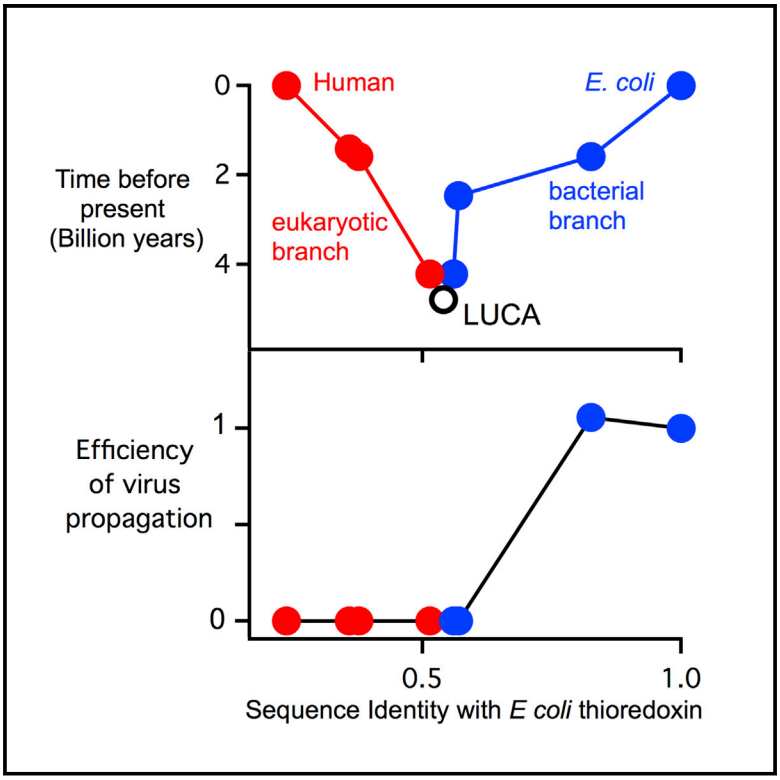
Article

Cell Reports

# Using Resurrected Ancestral Proviral Proteins to Engineer Virus Resistance

## Graphical Abstract




## Highlights

- Modern to ancestral replacement of a proviral factor prevents virus propagation
- This result points to an approach to the engineering of virus resistance
- This approach could potentially be applied to the engineering of plant virus resistance
- Ancestral reconstruction may probe the evolution of biomolecular interactions

## Authors

Asunción Delgado, Rocio Arco, Beatriz Ibarra-Molero, Jose M. Sanchez-Ruiz

## Correspondence

sanchezr@ugr.es

## In Brief

Proviral factors are host proteins hijacked by viruses for processes essential for virus propagation. Pathogens and their hosts co-evolve. Delgado et al. show that replacing a proviral factor with a functional ancestral analog may prevent viral propagation thus pointing to an approach to the engineering of virus resistance.

Delgado et al., 2017, Cell Reports *19*, 1247–1256
May 9, 2017 
http://dx.doi.org/10.1016/j.celrep.2017.04.037

Cell Reports
Article



# Using Resurrected Ancestral Proviral Proteins to Engineer Virus Resistance

Asunción Delgado,[1] Rocio Arco,[1] Beatriz Ibarra-Molero,[1] and Jose M. Sanchez-Ruiz[1,2,*]
[1]Facultad de Ciencias, Departamento de Quimica Fisica, Universidad de Granada, 18071 Granada, Spain
[2]Lead Contact
*Correspondence: sanchezr@ugr.es
http://dx.doi.org/10.1016/j.celrep.2017.04.037

## SUMMARY

**Proviral factors are host proteins hijacked by viruses for processes essential for virus propagation such as cellular entry and replication. Pathogens and their hosts co-evolve. It follows that replacing a proviral factor with a functional ancestral form of the same protein could prevent viral propagation without fatally compromising organismal fitness. Here, we provide proof of concept of this notion. Thioredoxins serve as general oxidoreductases in all known cells. We report that several laboratory resurrections of Precambrian thioredoxins display substantial levels of functionality within *Escherichia coli*. Unlike *E. coli* thioredoxin, however, these ancestral thioredoxins are not efficiently recruited by the bacteriophage T7 for its replisome and therefore prevent phage propagation in *E. coli*. These results suggest an approach to the engineering of virus resistance. Diseases caused by viruses may have a devastating effect in agriculture. We discuss how the suggested approach could be applied to the engineering of plant virus resistance.**

## INTRODUCTION

It has been estimated that direct losses caused by pathogens, animals, and weeds amount to 20%–40% of the global agricultural productivity (Savary et al., 2012). Plant viruses that affect economically important crops, in particular, often become a cause of severe hardship for large numbers of people (Strange and Scott, 2005; Rybicki, 2015). Several approaches to the engineering of plant viral resistance have been explored, and some crops with engineered resistance are grown commercially (Dasgupta et al., 2003; Prins et al., 2008; Collinge et al., 2010; Galvez et al., 2014; Whitham and Hajimorad, 2016). Briefly, introduction of viral genes into plants has been found in some cases to interfere with essential steps in the virus life cycle, thus giving rise to the so-called pathogen-derived resistance. Plant resistance genes encoding products that detect viral gene products and trigger resistance responses are known and have been shown to confer resistance to other plant species when used as transgenes. Finally, RNA-silencing systems in plants can be engineered to target viral genomes and transcripts.

Some host proteins play a permissive role for viral infection and some play a restrictive role ("proviral" and "antiviral" factors, respectively). There are abundant proviral factors as shown by recent studies (Friedel and Haas, 2011; Wang, 2015; Enard et al., 2016; Wilke and Sawyer, 2016; Sasvari and Nagy, 2016) thus pointing to many potential targets for resistance engineering. Very recent work has used gene-editing methodologies to suppress (knock out or render non-functional) proviral factors and achieve resