4), which contain some combination of eleven alternative

mutations in the genetic background T (Fig. 3C, D). We measured the green fluorescence intensity of the three typical genotypes as well as of the four alternative genotypes (*29*). Three of the alternative genotypes exhibited greater green fluorescence than all typical genotypes (Fig. 3E).

5   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 had significantly higher green fluorescence than all three typical genotypes.

Because V^C populations evolved faster in phase II than V^0 populations (Fig. 2A), we suspected

10   that some of their adaptive mutations or genotypes accumulated in phase I. We thus studied the phase I evolutionary dynamics of the four typical mutations and eleven alternative mutations (Fig. S4A). All fifteen mutations were already present above our phase I detection limit of 0.064-0.16 percent(*29*), and eleven of the fifteen 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

15   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

20   expected by mutation pressure alone during phase I (Two-way Ancova with Holm adjustment, $P=9.11\times10^{-5}$; Fig. S4B). Specifically, 93.3 percent (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 endpoint (Figs. 3E, S5, S6; 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 each of the typical genotypes (T1-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. 4A, B).

We then engineered and analyzed the mutations leading to the alternative genotypes A1-A4, and found that these genotypes are much less accessible (Figs. 4A, B, S5, 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 (Figs. 4B, 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 decrease green fluorescence both in the wildtype background, and also in the presence of the remaining two mutations, which suffices to block each path (Figs. 4A, S5; (*29*)).

We next examined our sequence data to study the order of mutations by which evolving populations approached 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-T3 (Fig. 4C). They first acquired either mutation G66S or Y204C, which arose to an average frequency of 20.1 % after the first generation 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 three (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 already 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 towards 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 provides further support for our hypothesis that cryptic variation can help explore

alternative trajectories and peaks (Figs. 4B, S7; (*29*)). One example involves the evolution of

alternative genotype A4 in population V_4^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 helps populations not only

traverse otherwise inaccessible trajectories to high fitness genotypes, it 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 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 novel phenotype. Such mutations can become stepping stones towards the novel 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 that can also uniquely solve evolution's problems (Fig. 3C, 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 tend to evolve distinct genotypes and, in consequence, achieve quite different fitness (*28, 44*).

When directed evolution relies on large populations 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 will not only accumulate substantial cryptic variation, but may also uncover different high fitness phenotypes during subsequent directional selection on a new phenotype. Our observations call for experiments where *many and large* populations are 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 novel biomolecule. This approach may work when conventional directed evolution fails.

Directed evolution experiments require high mutation rates in order to observe 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, large 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 cannot only help populations 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 (1956).

9. C. Lai, R. Lyman, A. Long, C. Langley, T. 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**, 1065–76 (2009).

16. L. E. Cowen *et al.*, *Science* **309**, 2185–2189 (2005).

17. E. Rigato, G. Fusco, *J. Exp. Zool. Part 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**, 6191–5 (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. Picton, 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 supplemental 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**, 837–40 (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**, 1389–410 (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**, 14878–83 (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**, 1985–91 (2006).

42.   S. Kryazhimskiy, D. P. Rice, E. R. Jerison, M. M. Desai, *Science* **344**, 1519–1522 (2014).

43.   H. Teotó Nio *et al.*, *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, *Proceedings. Biol. Sci.* **283**, 20160749 (2016).

52.   I. G. Szendro, J. Franke, J. A. G. M. de Visser, J. Krug, *Proc. Natl. Acad. Sci.* **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.   A. Khlebnikov, K. A. Datsenko, T. Skaug, B. L. Wanner, J. D. Keasling, *Microbiology* **147**, 3241–3247 (2001).

57.   D. J. Warren, *Anal. Biochem.* **413**, 206–207 (2011).

58.   C. Stanton *et al.*, *Nat Chem Biol* **10**, 99–105 (2014).

59.   S. Bershtein, M. Segal, R. Bekerman, N. Tokuriki, D. S. Tawfik, *Nature* **444**, 929–932 (2006).

60.   S. Bratulic, F. Gerber, A. Wagner, *Proc. Natl. Acad. Sci. U. S. A.* **112**, 12758–63 (2015).

61.   *Pacific Biosciences DevNet. Available at http://www.smrtcommunity.com/DevNet (2015)*.

62.   M. J. Chaisson, G. Tesler, *BMC Bioinformatics* **13**, 238 (2012).

63.   H. Li *et al.*, *Bioinformatics* **25**, 2078–2079 (2009).

64.   K. E. Kim *et al.*, *Sci. Data 1*, 140045 (2014).

65.   W. Shao *et al.*, *J. Virol. Methods* **203**, 73–80 (2014).

https://github.com/ZhengJiaCode/Cryptic_genetic_variation

(https://doi.org/10.5281/zenodo.3233964).

Legends

Fig. 1. Experimental evolution of yellow fluorescent protein. In phase I, we subjected four replicate populations of yellow fluorescent protein (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, λex=488nm and λem=530±15nm, (*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 (λex=405nm and λem=525±25nm, (*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 intensity relative to the ancestral YFP during phase II evolution (*29*). Error bars represent one 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 (405nm) 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 512nm (green vertical dashed line). **(C)** Fold-change of green fluorescence intensity relative to the ancestral YFP for each replicate population at the evolutionary endpoint. 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 line), and error bars denote SEM. **(B)** Average number of amino-acid differences (at the evolutionary endpoint) 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 grey circles indicate genotypes that were not observed in populations, or that were observed in at least two replicate populations at the endpoint, respectively. Sizes of grey 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 endpoint. 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 T (G66S+Y204C), and genotype A1 also harbors the mutation F47L. **(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). * $P<0.05$, ** $P<0.01$, *** $P<0.001$ and **** $P<0.0001$ (One-sided t-tests with Holm adjustments).

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

5   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

10   contains at least one inaccessible step, and consider a difference in green fluorescence between genotypes as significant if $P<0.05$ (Two-sided t-test with Holm adjustment). **(B)** Percentage 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

15   frequency changes of mutants G66S and Y204C, of genotype T (G66S+Y204C), of all high-fitness genotypes that had significantly higher green fluorescence than genotype T (Figs. 3E, S5), as well as 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)** A schematic illustration of how

20   cryptic genetic variation can accelerate adaptation and provide access to diverse adaptive peaks (see text for details).

Supplemental materials:

Materials and Methods

Supplementary Text

Figs. S1 to S12

5   Tables S1 to S24

References (56-65)

Phase I (generations 1–4)

Phase II (generations 1–4)

V^c YFP

Transform

Select

Mutate

20%

Frequency

Yellow fluorescence

Frequency

Green fluorescence

0.01%

V^0 YFP

A

Relative fluorescence

** (at II-1), * (at II-2), ** (at II-4)

- V⁰ — Green fluorescence: V^0
- Vᶜ — Green fluorescence: V^c
- V⁰ — Yellow fluorescence: V^0
- Vᶜ — Yellow fluorescence: V^c

Generation: II-1, II-2, II-3, II-4

B

Normalized green emission

V^0 Generation: II-1, II-2, II-3, II-4

V^c Generation: II-1, II-2, II-3, II-4

Wavelength(nm)

C

Generation: II-4

Population: 1, 2, 3, 4

- V^0
- V^c

Relative green fluorescence

A

T1 / A1

Relative green fluorescence vs. Numbers of Mutations

T1 panel labels: T+F65L, T, G66S, G66S+F65L, Y204C+F65L, Y204C, WT, F65L

A1 panel labels: T+F65S+N145S+V164A, T+F65S, G66S+N145S, T, T+V164A, G66S, T+N145S, Y204C, WT, V164A, N145S, F65S

B

Genotype	Accessible Paths(%)	Inferred Accessibility	Population
T1	83.3	+	$V_1^0, V_2^0, V_3^0, V_4^0$ & V_2^c
T2	33.3	+	V_2^0 & V_4^0
T3	50	+	V_1^0
A1	0	−	V_1^c
A2	3.3	−	V_1^c
A3	≤ 8.3	−	V_3^c
A4	33.3	−	V_4^c

+: Accessible, −: Inaccessible

C

Legend: G66S or Y204C; T; Intermediate inaccessible genotypes; High−fitness genotypes

V^0 and V^c panels: Frequency(%) vs. I−1, I−2, I−3, I−4, II−1, II−2, II−3, II−4

D

V^0 (blue), V^c (red)

Environmental change

Ancestor

Science

AAAS
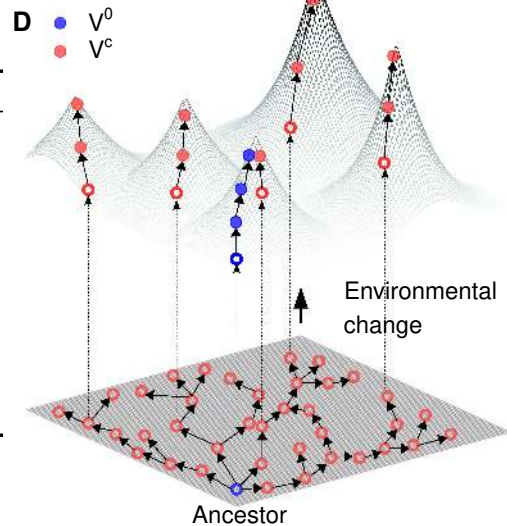
Supplementary Materials for

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

Jia Zheng, Joshua L. Payne & Andreas Wagner
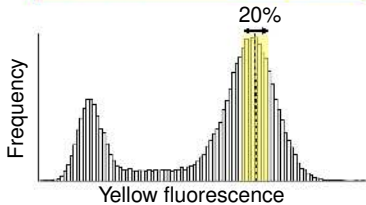Correspondence to: andreas.wagner@ieu.uzh.ch

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S24

Materials and Methods

Strains and Plasmids

We used *E. coli* strain BW27783 (CGSC 12119) for preparing YFP mutant libraries, which enables the homogeneous expression of the arabinose-inducible araBAD promoter in *E. coli* (*56*). We used the plasmid pBAD202/D-TOPO® (K4202-01, Invitrogen) as the vector for YFP evolution. It carries an arabinose-inducible araBAD promoter and a Kanamycin resistance marker.

Preparation of electro-competent cells

We prepared electro-competent cells using glycerol/mannitol step centrifugation (*57*). Briefly, we grew *E. coli* strain BW27783 in 5 mL SOB medium at 37°C and 250 rpm overnight. We transferred 3 mL culture into 300 mL SOB medium the next morning and continued to incubate the transferred culture at 37°C and 250 rpm until its OD_{600} had reached a value between 0.4 and 0.6 (optical path length: 1 cm, 2-4 hours). We cooled the culture on ice for 15 min and collected cells at 4°C by centrifuging at 1,500 g for 15 min. We used 60 mL ice-cold ddH_2O to suspend the cells and distributed t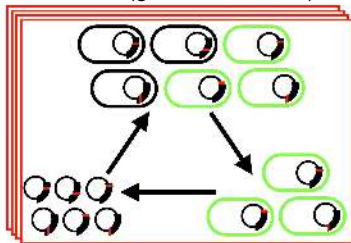hem into three 50 mL tubes. Then we slowly added 10 mL ice-cold glycerol/mannitol solution (20% glycerol (w/v) and 1.5% mannitol (w/v)) to the bottom of each tube by using a 10 mL pipette. We centrifuged the tubes at 1,500 g and 4°C for 15 min in a centrifuge (Eppendorf 5810/5810 R) by setting acceleration/deceleration to zero. We removed the supernatant and suspended the cells in 3.0 mL ice-cold glycerol/mannitol solution. We transferred 100 µL of the resulting suspensions into pre-cooled 1.5 mL tubes and incubated them in a dry ice-ethanol bath for ~1 min. Then we stored the suspensions at -80°C for transformation experiments.

Electro-transformation

We mixed 4 µL of ligation product with 100 µL of electro-competent BW27783 cells. We transferred the mixture into a 0.2 cm cuvette (EP202, Cell Projects, UK), and then used a Micropulser electroporator (Bio-Rad) set at EC3 (15k V/cm) for transformation. We immediately added 1 mL of pre-warmed SOC medium after electroporation and transferred the culture into a 10 mL tube. We incubated the culture for 1.5 h at 37°C with shaking at 220 rpm in a shaking incubator (INFORS HT, Switzerland), and used the recovered transformants for further experiments.

Construction of the expression plasmid

We designed the primers Fbad/Rbad (Table S8), which contain *Xho*I and *Hind*III restriction sites, to amplify the vector backbone of pBAD202/D-TOPO® by PCR. The resulting amplification product contained novel *Xho*I and *Hind*III restriction sites upstream of the original restriction sites *Nco*I and *Pme*I, and harbored a deletion of the DNA fragment between the latter two restriction sites (including the *Nco*I site). In addition, we designed the primers F203/R203 (Table S8) to amplify the coding region of YFP from the plasmid pAND (*58*) and to introduce *Xho*I and *Hind*III restriction sites upstream of its ATG start codon and downstream of its TGA stop codon, respectively.

To amplify both DNA molecules, we performed a 50 µL PCR reaction, consisting of 10 ng template plasmid, 400 µM dNTPs, 5.0 U Phusion Hot Start II High-Fidelity DNA Polymerase, 10 µL 5× Phusion HF Buffer and 400 nM each primer. We performed the PCR reaction for amplifying the vector backbone of pBAD202/D-TOPO® with the following program: 98°C/30 s; 10 cycles of 98°C/10 s, 54°C/15 s and 72°C/80 s; 15 cycles of 98°C/10 s, 68°C/15 s and 72°C/80 s; 72°C/5 min. And we used the following the program to amplify the coding region of YFP: 98°C/30 s; 10 cycles of 98°C/10 s, 66°C/15s and 72°C/30s; 15 cycles of 98°C/10 s and 72°C/30 s; 72°C/5 min.

To remove the template plasmid from the PCR product, we incubated the product with 10 U of *Dpn*I (R0176S, NEB) at 37°C for 2 h. To inactivate DNA polymerase, we added 0.6 U of proteinase K (EO0491, Thermo Scientific), followed by incubation at 50°C for 1h, and by 15 min of inactivation at 80°C. We then added 20 U *Xho*I/*Hind*III-HF (R0146L/R3104S, NEB), incubated at 37°C overnight, and inactivated the restriction enzymes at 80°C for 20 min. We used the QIAquick Gel Extraction Kit (Qiagen, Germany) to purify the digested vector backbones, and used the QIAquick PCR purification kit (Qiagen, Germany) to purify the digested YFP products.

We mixed 60 ng of purified YFP fragment, 100 ng of purified vector backbone, 10 U of T4 DNA ligase, and 2 µL of 10×Ligation buffer (M0202L, NEB) in a 20 µL of ligation reaction. We incubated the mixture at 20-22°C for ~16 h, followed by 10 min of inactivation at 65°C. To purify the ligation product, we first precipitated it by adding 1 µL of glycogen (R0551, Thermo Scientific), 50 µL of 7.5 M ammonium acetate (A2706-100ML, Sigma), 375 µL of ice-cold absolute ethanol, and 80 µL of ddH$_2$O. After incubation at -20°C for 20 min, we centrifuged the mixture at 18,000 g for 20 min. We washed the precipitate twice using 800 µL of cold ethanol (70%). After drying the precipitate using an Eppendorf concentrator 5301, we dissolved it in 10 µL of ddH$_2$O.

We used 4 µL of purified ligation product for transformation according to the protocol described above. Then we diluted 10 µL of an aliquot of transformed cells using saline, and plated the serially diluted aliquot on LB agar containing 25 µg/mL of kanamycin. After incubation at 37°C overnight, we choose three colonies for colony PCR to identify transformants harboring the appropriate vector construct. We extracted the plasmids of these colonies and subjected them to double restriction-enzyme digestion and Sanger sequencing. Finally, we chose one correctly constructed plasmid (named pBAD-EY) with YFP inserted for the following experiments.

Mutagenic PCR

For the evolution of YFP populations, we introduced mutations into the coding region of YFP using mutagenic PCR. In brief, we performed mutagenic PCR in a 100 µL volume with 10 ng of template plasmid, 400 µM of dNTPs (R0192, Thermo Scientific), 3 µM of 8-oxo-GTP/dPTP (Trilink Biotechnologies), 2.5 U of *Taq* DNA polymerase (M0267L, NEB), 10 µL of 10×ThermoPol buffer (M0267L, NEB), and 400 nM of each primer (MutafpF/MutafpR, Table S8). We performed 25 cycles of PCR, as follows: 95°C/30 s; 25 cycles of 94°C/20 s, 46°C/30 s and 68°C/50 s; 68°C/5 min. We treated PCR products by following the procedure described in *Construction of the expression plasmid* to obtain linearized inserts (mutated YFP pools). We also

digested and purified the plasmid pBAD-EY following the procedure described in *Construction of the expression plasmid*. We then used 5 U of Antarctic Phosphatase (M0289S, NEB) to dephosphorylate the purified vector backbone and purified the dephosphorylated vector backbone using the QIAquick PCR purification kit. We followed the procedure described in *Construction of the expression plasmid* to perform the ligation reaction and purify the ligation products. The ligation products served as substrates for the *E.coli* transformation described above. SMRT sequencing indicated a mutation rate of 1.04 nucleotide changes per YFP molecule and generation of mutagenesis (Table S 1), and allowed us to estimate the extent of the A→G and T→C mutation bias (Table S9) that PCR mutagenesis is well-known to cause (*59*).

Evolution of YFP under selection for yellow fluorescence

We performed electro-transformation as described above, using 4 µL of the mutagenized YFP library ligated to the vector backbone. Then we added 10 mL of LB medium supplemented with 30 µg/mL of kanamycin to the recovered culture in a 50 mL tube, and incubated the culture at 37°C overnight, with shaking at 220 rpm. To estimate library size during every generation of directed evolution, we also sampled 100 µL of the recovered culture and plated the serially diluted aliquot (in saline) on LB agar with 25 µg/mL of kanamycin. This procedure yielded a library size of ~10^6 colonies. Subsequently, we transferred 1 mL of the overnight culture into 10 mL LB medium with 55 µg/mL of kanamycin, and incubated the cell culture at 37°C with shaking at 220 rpm until it had reached an OD_{600} of 1.0~1.5 (optical path length: 1 cm). Then we transferred 200 µL of culture to a 10 mL tube containing 2 mL LB medium (supplemented with 55µg/mL of kanamycin), and incubated further until the culture had reached an OD_{600} of ~0.6. We added 22 µL of 20% arabinose and incubated the culture at 37°C with shaking at 220 rpm to induce expression of YFP. After 12 h of incubation, we sampled 400 µL of the culture and centrifuged it at 8000 g and 4°C for 3 min. We washed cells once using 1 mL of cold PBS buffer and suspended them in 1 mL of cold PBS buffer for fluorescence-activated cell sorting (FACS).

We sorted cells at 4°C with an Aria III cell sorter (BD Biosciences), using the FITC channel (λ_{ex}=488nm, λ_{em}=530±15nm) and a sorting speed of ~4×10^4 events/s. We used the sort precision of 4-Way Purity to make sure that only droplets free of contaminating particles would be sorted. We selected cells according to the selection criteria described in Fig. 1. We note that the ability to select cells within a narrow phenotypic bandwidth (Fig.1) is a great advantage of fluorescent proteins for directed evolution. It comes from the ability to measure fluorescence in individual cells and would be difficult to match with other proteins, such as antibiotic resistance proteins.

For each replicate population, we collected 10^6 selected cells in ~1 mL of cold PBS buffer (4°C). We placed selected cells on ice to prevent cell proliferation or death before the subsequent steps. As the first of these steps, we transferred the selected cells into 1 mL of LB medium in a 10 mL tube, and incubated them at 37°C for 30 min without shaking. We sampled 5 µL of this culture and diluted it 200-fold in saline. We plated this diluted cell aliquot on LB agar (25µg/mL of kanamycin) to estimate the post-selection library size, which we did for every generation of directed evolution. We incubated the remainder of the culture at 37°C and 220 rpm for 2 h, and added 3 mL LB medium (supplemented with 50µg/mL of kanamycin) to each tube for further incubation overnight. We added 400 µL of 50% glycerol to 600 µL of overnight culture, and stored the mixture at -80°C for the fluorescence assays described further below. In addition, we sampled 2 mL of overnight culture and extracted the plasmids using the QIAprep spin miniprep

kit (Qiagen, Germany). We used the isolated plasmids as templates for the next mutation-selection cycle. After four mutation-selection cycles, we used the isolated plasmids as starting templates for evolution under selection for green fluorescence, and also for SMRT sequencing.

In addition to these experiments, we also evolved four replicate populations (dubbed $Nsel_{1-4}$), following the same procedure as for populations V^C except for collecting all generated variants in each replicate population and in each generation, regardless of their fluorescence. In comparison to populations V^C, populations $Nsel_{1-4}$ thus did not experience selection on yellow fluorescence. They served as controls to determine changes in mutant frequencies that occurred in the absence of selection.

Evolution of YFP under selection for green fluorescence

We performed electro-transformation as described above, using 4 µL of the ligation product between a mutagenized YFP library and the vector backbone. Specifically, we electro-transformed 4 µL of ligation product into 100 µL of electro-competent BW27783 cells and induced YFP expression by arabinose, following the same procedure as described above. We sorted fluorescing cells as described above, except that we used the AmCyan channel (λex=405 nm and λem=525±25 nm) and the criterion for selecting green fluorescing variants depicted in Fig.1. We used the sort precision of Single Cell Purity to make sure that only droplets free of contaminating particles would be sorted, and that only particles centered within the sorted drop would be deflected. During selection for green fluorescence, we selected 5×10^4 cells in the top 1% of green fluorescence intensity (measured as AmCyan-H). We re-grew the sorted cells and repeated the sorting process by selecting 10^4 cells in the top 1% of green fluorescence intensity. Just as for selection for yellow fluorescence, we used plasmids isolated from sorted cells as templates for the next mutation-selection cycle, and stored sorted cells at -80°C in preparation for the fluorescence assays described below. We used YFP variants isolated after each generation of evolution for SMRT sequencing.

Engineering YFP variants

We designed primers (Table S10) that carried the necessary mutations to engineer single mutants and some double mutants (F65L+G66S, F65S+G66S, G66S+F72C and G66S+F72I) by whole plasmid PCR. To engineer these individual mutations in the genetic background of G66S, Y204C, and G66T+Y204C, we performed whole plasmid PCR by using the variants carrying G66S, Y204C or G66T+Y204C as templates, employing appropriate primers (Table S10). A 40 µL PCR reaction consisted of 10 ng template plasmid, 0.8µL 10mM dNTPS, 0.4 µL Phusion Hot Start II High-Fidelity DNA Polymerase (F549L, Thermo Scientific), 8 µL Phusion HF buffer, 1.2 µL DMSO and 0.8 µL 10 µM primers. We performed the PCR reaction with the following program: 98°C/1min, 8 cycles of 98°C/20s, 68°C/20s and 72°C/2min, 20 cycles of 98°C/20s and 72°C/2min, 72°C/5min. We digested the PCR product with DpnI at 37°C for 2 hours to remove the template plasmid, followed by incubation at 80°C for 20min to inactivate the enzyme. Then we directly used 1 µL of the reaction for electroporation.

We used the primers G66Sf/Y204Cr and Y204Cr/G66Sf ("f" and "r" indicate the forward (f) and reverse (r) primers carrying the corresponding mutation) to amplify the plasmid pBAD-EY carrying the ancestral YFP, and assembled the resulting two DNA fragments by Gibson

PCR, which yielded the double mutant G66T+Y204C. We engineered the rest of the multiple mutants by Gibson PCR using the appropriate primers in Table S10. We used the same PCR reaction system and program as in whole plasmid PCR, except for the PCR extension time, which equaled 30s per 1 kb and varied from 10s to 2min, depending on the length of amplified DNA fragments. We digested the PCR product with DpnI at 37°C for 2 hours to remove the template plasmid, followed by column purification using the QIAquick PCR purification kit (Qiagen, Germany). We added 2 µL mixture of purified DNA fragments with overlapping ends to 2 µL 2×Gibson Assembly Master Mix (using the recommended amount of fragments in the protocol, E2611S, NEB), and incubated at 50°C for 20 min in a thermocycler. Then we used 1 µL of the reaction for electro-transformation.

After electro-transformation, we diluted the recovery culture ten-fold (for cells transformed with Gibson PCR product) or 100-fold (for cells transformed with whole plasmid PCR product). We plated the resulting dilution on LB agar with 25 µg/mL kanamycin (Sigma 60615), and incubated at 37°C overnight. We choose three or six colonies from each transformation to validate the construction by colony PCR and Sanger sequencing. We inoculated correctly engineered mutants into 2 mL LB with 50 µg/mL of kanamycin, and incubated at 37°C and 220 rpm overnight. We mixed 900 µL overnight cultures with 600 µL 50% glycerol, and stored the glycerol stock at -80°C for the further experiments.

Inducible expression of evolving populations and engineered variants

To induce the expression of YFP variants in evolving populations, we allowed the glycerol stock of sorted cells from each generation of directed evolution to thaw on ice, and then transferred 100 µL of this stock into 2 mL LB medium (supplemented with 33µg/mL of kanamycin). We incubated the resulting cell suspension at 37°C with shaking at 220 rpm overnight. Then we transferred 200 µL of the resulting culture into 2 mL LB medium (supplemented with 55 µg/mL of kanamycin) for further incubation at 37°C with shaking at 220 rpm for 4 hours. We transferred 20 µL of the culture into 200 µL LB medium (55 µg/mL of kanamycin) in a Thermo Scientific Nunc Edge 2.0 96-well plate (to minimize evaporation), and incubated it in a Tecan Infinite F200 microplate reader at 37°C (2 mm orbital shaking), until the culture had reached an OD_{600} of ~0.6 (optical path length: 1 cm). Subsequently, we added 4 µL of 10% arabinose to induce YFP expression, and continued the incubation for 16 hours at 37°C. After these 16 hours, we sampled the culture for fluorescence assays using flow cytometry and fluorescence emission scanning with a Tecan microplate reader. To induce the expression of YFP variants in engineered variants, we followed the same procedure, except that we incubated them in a microplate incubator at 800 rpm (To prevent evaporation, we sealed plates using Parafilm and removed it after 15 hours of incubation, and then continued the incubation for one hour).

Fluorescence assay using flow cytometry

Sixteen hours after induction of expression, we suspended a 40 µL sample of the culture in 160 µL of cold PBS buffer, transferred 20 µL of the suspension to 180 µL cold PBS buffer, and mixed thoroughly to measure yellow fluorescence at the original wavelength (λ_{ex}=488nm and λ_{em}=530±15nm) and green fluorescence at the new wavelength (λ_{ex}=405nm and λ_{em}=525±25nm) using flow cytometry. We performed the cytometry assay at room temperature using a Fortessa

cell analyzer (BD Biosciences) with a flow rate of ~3000 events/s. We performed three biological replicates for each assay and analyzed 10^4 cells for each replicate. We placed all samples on ice until we had finished the assay to prevent cell proliferation or death.

Fluorescence emission scan

To measure the emission spectrum of evolving populations in each generation of evolution, we sampled 40 µL of culture after 16 hours of expression induction, suspended the sample in 160 µL of cold PBS buffer, and mixed thoroughly. Subsequently, we measured fluorescence (i) at the excitation wavelength λ_{ex}=485nm for emission wavelengths varying between 520 nm and 560 nm, and (ii) at the excitation wavelength λ_{ex}=405 nm for emission wavelengths varying between 438 nm and 560 nm in a Tecan Infinite F200 microplate reader. We performed this fluorescence emission scan at room temperature (24-27°C) in 2 or 3 nm increments.

Flow Cytometry Data Analysis

We used FlowJo V10.4.2 (LLC) to perform flow cytometry data analysis. In brief, we used forward scatter height (FSC-H) versus side scatter height (SSC-H) density plots to select a homogenous cell population, and used side scatter area (SSC-A) versus side scatter height (SSC-H) density plots to exclude doublets. Finally, we used FITC-Height versus AmCyan-Height density plots to further exclude dust, cellular debris and micro-particles (see the details in Fig. S8). We used the resulting filtered data for calculating the mean fluorescence intensity of each biological replicate.

SMRT sequencing

We barcoded YFP variants of each replicate population for single molecule real-time (SMRT) sequencing using two-step PCRs, as described previously (60). We performed PCR amplification using high-fidelity Phusion DNA polymerase to decrease the incidence of mutations during this process. Briefly, we first performed a 12-cycle PCR to amplify YFP variants from each replicate population using primers FP_smrtF/pBAD849R (Table S11) in a 30 µL PCR reaction, consisting of 1 ng template plasmid, 400 nM of each primer, 400 µM dNTPs, 6 µL 5× Phusion HF Buffer, 0.3 µL 100% DMSO and 1.5 U Phusion Hot Start II High-Fidelity DNA Polymerase (F-549L, Thermo Scientific). We executed the PCR reaction with the following thermocycler program: 98°C/30 s, 12 cycles of 98°C/15 s, 65°C/15 s and 72°C/15 s, 72°C/1 min. We treated the above PCR product with 5 U DpnI and 5 U Exonuclease I (EN0581, Fermentas) at 37°C for 1 h to digest the template plasmid and the primers, and inactivated these enzymes at 85°C for 20 min. Then we used the resulting product as template for a barcoding PCR using the barcode-tagged primers BCXX and pBAD849R (Table S11). Each barcode-tagged primer harbors a unique 6 bp sequence. We performed the barcoding PCR in a 50 µL volume by mixing 2 µL of template, 400 nM primers BCXX and pBAD849r, 400 µM dNTPs, 2.5 U of Phusion Hot Start II High-Fidelity DNA Polymerase, and 10 µL of 5× Phusion HF Buffer. We executed the PCR amplification with the following thermocycler program: 98°C/30 s; 28 cycles of 98°C/15 s, 63°C/15 s and 72°C/15 s; 72°C/1 min.

We purified the barcode-tagged PCR products with a QIAquick PCR purification kit, and checked their quality and concentration using a UV-Vis spectrophotometer (NanoDrop, Thermo

Fisher Scientific), a Qubit 4 Fluorometer (Invitrogen), and agarose gel electrophoresis. To estimate potential errors that might occur during library preparation, we used the same procedure to amplify the ancestral YFP gene. After that, we pooled 20 ng DNA of each population from the same generation of evolution in a single test tube (eight pooled samples in total for our eight generations of evolution). We constructed the SMRTbell library using the DNA Template Prep Kit 2.0 (250 bp - 3 Kb, Pacific Biosciences). We ligated a blunt end adapter that contains a unique 16-bp barcode (Table S12) to amplicons of each sample, and used the ligation products as SMRTbell templates. We used the DNA/Polymerase P4 binding kit (Pacific Biosciences) to create ready-to-sequence SMRTbell-polymerase complexes, and pooled these complexes to create one amplicon pool. Subsequently, we sequenced this amplicon pool on eight SMRT cells (v3.0) using the Pacific Biosciences RS2 instrument (Pacific Biosciences). We used P4/C2 chemistry and the magnetic bead loading method, and recorded two movies of 3 hour duration for each cell.

Primary Data Analysis

We analyzed SMRT sequencing data using a previously described pipeline (*60*). Briefly, we used the protocol RT_ReadsOfInsert in SMRTAnalysis v2.3 package to demultiplex the PacBio RSII raw data and assemble consensus reads from sub-reads (*61*). We set the full-pass subread number to ≥4, the insert length to 650-1000 bp, and the predicted consensus accuracy to ≥0.9, in order to filter reads of YFP inserts. We set a minimum barcode score of 23 (resulting in 99.5% calling accuracy) to demultiplex the resulting consensus reads from the same generation of evolution into a single file. We mapped reads from each single file to the ancestral YFP sequence (*58*) using BLASR (*62*) by setting the mapped length to ≥850 bp and the mapping accuracy to ≥0.9. We only used mapped reads that span the entire YFP coding region and that had an average Phred quality above 20 for further analysis. We used SAMtools (*63*) for converting the resulting sam files to bam files, and de-multiplexed the filtered, mapped reads according to each read's barcode using custom Python scripts. We excluded those sequences from further analysis that lacked a start codon or stop codon. These primary processing steps yielded hundreds of sequences from each replicate population for further analyses (Table S4).

Identification of SNPs and mutation combinations

Because single-nucleotide indels account for more than 90% of sequencing errors during SMRT sequencing (*64*), and because most indels render YFP non-functional, we ignored indels and only analyzed point mutations. A mismatch of an YFP variant sequence to the YFP reference sequence was considered a true SNP (point mutation) only if its Phred quality score was above 20. We wrote Python scripts (Python 2.7.12) to identify point mutations and their combinations, and calculated their frequencies in each replicate population from each generation of evolution. Because we genotyped 613-1574 variants for each replicate population and in each phase I generation, our detection limit for SNPs or mutation combinations varied with the reciprocal of these values between 0.064% to 0.163%, depending on the population and generation.

Calculation of genetic diversity within populations

We used the pairwise alignment positional nucleotide counting (PAPNC) method (*65*) for calculating the per-site and average genetic diversity, as described previously (*60*). We used Python 2.7.12 to compute sequence identity and similarity from sequence alignments.

Calculation of the mean distances between populations

We determined the mean genetic distance between all pairs of sequences whose members are taken from two different populations at the evolutionary end point. To this end, we used MEGAX 10.0.5 (https://www.megasoftware.net/) to calculate the number of amino acids that differed between two sequences, and averaged these differences for all pairs of sequences in the two populations.

Creation of genotype networks

To create the networks shown in Fig. 3C, we determined the set of amino acid sequences (genotypes) that appeared in at least one generation of at least one replicate population. A genotype consisted of the amino acid residues at those fifteen positions in a fluorescent protein sequence (positions 2, 43, 47, 65, 66, 72, 102, 129, 141, 145, 164, 167, 168, 172 and 204) where four typical mutations, eleven alternative mutations, two expression-improving mutations (see Supplementary Text S1) and the mutation L43M arose. We represented each genotype as a vertex in the network and connected two vertices by an edge if their corresponding sequences differed by a single amino acid. The layout of the network was generated with a force-directed algorithm ("force atlas 2") in Gephi (https://gephi.org/).

Statistical analysis

Unless specified otherwise, we used a two-tailed t-test to conduct pairwise comparisons or multiple comparisons with Holm adjustment to control the family-wise error rate. To test the statistical significance of the fluorescence intensity change caused by every single mutation at different genetic backgrounds, we performed a two-sided t-test with Holm corrections. To determine whether each V^C replicate population had significantly higher green fluorescence increase than all V^0 populations at the evolutionary endpoint, we performed a two-way ANOVA with post hoc Tukey's test. We performed all statistical analysis using R version 3.4.4.

Supplementary Text

Supplementary Text S1 Expression-improving mutations were enriched during evolution for green fluorescence

During evolution for green fluorescence, several mutations located at the N-terminus of the protein achieved high or medium frequency in both V^0 and V^C populations (the first twenty residues, Fig. S3). For example, the mutations V2A and V2M (the latter change forms two consecutive start codons) reached a frequency of more than 10% in most replicate populations of V^0 and V^C. We engineered these two mutations into the ancestral YFP and measured their effects on both yellow fluorescence and green fluorescence. They can improve both yellow fluorescence and green fluorescence by a similar amount (Fig. S9A). Also, they can improve green fluorescence by almost the same amount in the genetic backgrounds G66S, Y204C and G66S+Y204C (Fig. S9A). Based on the above observations, we suspect that they affect fluorescent protein expression. Notably, these mutations did not attain significantly higher frequency in V^C populations than in V^0 populations (Fig. S9B). Thus, these mutations are not main contributors to faster phenotypic and genetic evolution in V^C populations. However, the modest increase in yellow fluorescence after the first generation of phase II could be caused by the accumulation of these expression-improving mutations (Fig. 2A).

Supplementary Text S2 Sign epistasis constrains the accessibility of mutational paths to high-fitness genotypes in V^0 populations

Although each of the eleven alternative mutations and the typical mutation F47L do not improve green fluorescence individually (Fig. S10), these mutations or combinations thereof can significantly enhance green fluorescence when they co-occur with the typical mutations G66S and Y204C (Fig. S10 & Fig. 3E). Such epistatic interactions can constrain the mutation paths to high-fitness genotypes in V^0 populations. To find out whether this is the case, we engineered selected mutations and their combinations described below into ancestral YFP and measured their fluorescence.

We first analyzed the accessibility of mutational paths leading to the typical genotype T1 (T+F65L), by measuring fluorescence of all eight genotypes leading to this genotype. Five of six mutational paths are accessible (Fig. 4A). In contrast, the same approach shows that only two and three of six potential paths to the typical genotype T2 (T+F47L) and T3 (T+L43M) are accessible, respectively (Fig. S5).

We next analyzed the alternative genotype A1 (Fig. 4A & Fig. S5), and noted that none of the four individual mutations that constitute it (F65S, K102R, N145S and V164A) can significantly improve green florescence in the genetic background of T or in the ancestral YFP. Thus, there is no accessible path to this genotype.

For the alternative genotype A2 (T+F72I+I172V+K167E; Fig. S6 & S10), F72I, I172V, or K167E do not significantly improve green florescence in the genetic backgrounds of G66S, Y204C or ancestral YFP. However, F72I or I172V can significantly improve green florescence in the T background, whereas K167E can significantly improve green florescence in the T+F72I+I172V background. These measurements lead us to conclude that only four out of one

hundred and twenty possible mutational paths to this alternative genotype are accessible from the wild-type.

For the alternative genotype A3 (T+F47L+I129T+K141R), we engineered mutations I129T and K141R individually into the background of T, and found that these mutations individually do not improve green fluorescence (Fig. S5 & S10). An analogous experiment shows that the joint mutation I129T+K141R slightly improves green florescence by 6.4% in background T2 (T+F47L), but significantly increases green florescence by 14.6% in the background of T2+V2M (Fig. S11). Thus, the joint mutation I129T+K141R is slightly beneficial in background T2, but more strongly beneficial in the presence of the additional mutation V2M, which can improve fluorescent protein expression (Fig. S9). These observations are also consistent with the observation that 95.61% of A3 genotypes also harbor the mutation V2M at the evolutionary endpoint (Fig. S11). To simplify the analysis of accessibility of mutation paths to A3, we do not consider the effect of V2M and treat the joint mutation I129T+K141R as a beneficial mutation in background T2. Although we did not engineer I129T and K141R individually into the background of T2, the following considerations allow us to assert that no more than 8.33% of paths to A3 are accessible. First, if neither I129T nor K141R separately can significantly improve green florescence in background T2, no path would be accessible to genotype A3 from T2 (0% accessible paths). Second, if only one of the two mutations I129T and K141R were to improve green florescence in background T2, then only two of twenty-four possible paths to T2+I129T or T2+K141R would be accessible (8.33% accessible paths), because either I129T or K141R would not contribute to green fluorescence of A3. Lastly, if I129T or K141R separately can significantly improve green florescence in background T2, it follows from our measurements that no more than four out of one hundred and twenty paths (3.33%) to A3 can be accessible.

For the alternative genotype A4 (T+F72C+I168V), we engineered and measured thirteen genotypes leading to this genotype. These measurements show that I168V cannot increase green fluorescence in the background T+F72C (Fig. S12). However, it compensates for the deleterious effects of the mutations G66S, T and T+F72C on yellow fluorescence (Fig. S12). It is thus likely that I168V already occurred in phase I to compensate the deleterious effects of mutations G66S, Y204C, F72C (or combinations thereof). For this reason, it would be inappropriate to include I168V when constructing potential mutational paths to the alternative genotype A4 based on its effects on green fluorescence. Of the remaining six paths to T+F72C, two are accessible (Fig. S5).

Supplementary Text S3 Evolutionary trajectories of high-fitness genotypes in V^0 and V^C populations

Our experimental data on (i) genotype frequency (Table S6, S13-24) and (ii) the rate at which mutation pressure alone causes specific mutations to increase in frequency during evolution enabled us to infer the evolutionary trajectories of high-fitness genotypes in each replicate population. Consider this example for typical genotype T1 in replicate population V_1^0. Genotype T1 first appeared during the third generation (Table S13). In the second generation, we detected the three constituent mutations of this genotype individually at frequencies of 0.16% (F65L), 2.17% (Y204C), and 62.95% (G66S), as well as two out of the three possible double mutants at frequencies 0.16% (G66S+F65L) and 11.32% (G66S+Y204C, which is the T genotype). From our evolution experiments in the absence of selection (Fig. S4), we can quantify

11

how fast individual mutations and genotypes are expected to arise through mutation pressure alone (Table S7). From this experimental data, we can compute the probability P of each possible mutational path (or genotype) leading to T1 as shown below, where f_X indicates the observed frequency of mutation X in the second generation, and μ_X indicates the rate at which a mutation X arises in the absence of selection, i.e., through mutation alone.

$$P_{F65L \to T1} = f_{F65L} \times \mu_{Y204C} \times \mu_{G66S} = 0.16\% \times 0.258\% \times 0.0868\% = 3.6 \times 10^{-9}$$

$$P_{Y204C \to T1} = f_{Y204C} \times \mu_{F65L} \times \mu_{G66S} = 2.17\% \times 0.22\% \times 0.0868\% = 4.1 \times 10^{-8}$$

$$P_{G66S \to T1} = f_{G66S} \times \mu_{Y204C} \times \mu_{F65L} = 62.95\% \times 0.258\% \times 0.22\% = 3.6 \times 10^{-6}$$

$$P_{G66S+F65L \to T1} = f_{G66S+F65L} \times \mu_{Y204C} = 0.16\% \times 0.258\% = 4.1 \times 10^{-6}$$

$$P_{T \to T1} = f_T \times \mu_{F65L} = 11.32\% \times 0.22\% = 2.5 \times 10^{-4}$$

Note that these calculations allow simultaneous appearance of two but not of three mutations because the simultaneous occurrence of three mutations is very unlikely ($<10^{-7}$, estimated by raising the mean incidence rates of mutations in Table S7 to the third power). The above computation shows that the most likely mutational path to reach T1 starts from T because its probability ($P_{T \to T1}$) is much higher than all others. Next, we used the same approach to estimate the possible mutational paths (or genotypes) leading to T itself, which showed that T would most likely first appear at a frequency of 0.45% in the first generation, which involves the ancestor acquiring two mutations at once. Taken together, these analyses show that the most likely evolutionary trajectory leading to T1 in V_1^0 reaches it from the wild-type ancestor via genotype T in two steps (WT→T→T1; Fig. S7).

Using the above strategy, we inferred the most likely paths to all typical and alternative genotypes in each replicate population. Fig. S7 summarizes these analyses and shows that all the typical genotypes, which dominate V^0 populations, are most likely reached by selectively accessible trajectories. In contrast, all except one of the high-fitness genotypes in V^C populations (T1 in V_2^C), are reached by trajectories that would be inaccessible to V^0 populations. Thus, the cryptic variation that accumulated during phase I evolution is important to reach otherwise unattainable high fitness genotypes.
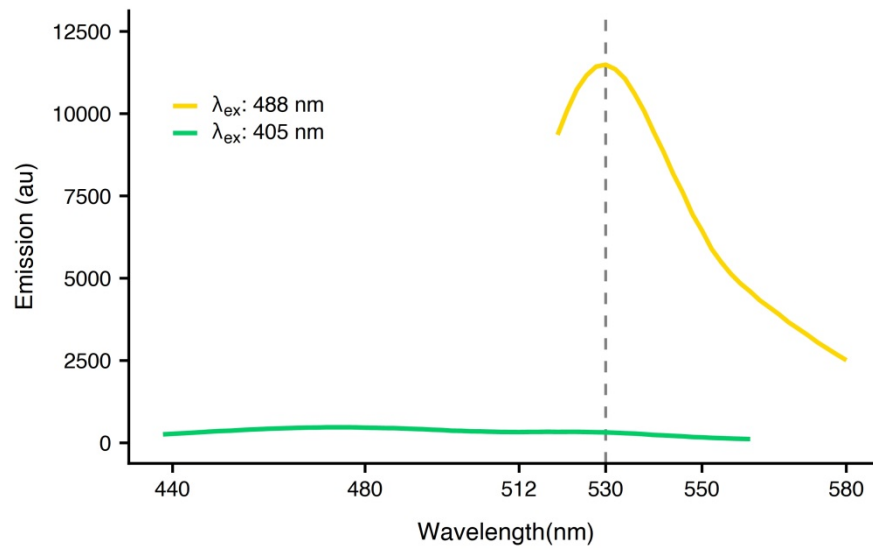
Fig. S1.
The emission spectrum of ancestral YFP at the native excitation wavelength (yellow line, λex=488nm) and at the phase II excitation wavelength (green line, λex=405nm). The vertical axis indicates fluorescence intensity (arbitrary units) at the emission wavelength (horizontal axis).
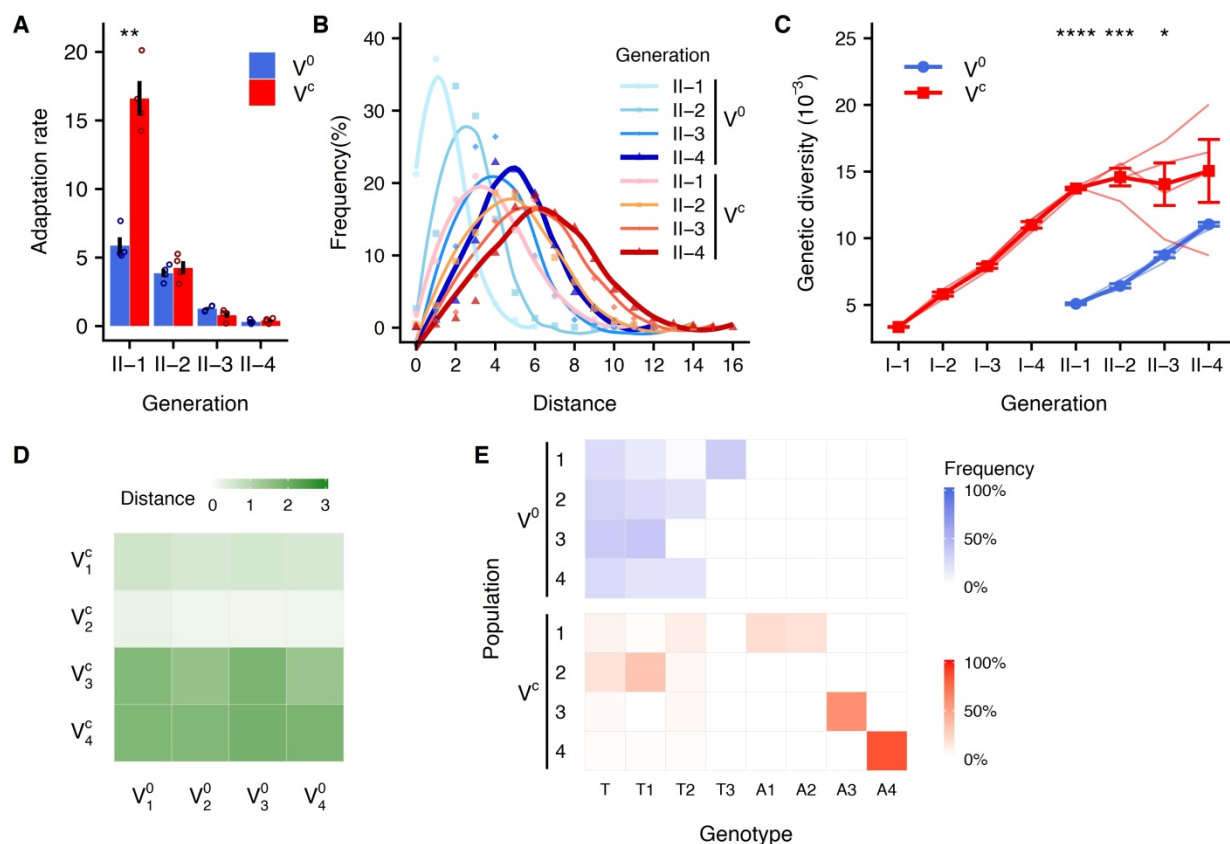
Fig. S2.

(A) Rate of evolutionary adaptation (shown as the increase of green fluorescence intensity in each generation) of populations V^0 and V^C during phase II evolution. Each circle represents a replicate population. We performed two-sided t-tests to determine whether the adaptation rates between V^0 and V^C are significantly different in each generation. **indicates *P*<0.01. **(B) Distribution of the number of amino acid changing mutations per individual molecule in V^0 and V^C populations during the phase II evolution.** Each circle represents the mean of four replicate populations. Lines were generated by LOESS curve fitting. Thin lines represent the frequency distribution from generation 1 to 3, and thick lines indicate the frequency distribution at the end of phase II (generation 4). To count the number of amino acid changing mutations per sequence, we aligned each individual amino acid sequence inferred from DNA sequences of each population to the ancestral YFP reference sequence. **(C) Average pairwise sequence distance (see Materials and Methods) of evolving V^0 (thick blue lines) and V^C (thick red lines) populations.** Error bars represent one SEM of four replicate populations. Each thin blue and red line represents a replicate population of V^0 and V^C, respectively. We performed one-sided t-tests to determine if the genetic diversity of the V^C populations was higher than the genetic diversity of the V^0 populations (t=-70.43, df=5.26, *P*=2.46×10^{-9} for generation II-1, t=-12.04, df=3.37, *P*=0.00035 for generation II-2, t=-3.29, df=3.11 and *P*=0.022 for generation II-3 and t=-1.69, df=3.02 and *P*=0.094 for generation II-4). *, *** and **** indicate *P*<0.05, 0.001 and 0.0001. **(D) Average number of amino-acid differences (at the evolutionary endpoint) between all protein sequences in the labeled populations (see Materials and Methods). (E)**

The frequency of genotypes in each replicate V^0 and V^C population at the endpoint of evolution. Estimates of the frequency of A1 and A2 also include genotypes derived from A1 or A2 but with one fewer mutation (see Table S5).
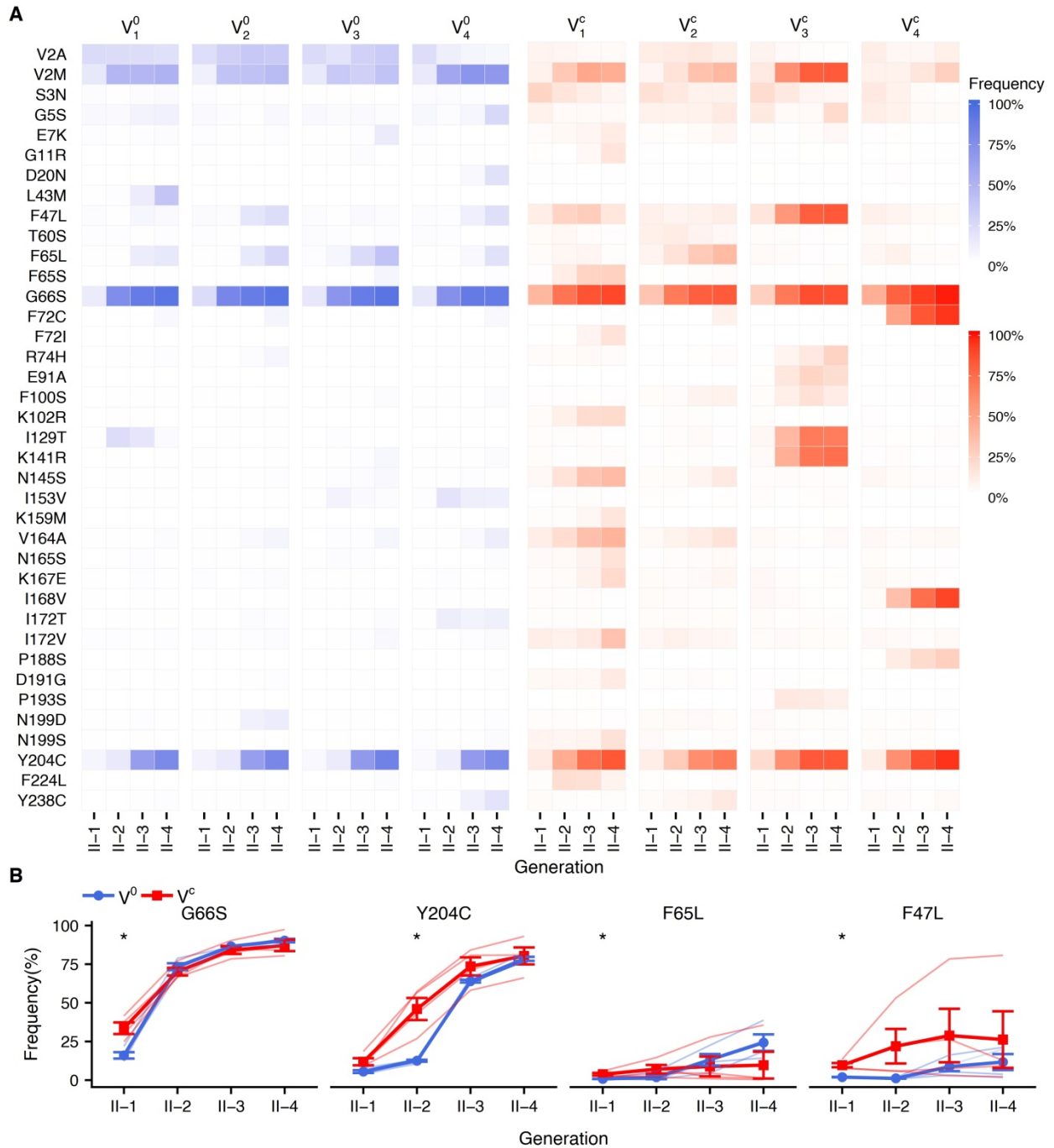
Fig. S3.

Dynamics of amino acid changing single nucleotide polymorphism (SNP) frequencies in evolving populations V^0 and V^C during the evolution of green fluorescence (phase II).
(A) The horizontal axis indicates time (generations of directed evolution), and the vertical axis indicates individual mutations. Colored squares indicate mutation frequencies (see color legend). Only those SNPs which achieved a frequency exceeding 10% in at least one replicate population of V^0 and V^C are shown. **(B)** The horizontal axis indicates time (generations of directed evolution), and the vertical axis indicates the frequencies of individual mutations. Thick blue and

16

red lines indicate averages for populations V^0 and V^C, respectively, and error bars represent one SEM based on four replicate populations. Each thin blue or red line indicates data from a single replicate population V^0 or V^C, respectively. For each generation, we performed two-sided t-tests to determine whether the frequencies of individual mutations differed between V^0 and V^C. '*' indicates $P<0.05$.
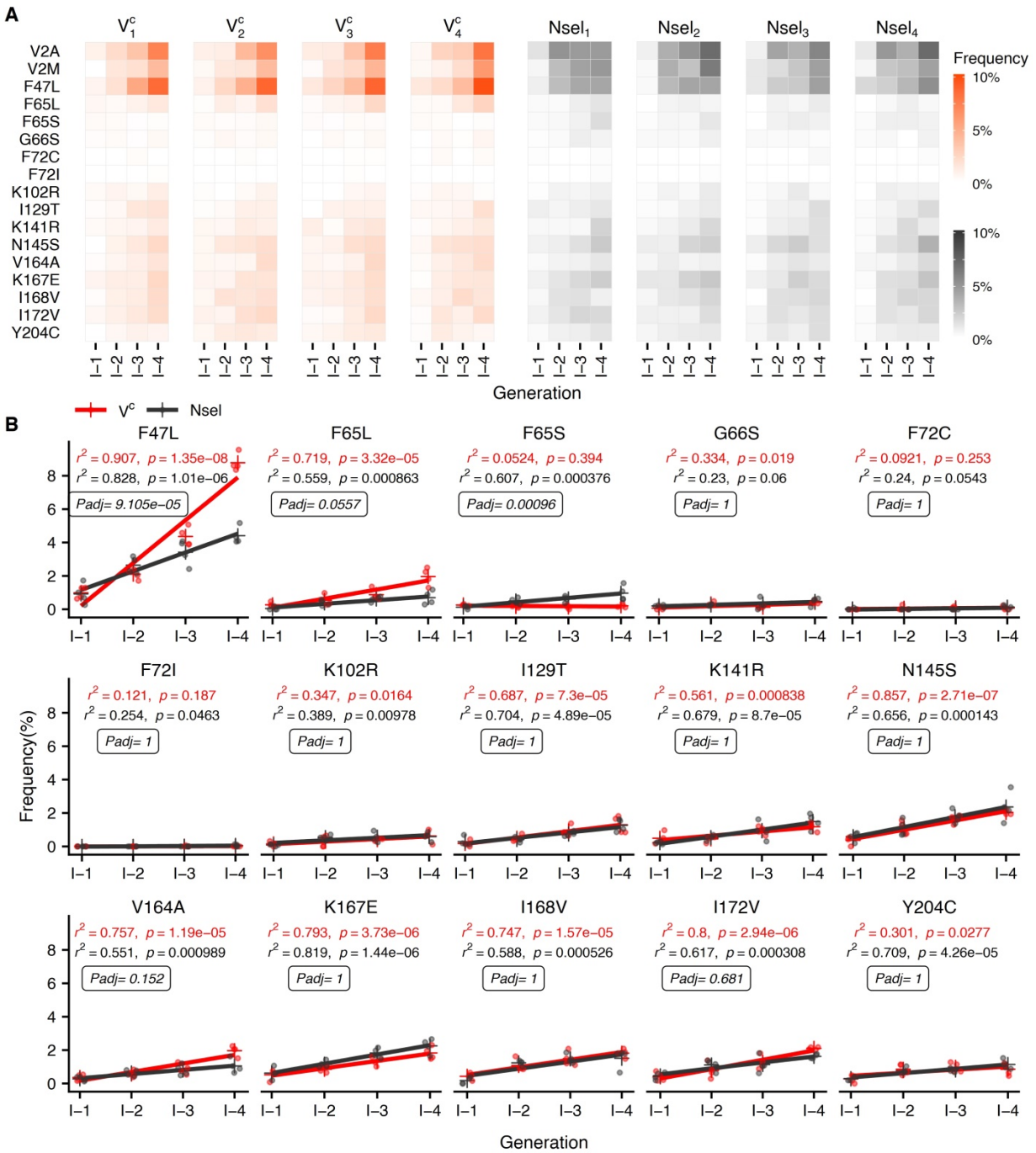
Fig. S4.

Most mutations that reach appreciable frequency in phase I do not accumulate at rates greater than expected from the mutation rate alone. The populations $NseI_{1-4}$ (No selection) underwent the same directed evolution protocol as populations V^C, except that we collected all gene variants encoding fluorescent protein in each generation, regardless of the fluorescence intensity of the encoded proteins. In other words, populations $NseI_{1-4}$ did not experience selection for yellow fluorescence, and thus serve as controls to identify candidate mutations influenced by

selection in V^C populations, because they increase more rapidly in frequency in V^C populations than in Nsel populations. In panel A**,** the horizontal axes indicate time (generations of directed evolution), and the vertical axis lists individual mutations. The color scale indicates mutation frequency. In panel B, the horizontal axes indicate time (in generations of directed evolution), and the vertical axes indicate mutation frequency. Each grey or blue circle represents frequency data from a replicate V^C or Nsel population, respectively. Error bars represent one SEM over four replicate populations. We used F-statistics to determine whether increases in mutant frequencies are significantly different from zero, as indicated by the p-values shown in grey for Nsel populations and in red for V^C populations. Values of r^2 indicate the proportion of variance explained by each regression. Importantly, we also conducted two-way Ancovas with Holm adjustment to compare the two linear regression models for each mutation, which are generated based on the frequency data of each mutation from V^C or Nsel populations in phase I, respectively. The resulting P values are highlighted in boxes (*Padj*). This analysis showed that only the F47L mutation frequency increased faster than expected from the mutation rate alone during stabilizing selection for yellow fluorescence in phase I.
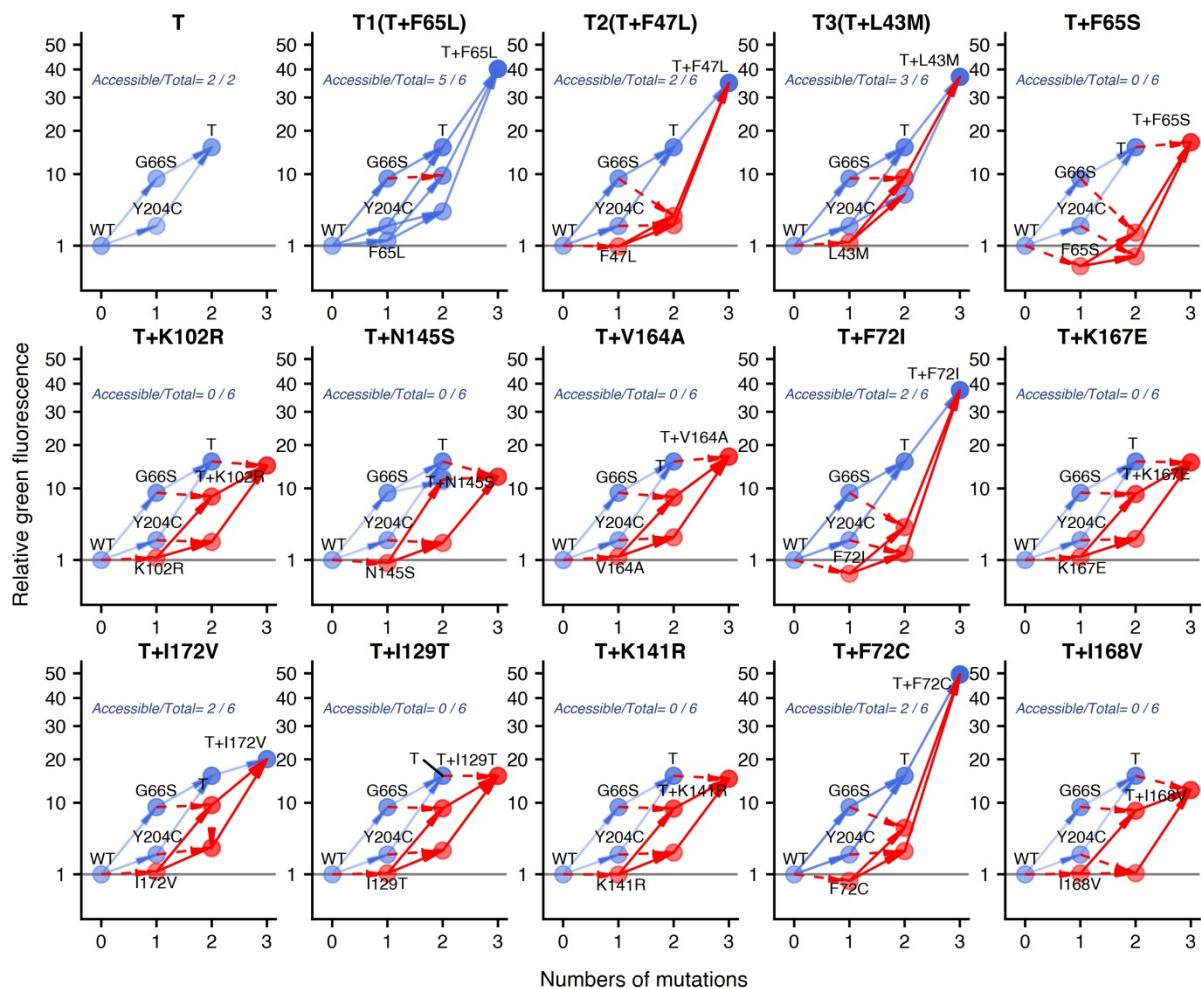
Fig. S5.

Accessible and inaccessible mutational paths to typical and alternative high fluorescence genotypes. We distinguish three kinds of steps in a mutational path, an accessible mutational step (blue arrows) that increases green fluorescence significantly, an inaccessible step (dashed red arrows) that does not increase green fluorescence significantly, and a conditionally accessible step (solid red arrows) that increases green fluorescence significantly in the genetic background where it occurs, but where the ancestral YFP must first traverse one or more inaccessible steps to create this genetic background. Blue circles indicate genotypes created by an accessible step, whereas red circles indicate genotypes created by an inaccessible or a conditionally accessible step. A path is inaccessible if it contains at least one inaccessible step. To determine accessibility, we considered a difference in green fluorescence between genotypes significant only if $P<0.05$, based on a two-sided t-test with Holm adjustment. For this test, we performed three biological replicate fluorescence measurements, and analyzed 10^4 cells for each biological replicate. The numbers of accessible paths and total paths are shown in each panel.
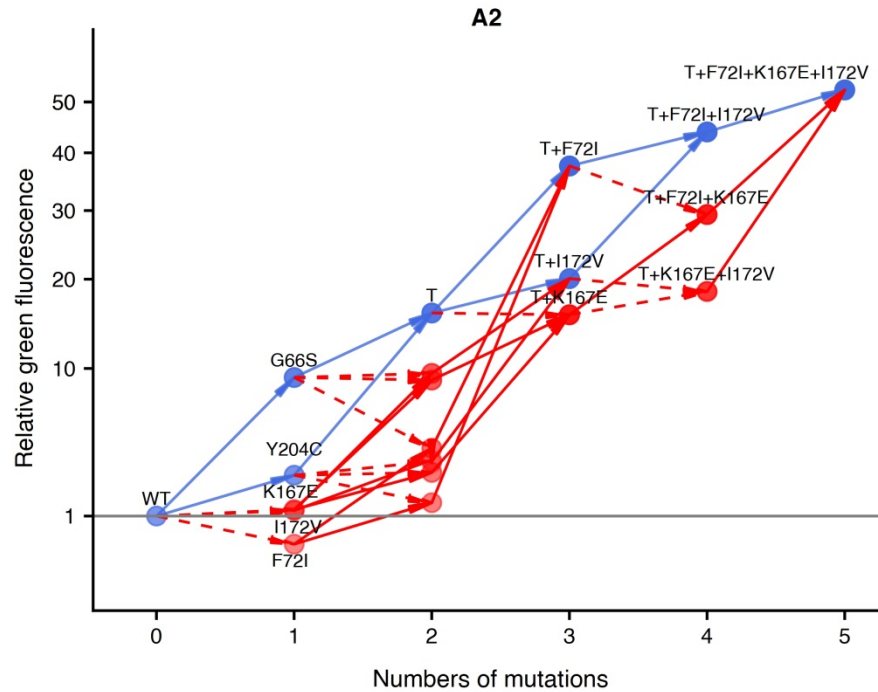
Fig. S6.

Accessibility of mutational paths to the alternative genotype A2 (see also Supplementary Text S2). Blue lines indicate an accessible mutational step, which increases green fluorescence significantly, and dashed red lines indicate an inaccessible step, which does not significantly increase green fluorescence. 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. A path is inaccessible if it contains at least one inaccessible step. We considered a difference in green fluorescence between genotypes as significant if $P<0.05$ (Two-sided t-test with Holm adjustment).
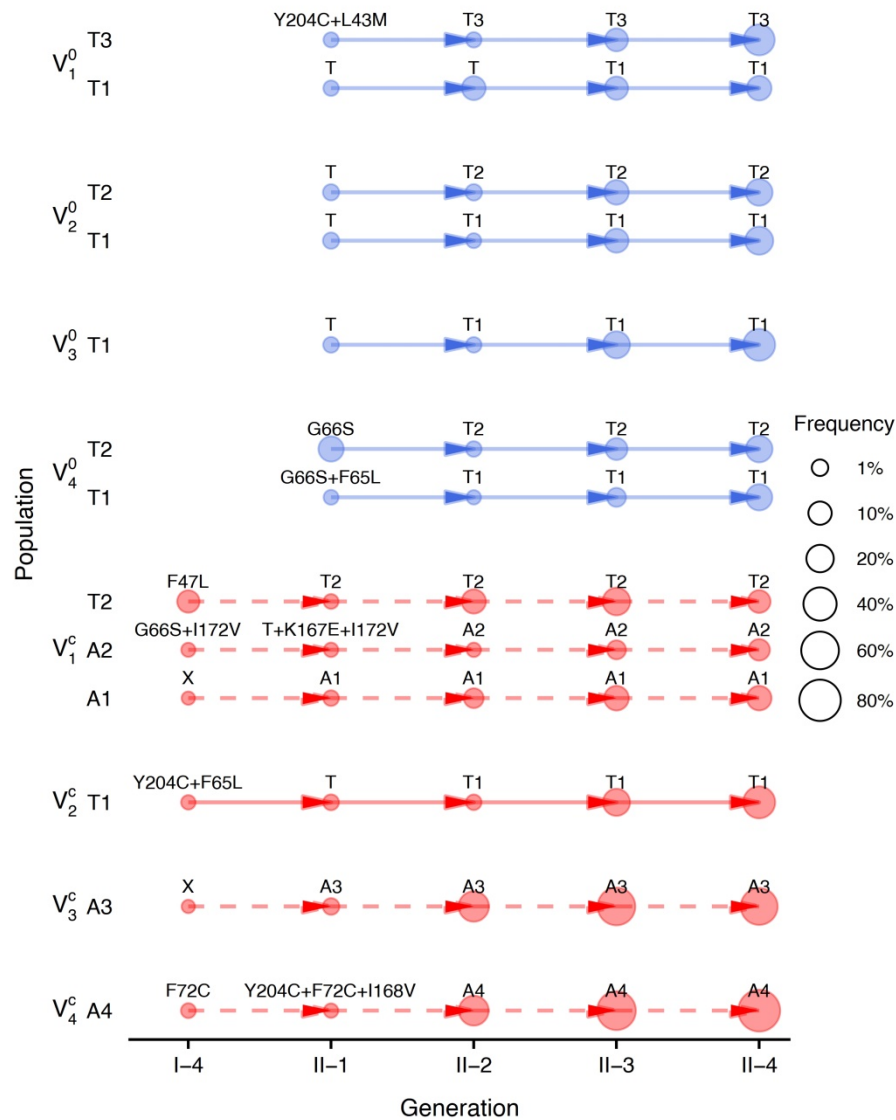
Fig. S7.

Inferred evolutionary trajectories of typical and alternative high fluorescence genotypes in V^0 or V^C populations. We inferred the mutational paths to typical and alternative high fluorescence genotypes based on frequency data of genotypes in each population and on the incidence of mutations during evolution (see Supplementary Text S3). The horizontal axis indicates time (generation). The vertical axes indicate individual replicate populations (V^0 or V^C) and the trajectories of high fitness genotypes that attain high frequency in these populations. Circle area reflects the frequency of the genotype that labels the circle. 'X' indicates a genotype that is likely to exist but that cannot be detected, and it was given the minimum circle area, because its frequency is below the detection limit (See Materials and Methods). Solid and dashed arrows indicate selectively accessible and inaccessible paths, respectively. A path is inaccessible if it contains at least one inaccessible step (See Fig. S5 and Supplementary Text S2 for details). The figure shows that all three typical genotypes attained by V^0 populations evolve by following

22

selectively accessible trajectories, whereas all alternative high-fitness genotypes (except one typical genotype T1 in V_2^C) evolved in V^C populations by following trajectories that would be selectively inaccessible in V^0 populations (see Supplementary Text S2 for details).
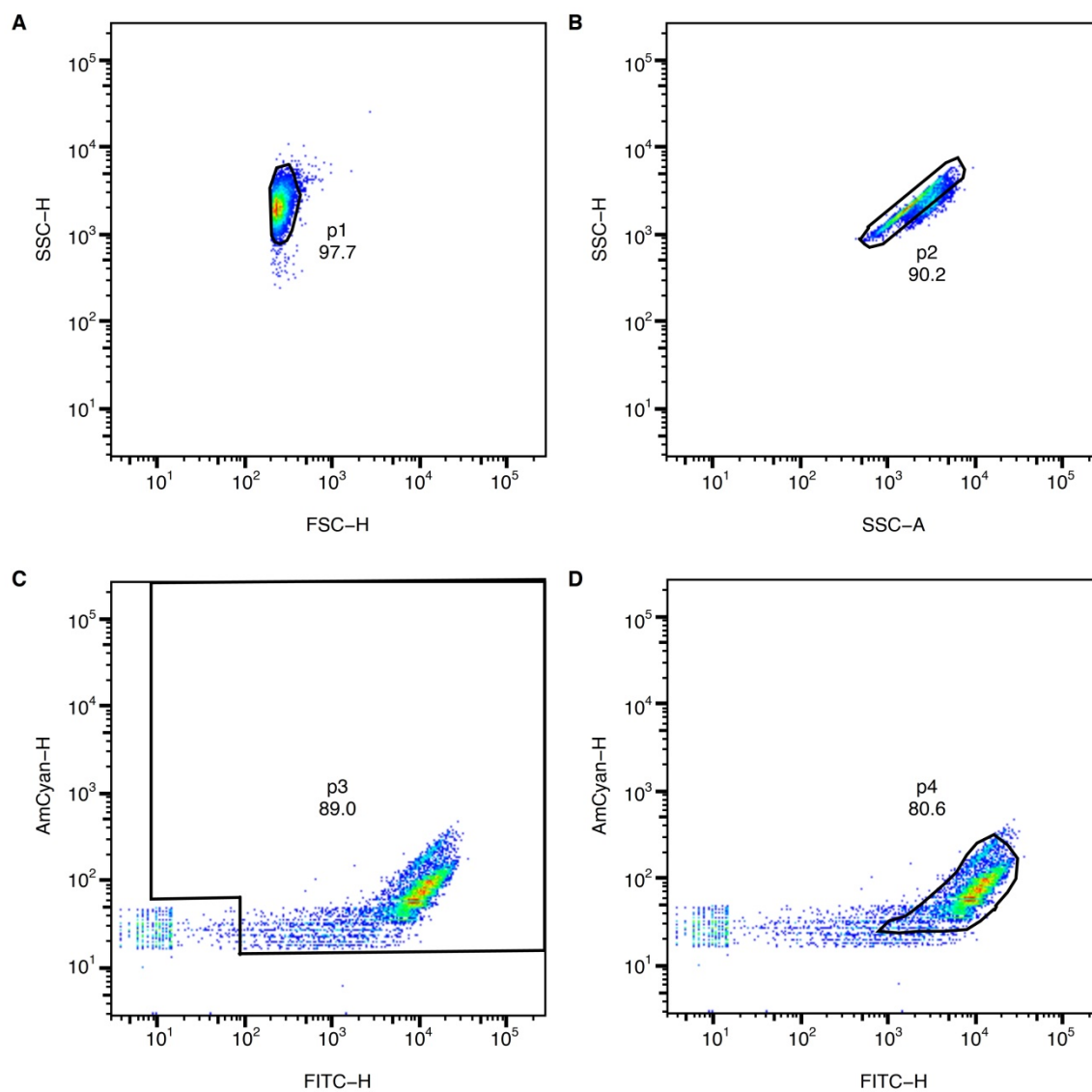
Fig. S8.
Representative gates for flow cytometry data analysis. We used forward scatter height (FSC-H) versus side scatter height (SSC-H) density plots to select a homogenous cell population (p1 in panel A), and used side scatter area (SSC-A) versus side scatter height (SSC-H) density plots to exclude doublets (p2 in panel B). For each evolving replicate population, we used FITC-Height (FITC-H) versus AmCyan-Height (AmCyan-H) density plots to further exclude dust, cellular debris and micro-particles, which showed fluorescence characteristic of the negative control population (ancestral YFP without adding arabinose for YFP induction; p3 in panel C). For each engineered variant, we used FITC-Height versus AmCyan-Height density plots to select the dominant cell population (p4 in panel D). We used the resulting filtered data (p3 or p4) for calculating the mean fluorescence intensity of each biological replicate.

24
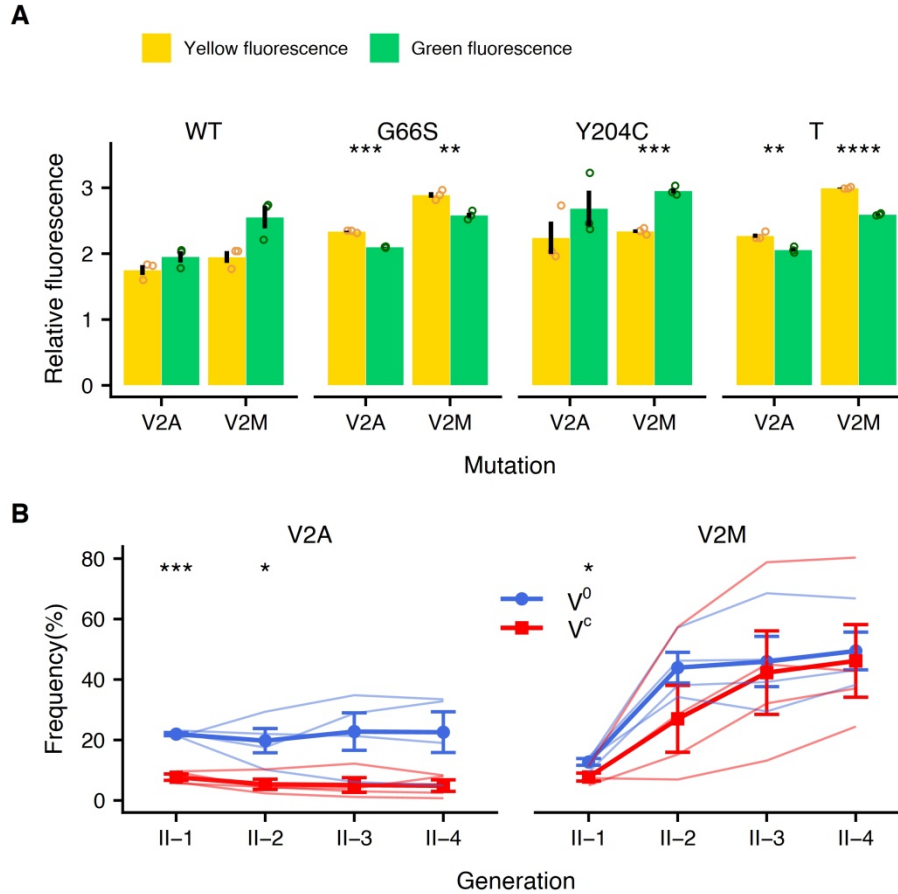
Fig. S9.

(A) Effects of mutations V2A and V2M on yellow fluorescence and green fluorescence.
The horizontal axis indicates genotypes V2A and V2M, and the vertical axis indicates yellow and green fluorescence intensity, expressed as a fold-change relative to ancestral YFP at different genetic backgrounds. Error bars represent one standard deviation of three biological replicate fluorescence measurements. We analyzed 10^4 cells for each such replicate. We performed two-sided t-tests (*P*-values shown above bars) to determine whether green and yellow fluorescence intensities differed significantly. **(B) Frequency of V2A and V2M in evolving population V^0 and V^C during the evolution of green fluorescence.** Thick blue and red lines indicate means for populations V^0 and V^C, respectively, and error bars represent one SEM of four replicate populations. Each thin blue and red line indicates data from a replicate population of V^0 and V^C, respectively. We performed two-sided t-test to determine the statistical significance of frequency differences in V2A or V2M between V^0 and V^C in each generation of evolution. *, ** or *** indicate that *P*<0.05, 0.01 or 0.001 respectively.
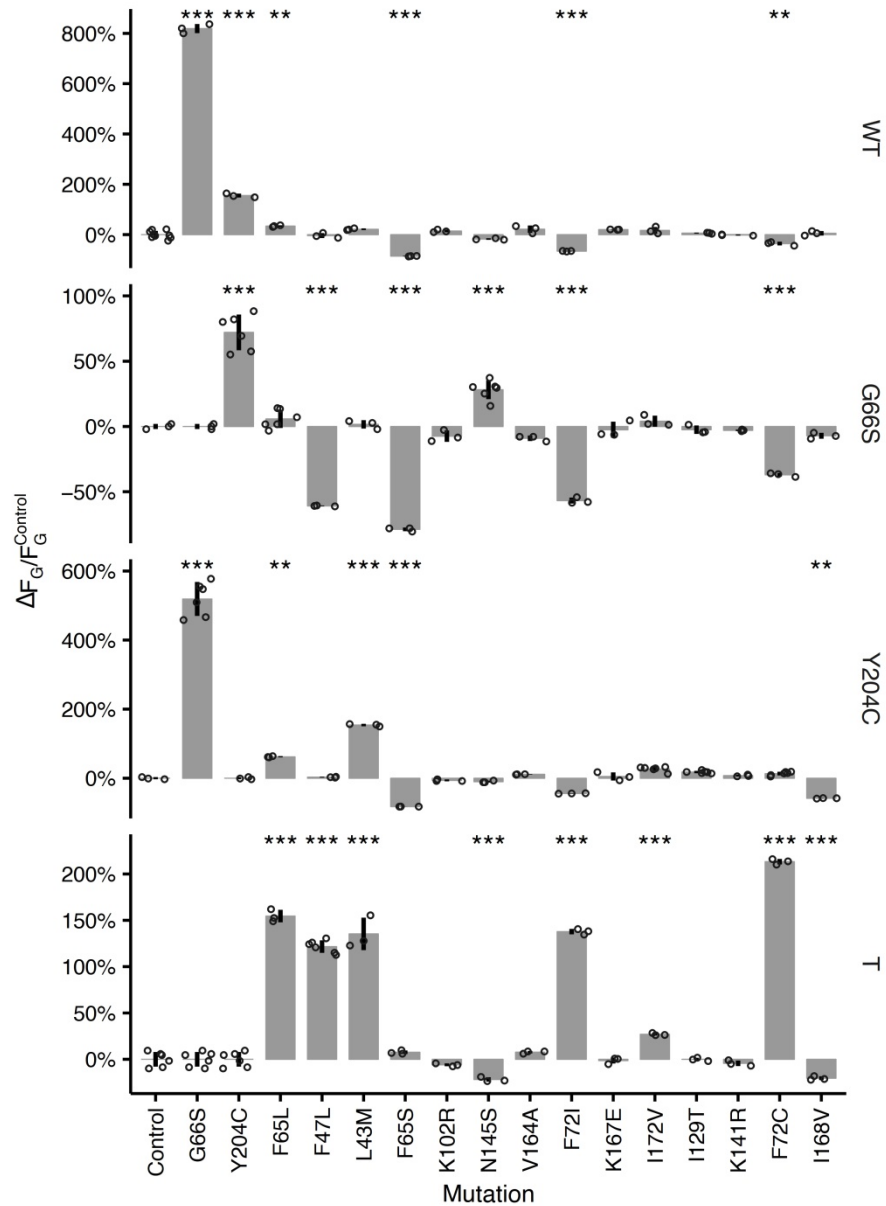
Fig. S10.

Effects of single mutations on green fluorescence in different genetic backgrounds.
The left vertical axes indicate the fluorescence intensity change of each mutation (horizontal
axes) in different genetic backgrounds (right vertical axes) relative to proteins with that genetic
background, i.e., lacking the mutation. Error bars represent one standard deviation and are based
on three or six biological replicate measurements. We analyzed 10^4 cells for each biological
replicate. Each dot represents a biological replicate measurement. We performed two-sided t-test
with Holm adjustment to test the null hypothesis that the fluorescence intensity of a given mutant
is identical to that of the background. *, ** and *** indicate $P<0.05$, 0.01 and 0.001.
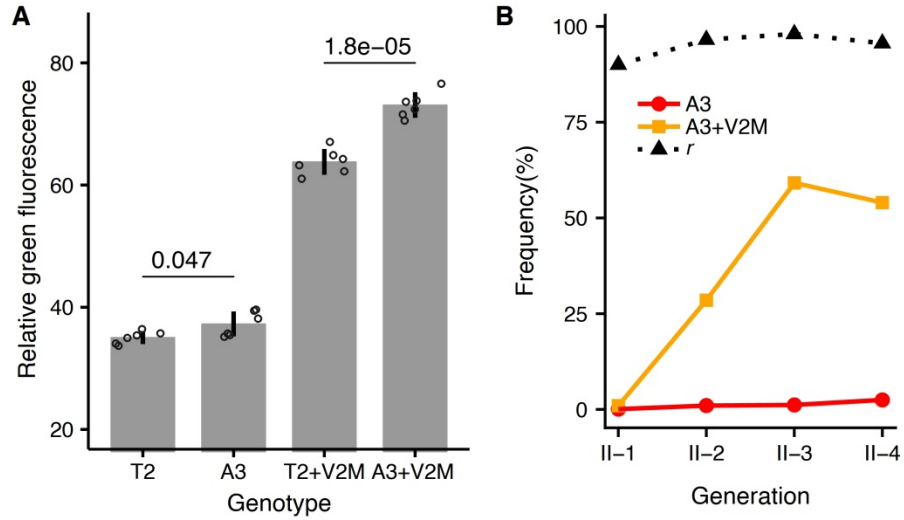
Fig. S11.

The beneficial effect of the joint mutation I129T+K141R becomes pronounced in background T2+V2M. Panel A shows the relative fluorescence intensity of each indicated genotype (horizontal axis) relative to the ancestral YFP. Error bars represent one standard deviation and are based on six biological replicate measurements. We analyzed 10^4 cells for each biological replicate. Each dot represents a biological replicate measurement. We performed two-sided t-tests (*P*-values shown above bars) to determine whether the green fluorescence intensities between T2 and A3 or between T2+V2M and A3+V2M differed significantly. Panel B shows the frequency of two genotypes in the population V_3^C during Phase II, as well as the ratio r of the frequency of A3+V2M to the frequency of A3 (dashed line).
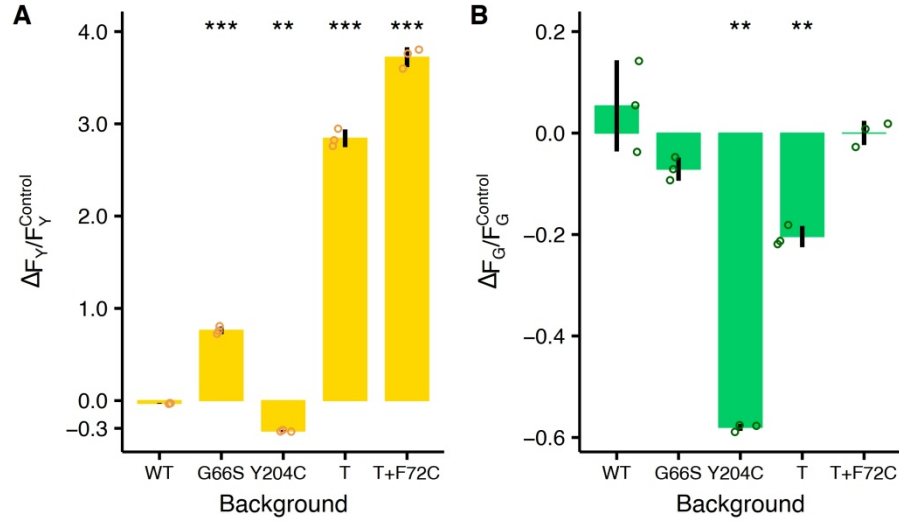
Fig. S12.

Effects of the mutation I168V on yellow (A) and green (B) fluorescence in different genetic backgrounds. The horizontal axis indicates genotypes, and the vertical axis indicates the fluorescence intensity change caused by the acquisition of the mutation I168V by the corresponding genotypes. Error bars represent one standard deviation based on three biological replicate fluorescence measurements, and on 10^4 analyzed cells for each such replicate. Each green or yellow dot represents a biological replicate measurement. We performed two-sided t-test with Holm adjustment to test the null hypothesis that the fluorescence intensity after acquiring the mutation I168V did not significantly change. *P* values are shown above the bars. *, ** and *** indicate *P*<0.05, 0.01 and 0.001.

Table S1

Mean number of SNPs per YFP molecule for evolving YFP populations during each generation of mutagenesis as determined by SMRT sequencing.

Population	SNPs per read in each generation of mutagenesis							
	I-1	I-2	I-3	I-4	II-1	II-2	II-3	II-4
V_1^0					2	3.19	4.8	5.87
V_2^0					1.98	3.18	4.85	6.05
V_3^0					1.96	3.18	4.61	6.2
V_4^0					1.88	3.27	4.71	6.05
V_1^C	1.19	2.13	2.97	4.19	5.28	6.22	7.52	8.34
V_2^C	1.23	2.24	2.98	4.19	5.22	6.3	7.57	8.87
V_3^C	1.22	1.96	2.73	3.84	5.08	8.15	10.42	11.06
V_4^C	1.19	2.09	2.81	3.92	5.33	6.28	7.31	8.04
$Nsel_1$	1.21	3.32	5	6.04				
$Nsel_2$	1.31	3.22	4.54	5.65				
$Nsel_3$	1.23	3.21	4.62	6.04				
$Nsel_4$	1.15	3.3	4.9	6.18				
Control 1[a]	1.04							
Control 2[a]	1.05							
YFP(ancestor)[b]	0.03							
YFP(ancestor)[b]	0.04							

[a] After constructing the mutation library, we sequenced the library before selection to estimate the mutation rate.
[b] We used ancestral YFP as a control and sequenced its encoding gene (see Materials and Methods).
We subjected populations $Nsel_{1-4}$ to the same procedure as populations V^C but without selection, i.e., we collected all generated variants in each generation, regardless of their fluorescence intensity.

Table S2

Mean number of amino-acid changes per YFP molecule for evolving YFP populations during each generation of mutagenesis as determined by SMRT sequencing.

Population	Amino-acid changing mutation per read for each generation of mutagenesis							
	I-1	I-2	I-3	I-4	II-1	II-2	II-3	II-4
V_1^0					1.42	2.70	4.05	4.74
V_2^0					1.45	2.59	4.10	5.09
V_3^0					1.40	2.48	3.80	5.09
V_4^0					1.37	2.75	4.15	5.46
V_1^C	0.95	1.59	2.19	3.03	3.86	5.07	6.35	6.90
V_2^C	0.98	1.71	2.13	2.98	3.70	4.44	5.63	6.78
V_3^C	0.96	1.48	2.00	2.80	3.74	5.66	6.91	7.45
V_4^C	0.96	1.62	2.10	2.90	3.64	4.39	4.98	5.49
$Nsel_1$	0.98	2.64	4.02	4.89				
$Nsel_2$	1.02	2.55	3.66	4.57				
$Nsel_3$	0.97	2.58	3.72	4.90				
$Nsel_4$	0.92	2.64	3.99	4.94				
Control 1[a]	0.83							
Control 2[a]	0.84							
YFP(ancestor)[b]	0.028							
YFP(ancestor)[b]	0.025							

[a] After constructing the mutation library, we sequenced the library before selection to estimate the mutation rate.
[b] We used ancestral YFP as a control and sequenced its encoding gene (see Materials and Methods).
We subjected populations $Nsel_{1-4}$ to the same procedure as populations V^C but without selection, i.e., we collected all generated variants in each generation, regardless of their fluorescence intensity.

TableS 3

Statistical significance of green fluorescence fold change between each replicate population of V^0 and V^C at the evolutionary endpoint [a]

Population	diff [b]	lwr [c]	upr [d]	P value (padj) [e]
V_3^C-V_4^C	-40.729	-58.328	-23.13	1.2E-05
V_2^C-V_4^C	-34.136	-51.736	-16.537	0.00011
V_1^C-V_4^C	-28.447	-46.046	-10.848	0.00082
V_4^0-V_4^C	**-72.574**	**-90.173**	**-54.975**	**0**
V_3^0-V_4^C	**-53.061**	**-70.66**	**-35.462**	**3E-07**
V_2^0-V_4^C	**-68.408**	**-86.007**	**-50.809**	**0**
V_1^0-V_4^C	**-63.887**	**-81.486**	**-46.288**	**0**
V_2^C-V_3^C	6.59228	-11.007	24.1914	0.88732
V_1^C-V_3^C	12.2821	-5.317	29.8812	0.29703
V_4^0-V_3^C	**-31.845**	**-49.444**	**-14.246**	**0.00024**
V_3^0-V_3^C	**-12.332**	**-29.931**	**5.2668**	**0.29279**
V_2^0-V_3^C	**-27.679**	**-45.278**	**-10.08**	**0.00109**
V_1^0-V_3^C	**-23.158**	**-40.758**	**-5.5594**	**0.00606**
V_1^C-V_2^C	5.68983	-11.909	23.2889	0.94297
V_4^0-V_2^C	**-38.437**	**-56.036**	**-20.838**	**2.5E-05**
V_3^0-V_2^C	**-18.925**	**-36.524**	**-1.3255**	**0.03051**
V_2^0-V_2^C	**-34.271**	**-51.87**	**-16.672**	**0.0001**
V_1^0-V_2^C	**-29.751**	**-47.35**	**-12.152**	**0.00051**
V_4^0-V_1^C	**-44.127**	**-61.726**	**-26.528**	**4.3E-06**
V_3^0-V_1^C	**-24.614**	**-42.213**	**-7.0153**	**0.00347**
V_2^0-V_1^C	**-39.961**	**-57.56**	**-22.362**	**1.5E-05**
V_1^0-V_1^C	**-35.441**	**-53.04**	**-17.842**	**6.8E-05**
V_3^0-V_4^0	19.5126	1.91356	37.1117	0.02443
V_2^0-V_4^0	4.1661	-13.433	21.7652	0.98925
V_1^0-V_4^0	8.68646	-8.9126	26.2855	0.6824
V_2^0-V_3^0	-15.347	-32.946	2.25254	0.11192
V_1^0-V_3^0	-10.826	-28.425	6.7729	0.43768
V_1^0-V_2^0	4.52036	-13.079	22.1194	0.98293

[a] we performed two-way ANOVAs with post hoc Tukey's test ($F_{7,16}$=46.5, P=1.99×10^{-9}). The comparisons between V^C and V^0 replicate populations are highlighted in bold.
[b] diff refers to the difference in the observed means.
[c] lwr refers to the lower end point of the confidence interval.
[d] upr refers to the upper end point of the confidence interval.
[e] padj refers to the p-value after adjustment for the multiple comparisons.

Table S4

Number of reads sequenced by SMRT sequencing for evolving YFP populations during each generation of evolution

Population	Reads in each generation							
	I-1	I-2	I-3	I-4	II-1	II-2	II-3	II-4
V_1^0					894	645	586	631
V_2^0					922	709	592	798
V_3^0					1002	593	607	460
V_4^0					835	745	569	220
V_1^C	630	1226	1084	1389	898	826	651	753
V_2^C	808	613	939	827	833	643	542	553
V_3^C	928	1045	1571	1382	925	589	764	524
V_4^C	1172	1081	1250	879	861	778	599	585
$Nsel_1$	1011	947	659	764				
$Nsel_2$	749	1290	810	929				
$Nsel_3$	795	872	744	787				
$Nsel_4$	521	700	862	677				
Control 1[a]	1764							
Control 2[a]	1188							
YFP(ancestor)[b]	910							
YFP(ancestor)[b]	629							

[a] after constructing the mutation library, we sequenced the library before selection to estimate the mutation rate.
[b] We used ancestral YFP as a control and sequenced its encoding gene (see Materials and Methods).
We subjected populations $Nsel_{1-4}$ to the same procedure as populations V^C but without selection, i.e., we collected all generated variants in each generation, regardless of their fluorescence intensity.

Table S5

Green fluorescence intensity of four alternative genotypes and their frequency at the endpoint of evolution

Genotype	Mutation	Frequency (%)	Population	Green fluorescence intensity [a]
A1	T+F65S+K102R+N145S+V164A	12.9	V_1^C	47.4±4.8
	T+F65S+N145S+V164A	4.9		52.8±6.3
A2	**T+F72I+K167E+I172V**	6.8	V_1^C	52.5±0.4
	T+F72I+I172V	1.5		44.0±1.2
	T+F72I+K167E	2		29.4±0.5
	T+K167E+I172V	5.4		18.4±1.3
A3	T+F47L+I129T+K141R	56.5	V_3^C	33.5±1.2
A4	T+F72C+I168V	82.7	V_4^C	49.6±1.2

[a] expressed as a fold-change relative to ancestral YFP

Table S6

Frequency of genotypes leading to A4 in V_4^C during evolution

Mutation				Generation				
G66S	Y204C	F72C	I168V	I-4	II-1	II-2	II-3	II-4
-	-	-	-	97.27	50.52	14.52	5.01	1.03
-	-	-	+	1.48	1.05	0.39	0.00	0.17
-	-	+	-	0.23	0.12	1.93	1.17	0.17
-	-	+	+	0.00	0.00	0.13	0.17	0.00
-	+	-	-	0.46	6.50	2.96	1.67	0.00
-	+	-	+	0.00	0.00	0.64	0.50	0.34
-	+	+	-	0.00	0.12	2.19	0.67	0.34
-	+	+	+	0.00	0.12	0.00	0.50	0.51
+	-	-	-	0.57	38.10	18.51	3.67	1.71
+	-	-	+	0.00	1.05	0.64	0.50	0.34
+	-	+	-	0.00	0.12	5.14	3.17	1.88
+	-	+	+	0.00	0.00	1.29	2.17	1.71
+	+	-	-	0.00	2.32	12.85	4.67	2.05
+	+	-	+	0.00	0.00	2.31	2.17	1.88
+	+	+	-	0.00	0.00	8.87	8.18	5.13
+	+	+	+	0.00	0.00	27.63	65.78	82.74

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3), and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S7

Incidence of mutations during each generation of evolution

Mutation	Incidence (%) [a]
L43M	0.003
F47L	1.12
F65L	0.22
F65S	0.26
G66S	0.09
F72C	0.03
F72I	0.02
K102R	0.15
I129T	0.33
K141R	0.42
N145S	0.60
V164A	0.25
K167E	0.56
I168V	0.43
I172V	0.35
Y204C	0.26

[a] We estimated the incidence of each mutation shown here by linear regression using frequency data of each mutation from our directed evolution experiments in the absence of selection (see Fig. S4). We considered the slopes of the linear regression lines describing the change in mutation frequency as the mutation incidence. A mutation with an incidence of one percent means that mutation pressure alone causes the mutation to appear at a frequency of one percent in the absence of selection.

Table S8

Primers for the construction of the expression plasmid and for mutagenic PCR

Primer	Sequence
Fbad	CCC aagctt GTTTAAACGGTCTCCAGCTTG
Rbad	CCGctcgagGTATATCTCCTTCTTAAAGTTAAACAAAAT
F203	CCG ctcgag ATGGTGAGCAAGGGCGAG
R203	CCCaagcttTCACTTGTACAGCTCGTCCATGC
MutfpF	GAAGGAGATATACctcgag
MutfpR	AGACCGTTTAAACaagctt

Lower case letters refer to restriction sites.

Table S9

Mutation incidence per nucleotide site and per cycle of mutagenesis [a]

→	A	T	C	G
A	9.97^{-1}	9.46^{-5}	1.02^{-4}	2.49^{-3}
T	7.33^{-5}	9.98^{-1}	1.97^{-3}	9.16^{-5}
C	4.41^{-5}	7.72^{-4}	9.99^{-1}	3.00^{-5}
G	7.50^{-4}	2.43^{-5}	1.11^{-5}	9.99^{-1}

[a] We calculated these mutation rates from pooled sequences of non-selected control libraries sequenced by SMRT.

Table S10

Primers for engineering YFP mutants

Position	Mutation	Forward primer	Reverse primer
2	V2M	CTCGAGATGaTGAGCAAGGGCGAGGAG	CCCTTGCTCAtCATCTCGAGGTATATCTC
2	V2A	CTCGAGATGGcGAGCAAGGGCGAGGAGCT	CCCTTGCTCgCCATCTCGAGGTATATCTC
43	L43M	TACGGCAAGaTGACCCTGAAGTTCATCTGCACCAC	TTCAGGGTCAtCTTGCCGTAGGTGGCATCGCC
47	F47L	CCCTGAAGcTCATCTGCACCACCGGCAAG	GTGCAGATGAgCTTCAGGGTCAGCTTGC
65	F65L	TCGTGACCACCcTCGGCTACGGCCTGCAATGC	GCCGTAGCCGAgGGTGGTCACGAGGGTGGGC
65	F65S	GACCACCTcCGGCTACGGCCTGCAATGC	CGTAGCCGgAGGTGGTCACGAGGGTGGG
66	G66S	TGACCACCTTCaGCTACGGCCTGCAATGCTTCG	CAGGCCGTAGCtGAAGGTGGTCACGAGGGTGGG
72	F72C	GCAATGCTgCGCCCGCTACCCCGACCAC	AGCGGGCGcAGCATTGCAGGCCGTAGCC
72	F72I	CTGCAATGCaTCGCCCGCTACCCCGACC	AGCGGGCGAtGCATTGCAGGCCGTAGCCG
102	K102R	ATCTTCTTCAgGGACGACGGCAACTACAAGAC	CGTCGTCCcTGAAGAAGATGGTGCGCTCC
129	I129T	GAAGGGCAcCGACTTCAAGGAGGACGGCAAC	TTGAAGTCGgTGCCCTTCAGCTCGATGCGGTTC
141	K141R	GGGGCACAgGCTGGAGTACAACTACAACAG	ACTCCAGCcTGTGCCCCAGGATGTTGCCG
145	N145S	GGAGTACAgCTACAACAGCCACAACGTC	CTGTTGTAGcTGTACTCCAGCTTGTGCCCCAG
164	V164A	GCATCAAGGcGAACTTCAAGATCCGCCAC	CTTGAAGTTCgCCTTGATGCCGTTCTTCTG
167	K167E	TGAACTTCgAGATCCGCCACAACATCGAG	GCGGATCTcGAAGTTCACCTTGATGCCGTTC
168	I168V	AACTTCAAGgTCCGCCACAACATCGAGGAC	GTGGCGGAcCTTGAAGTTCACCTTGATGCC
172	I172V	CGCCACAACgTCGAGGACGGCAGCGTGCAG	GTCCTCGAcGTTGTGGCGGATCTTGAAG
204	Y204C	ACTACCTGAGCTgCCAGTCCGCCCTGAGCAAAG	GGGCGGACTGGcAGCTCAGGTAGTGGTTGTCGGGC
65, 66	F65L+G66S[a]	CCACCcTCaGCTACGGCCTGCAATGCTTC	GGCCGTAGCtGAgGGTGGTCACGAGGGTGGG
65, 66	F65S+G66S[a]	GACCACCTcCaGCTACGGCCTGCAATGCTTCG	GCCGTAGCtGgAGGTGGTCACGAGGGTGGG
66, 72	G66S+F72C[b]	GCAATGCTgCGCCCGCTACCCCGACCAC	AGCGGGCGgAGCATTGCAGGCCGTAGCtG
66, 72	G66S+F72I[b]	CTGCAATGCaTCGCCCGCTACCCCGACC	AGCGGGCGAtGCATTGCAGGCCGTAGCtG

Lower case letters refer to mutations.

[a] we performed whole plasmid PCR by using the ancestral YFP as a template.

[b] we performed whole plasmid PCR by using the mutant G66S as a template.

Table S11

Primers for barcoding PCR

Primer	Sequence
BC01	GGTAGG<u>AGCAAT</u>GTAAAACGACGGCCAGT
BC02	GGTAGG<u>CCTGTT</u>GTAAAACGACGGCCAGT
BC03	GGTAGG<u>GGGTTT</u>GTAAAACGACGGCCAGT
BC04	GGTAGG<u>GAAGGC</u>GTAAAACGACGGCCAGT
BC05	GGTAGG<u>ATCTCA</u>GTAAAACGACGGCCAGT
BC06	GGTAGG<u>ATGGAT</u>GTAAAACGACGGCCAGT
BC07	GGTAGG<u>ATGTCT</u>GTAAAACGACGGCCAGT
BC08	GGTAGG<u>CGTGAC</u>GTAAAACGACGGCCAGT
BC09	GGTAGG<u>TTAGGT</u>GTAAAACGACGGCCAGT
BC10	GGTAGG<u>GTGCAT</u>GTAAAACGACGGCCAGT
BC11	GGTAGG<u>AACTTT</u>GTAAAACGACGGCCAGT
BC12	GGTAGG<u>GGATCG</u>GTAAAACGACGGCCAGT
BC13	GGTAGG<u>ATAAGG</u>GTAAAACGACGGCCAGT
BC14	GGTAGG<u>ATTGGT</u>GTAAAACGACGGCCAGT
BC15	GGTAGG<u>AGTGAG</u>GTAAAACGACGGCCAGT
BC16	GGTAGG<u>CCCACC</u>GTAAAACGACGGCCAGT
BC17	GGTAGG<u>CGATGC</u>GTAAAACGACGGCCAGT
BC18	GGTAGG<u>GATAGC</u>GTAAAACGACGGCCAGT
BC19	GGTAGG<u>GTCAGA</u>GTAAAACGACGGCCAGT
BC20	GGTAGG<u>TTAAGC</u>GTAAAACGACGGCCAGT
BC21	GGTAGG<u>AACCTG</u>GTAAAACGACGGCCAGT
BC22	GGTAGG<u>CTTTGC</u>GTAAAACGACGGCCAGT
BC23	GGTAGG<u>TGGAGA</u>GTAAAACGACGGCCAGT
BC24	GGTAGG<u>AATTGT</u>GTAAAACGACGGCCAGT
BC25	GGTAGG<u>TGACGA</u>GTAAAACGACGGCCAGT
BC26	GGTAGG<u>CAAATA</u>GTAAAACGACGGCCAGT
BC27	GGTAGG<u>GTTCAG</u>GTAAAACGACGGCCAGT
BC28	GGTAGG<u>CTTCAA</u>GTAAAACGACGGCCAGT
FP_smrtF	GTAAAACGACGGCCAGTCTTTAAGAAGGAGATATACCTCGAG
pBAD849r	AATCTTCTCTCATCCGCC

Underlined sequences refer to barcoding regions.

Table S12

Adapter and barcodes used for barcoding samples

lbc28	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>GAGACTCTGTGCGCGT</u>
lbc36	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>GCGTATATCTCATGCG</u>
lbc44	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>TGTCACTCATCTGAGT</u>
lbc52	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>GCAGACTCTCACACGC</u>
lbc66	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>CGACTACGTACAGTAG</u>
lbc68	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>ACAGTATGATGTACTC</u>
lbc76	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>ACGTGAGCTCACTCGC</u>
lbc84	ATCTCTCTCTTTTCCTCCTCCTCCGTTGTTGTTGTTGAGAGAGAT <u>GACTGCACATGCACGA</u>

Underlined sequences refer to barcoding regions.

Table S13

Frequency of genotypes leading to T1 in V_1^0 during evolution

Mutation			Generation			
G66S	Y204C	F65L	II-1	II-2	II-3	II-4
-	-	-	81.43	23.26	9.73	5.39
-	-	+	0.56	0.16	0.17	0.00
-	+	-	5.59	2.17	3.07	3.17
+	-	-	11.86	62.95	25.94	16.64
-	+	+	0.11	0.00	0.34	0.16
+	-	+	0.00	0.16	1.19	1.11
+	+	-	0.45	11.32	49.49	60.54
+	+	+	0.00	0.00	10.07	13.00

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S14

Frequency of genotypes leading to T3 in V_1^0 evolution

Mutation			Generation			
G66S	Y204C	L43M	II-1	II-2	II-3	II-4
-	-	-	81.88	23.26	9.73	5.39
-	-	+	0.11	0.16	0.17	0.00
-	+	-	5.37	2.02	2.73	2.85
+	-	-	11.86	62.79	25.94	13.31
-	+	+	0.34	0.16	0.68	0.48
+	-	+	0.00	0.31	1.19	4.44
+	+	-	0.45	10.85	50.00	41.52
+	+	+	0.00	0.47	9.56	32.01

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S15

Frequency of genotypes leading to T1 in V_2^0 during evolution

Mutation			Generation			
G66S	Y204C	F65L	II-1	II-2	II-3	II-4
-	-	-	70.50	18.19	11.32	5.39
-	-	+	0.65	0.28	0.00	0.13
-	+	-	6.83	2.68	2.36	2.88
+	-	-	21.04	67.98	24.83	13.16
-	+	+	0.11	0.00	0.17	0.25
+	-	+	0.00	0.28	1.69	2.88
+	+	-	0.87	10.30	47.80	53.13
+	+	+	0.00	0.28	11.82	22.18

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S16

Frequency of genotypes leading to T2 in V_2^0 during evolution

Mutation			Generation			
G66S	Y204C	F47L	II-1	II-2	II-3	II-4
-	-	-	69.41	18.19	10.30	4.89
-	-	+	1.74	0.28	1.01	0.63
-	+	-	6.72	2.68	2.53	2.76
+	-	-	20.72	67.84	24.32	14.16
-	+	+	0.22	0.00	0.00	0.38
+	-	+	0.33	0.42	2.20	1.88
+	+	-	0.87	9.59	46.62	57.02
+	+	+	0.00	0.99	13.01	18.30

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S17

Frequency of genotypes leading to T1 in V_3^0 during evolution

Mutation			Generation			
G66S	Y204C	F65L	II-1	II-2	II-3	II-4
-	-	-	80.44	27.82	9.39	6.30
-	-	+	0.70	1.52	0.33	0.22
-	+	-	3.59	2.19	3.62	1.96
+	-	-	14.27	54.30	21.75	8.26
-	+	+	0.00	0.00	0.49	0.00
+	-	+	0.40	3.71	2.47	3.04
+	+	-	0.60	9.95	42.67	44.78
+	+	+	0.00	0.51	19.28	35.43

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3). '+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S18

Frequency of genotypes leading to T1 in V_4^0 during evolution

Mutation			Generation			
G66S	Y204C	F65L	II-1	II-2	II-3	II-4
-	-	-	81.20	26.31	8.79	6.36
-	-	+	0.36	0.13	0.35	0.91
-	+	-	3.95	1.74	3.51	5.91
+	-	-	14.25	62.42	25.31	15.45
-	+	+	0.00	0.00	0.18	0.00
+	-	+	0.24	0.40	0.70	0.91
+	+	-	0.00	8.86	57.82	53.64
+	+	+	0.00	0.13	3.34	16.82

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3). '+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S19

Frequency of genotypes leading to T2 in V_4^0 during evolution

Mutation			Generation			
G66S	Y204C	F47L	II-1	II-2	II-3	II-4
-	-	-	79.88	26.17	8.79	6.82
-	-	+	1.68	0.27	0.35	0.45
-	+	-	3.95	1.74	3.51	5.45
+	-	-	14.49	62.28	24.43	15.45
-	+	+	0.00	0.00	0.18	0.45
+	-	+	0.00	0.54	1.58	0.91
+	+	-	0.00	8.46	53.43	52.27
+	+	+	0.00	0.54	7.73	18.18

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3). '+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S20

Frequency of genotypes leading to T2 in V_1^C during evolution

Mutation			Generation				
G66S	Y204C	F47L	I-4	II-1	II-2	II-3	II-4
-	-	-	90.64	47.33	18.28	9.83	5.71
-	-	+	8.35	7.57	4.48	2.00	0.66
-	+	-	0.65	7.24	5.81	4.45	5.71
+	-	-	0.36	33.07	29.42	14.29	10.49
-	+	+	0.00	0.22	1.45	1.08	1.20
+	-	+	0.00	1.45	4.72	2.61	1.86
+	+	-	0.00	2.90	23.73	45.01	65.34
+	+	+	0.00	0.22	12.11	20.74	9.03

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3), and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S21

Frequency of genotypes leading to A1 in V_1^C during evolution

Mutation						Generation				
G66S	Y204C	F65S	K102R	N145S	V164A	I-4	II-1	II-2	II-3	II-4
-	-	-	-	-	-	94.17	50.56	20.10	9.52	4.38
-	-	-	-	-	+	1.51	2.23	0.48	0.46	0.80
-	-	-	-	+	-	2.16	1.67	0.97	0.77	0.13
-	-	-	-	+	+	0.00	0.00	0.12	0.31	0.40
-	-	-	+	-	-	1.01	0.22	0.73	0.46	0.40
-	-	-	+	-	+	0.00	0.00	0.00	0.00	0.00
-	-	-	+	+	-	0.00	0.00	0.00	0.00	0.13
-	-	-	+	+	+	0.00	0.00	0.00	0.00	0.00
-	-	+	-	-	-	0.14	0.22	0.24	0.31	0.00
-	-	+	-	-	+	0.00	0.00	0.00	0.00	0.13
-	-	+	-	+	-	0.00	0.00	0.00	0.00	0.00
-	-	+	-	+	+	0.00	0.00	0.00	0.00	0.00
-	-	+	+	-	-	0.00	0.00	0.00	0.00	0.00
-	-	+	+	-	+	0.00	0.00	0.12	0.00	0.00
-	-	+	+	+	-	0.00	0.00	0.00	0.00	0.00
-	-	+	+	+	+	0.00	0.00	0.00	0.00	0.00
-	+	-	-	-	-	0.65	6.57	5.69	4.15	3.85
-	+	-	-	-	+	0.00	0.45	0.24	0.46	0.66
-	+	-	-	+	-	0.00	0.33	0.48	0.00	0.53
-	+	-	-	+	+	0.00	0.00	0.36	0.77	0.80
-	+	-	+	-	-	0.00	0.11	0.12	0.00	0.13
-	+	-	+	-	+	0.00	0.00	0.00	0.00	0.00
-	+	-	+	+	-	0.00	0.00	0.00	0.00	0.00
-	+	-	+	+	+	0.00	0.00	0.36	0.15	0.80
-	+	+	-	-	-	0.00	0.00	0.00	0.00	0.00
-	+	+	-	-	+	0.00	0.00	0.00	0.00	0.00
-	+	+	-	+	-	0.00	0.00	0.00	0.00	0.00
-	+	+	-	+	+	0.00	0.00	0.00	0.00	0.13
-	+	+	+	-	-	0.00	0.00	0.00	0.00	0.00
-	+	+	+	-	+	0.00	0.00	0.00	0.00	0.00
-	+	+	+	+	-	0.00	0.00	0.00	0.00	0.00
-	+	+	+	+	+	0.00	0.00	0.00	0.00	0.00
+	-	-	-	-	-	0.36	27.62	26.88	9.83	6.91
+	-	-	-	-	+	0.00	4.90	3.27	1.69	1.33
+	-	-	-	+	-	0.00	0.67	0.48	0.92	0.66
+	-	-	-	+	+	0.00	0.00	0.85	0.31	0.66
+	-	-	+	-	-	0.00	0.45	0.00	0.46	0.13
+	-	-	+	-	+	0.00	0.89	1.09	1.08	0.80
+	-	-	+	+	-	0.00	0.00	0.00	0.00	0.00
+	-	-	+	+	+	0.00	0.00	0.24	0.00	0.27
+	-	+	-	-	-	0.00	0.00	0.36	0.46	0.40
+	-	+	-	-	+	0.00	0.00	0.24	0.31	0.13
+	-	+	-	+	-	0.00	0.00	0.12	0.15	0.13
+	-	+	-	+	+	0.00	0.00	0.24	0.77	0.66

+	-	+	+	-	-	0.00	0.00	0.24	0.00	0.13
+	-	+	+	-	+	0.00	0.00	0.12	0.00	0.00
+	-	+	+	+	-	0.00	0.00	0.00	0.00	0.00
+	-	+	+	+	+	0.00	0.00	0.00	0.92	0.13
+	+	-	-	-	-	0.00	2.23	22.88	34.25	35.99
+	+	-	-	-	+	0.00	0.00	0.85	3.69	4.65
+	+	-	-	+	-	0.00	0.11	1.45	3.07	3.32
+	+	-	-	+	+	0.00	0.00	2.18	3.07	5.44
+	+	-	+	-	-	0.00	0.00	0.12	0.00	0.40
+	+	-	+	-	+	0.00	0.00	0.12	0.15	0.40
+	+	-	+	+	-	0.00	0.00	0.12	0.15	0.27
+	+	-	+	+	+	0.00	0.11	0.24	1.54	1.73
+	+	+	-	-	-	0.00	0.00	0.48	1.54	1.73
+	+	+	-	-	+	0.00	0.00	0.48	0.77	1.20
+	+	+	-	+	-	0.00	0.00	0.73	0.46	1.06
+	+	+	-	+	+	0.00	0.11	1.82	3.53	4.91
+	+	+	+	-	-	0.00	0.00	0.00	0.00	0.00
+	+	+	+	-	+	0.00	0.00	0.12	0.15	0.13
+	+	+	+	+	-	0.00	0.00	0.00	0.31	0.27
+	+	+	+	+	+	0.00	0.56	4.24	13.06	12.88

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3),
and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S22

Frequency of genotypes leading to A2 in V_1^C during evolution

Mutation					Generation				
G66S	Y204C	F72I	K167E	I172V	I-4	II-1	II-2	II-3	II-4
-	-	-	-	-	94.82	51.78	21.31	11.06	5.31
-	-	-	-	+	1.80	2.45	0.61	0.31	0.53
-	-	-	+	-	2.16	0.67	0.85	0.46	0.00
-	-	-	+	+	0.14	0.00	0.00	0.00	0.27
-	-	+	-	-	0.07	0.00	0.00	0.00	0.00
-	-	+	-	+	0.00	0.00	0.00	0.00	0.00
-	-	+	+	-	0.00	0.00	0.00	0.00	0.27
-	-	+	+	+	0.00	0.00	0.00	0.00	0.00
-	+	-	-	-	0.65	5.35	6.42	4.45	5.31
-	+	-	-	+	0.00	1.89	0.61	0.92	0.66
-	+	-	+	-	0.00	0.11	0.12	0.00	0.40
-	+	-	+	+	0.00	0.11	0.12	0.15	0.27
-	+	+	-	-	0.00	0.00	0.00	0.00	0.13
-	+	+	-	+	0.00	0.00	0.00	0.00	0.00
-	+	+	+	-	0.00	0.00	0.00	0.00	0.13
-	+	+	+	+	0.00	0.00	0.00	0.00	0.00
+	-	-	-	-	0.29	30.51	29.78	14.29	9.43
+	-	-	-	+	0.07	3.67	3.75	1.38	0.93
+	-	-	+	-	0.00	0.33	0.48	0.77	0.27
+	-	-	+	+	0.00	0.00	0.00	0.00	0.00
+	-	+	-	-	0.00	0.00	0.00	0.31	0.66
+	-	+	-	+	0.00	0.00	0.12	0.00	0.53
+	-	+	+	-	0.00	0.00	0.00	0.00	0.40
+	-	+	+	+	0.00	0.00	0.00	0.15	0.13
+	+	-	-	-	0.00	2.45	33.41	54.84	38.91
+	+	-	-	+	0.00	0.56	1.82	4.45	14.87
+	+	-	+	-	0.00	0.00	0.48	1.08	2.92
+	+	-	+	+	0.00	0.11	0.00	0.31	5.44
+	+	+	-	-	0.00	0.00	0.00	0.61	1.99
+	+	+	-	+	0.00	0.00	0.00	0.46	1.46
+	+	+	+	-	0.00	0.00	0.00	0.77	1.99
+	+	+	+	+	0.00	0.00	0.12	3.23	6.77

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3), and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).

'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S23

Frequency of genotypes leading to T1 in V_2^C during evolution

Mutation			Generation				
G66S	Y204C	F65L	I-4	II-1	II-2	II-3	II-4
-	-	-	96.86	58.82	26.59	12.92	12.84
-	-	+	1.69	3.24	1.56	1.11	0.90
-	+	-	0.97	6.84	4.98	6.46	5.42
+	-	-	0.36	27.13	38.26	23.25	16.09
-	+	+	0.12	0.96	0.31	1.11	0.36
+	-	+	0.00	1.56	6.69	4.61	3.98
+	+	-	0.00	1.32	15.55	29.52	30.02
+	+	+	0.00	0.12	6.07	21.03	30.38

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3), and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.

Table S24

Frequency of genotypes leading to A3 in V_3^C during evolution

Mutation					Generation				
G66S	Y204C	F47L	I129T	K141R	I-4	II-1	II-2	II-3	II-4
-	-	-	-	-	87.92	53.19	19.02	6.02	5.73
-	-	-	-	+	1.23	0.43	1.53	0.26	0.38
-	-	-	+	-	0.80	1.08	0.34	0.65	0.19
-	-	-	+	+	0.00	0.00	0.00	0.00	0.38
-	-	+	-	-	8.18	7.14	3.74	2.23	2.10
-	-	+	-	+	0.22	0.00	0.00	0.26	0.19
-	-	+	+	-	0.07	0.22	0.51	0.13	0.19
-	-	+	+	+	0.00	0.11	0.00	0.13	0.00
-	+	-	-	-	1.09	11.14	3.90	2.62	2.29
-	+	-	-	+	0.00	0.00	0.51	0.65	0.38
-	+	-	+	-	0.00	0.32	0.17	0.26	0.57
-	+	-	+	+	0.00	0.00	0.85	0.00	1.34
-	+	+	-	-	0.14	1.19	1.19	0.79	1.34
-	+	+	-	+	0.00	0.11	0.34	0.39	0.00
-	+	+	+	-	0.00	0.00	0.34	0.26	0.00
-	+	+	+	+	0.00	0.22	0.68	0.39	0.95
+	-	-	-	-	0.36	16.11	10.19	4.97	2.10
+	-	-	-	+	0.00	0.43	0.34	0.26	0.00
+	-	-	+	-	0.00	0.43	0.34	0.39	0.38
+	-	-	+	+	0.00	0.11	0.00	0.13	0.19
+	-	+	-	-	0.00	1.73	4.58	2.36	3.63
+	-	+	-	+	0.00	0.22	0.85	0.52	0.76
+	-	+	+	-	0.00	0.11	0.51	0.13	0.76
+	-	+	+	+	0.00	0.11	1.53	0.79	2.10
+	+	-	-	-	0.00	2.59	6.79	3.53	3.24
+	+	-	-	+	0.00	0.11	0.85	0.65	1.34
+	+	-	+	-	0.00	0.32	1.02	0.65	0.19
+	+	-	+	+	0.00	0.11	1.02	0.52	0.57
+	+	+	-	-	0.00	1.08	5.43	4.58	3.44
+	+	+	-	+	0.00	0.11	3.57	3.27	7.25
+	+	+	+	-	0.00	0.22	0.34	1.83	1.53
+	+	+	+	+	0.00	1.08	29.54	60.34	56.49

Note that the data in green indicate the inferred trajectories with highest probability (see Supplementary Text S3),
and the data in yellow indicate the frequency of cryptic variation at the end of phase I (I-4).
'+' and '-' indicate an incidence or not incidence of a corresponding mutation.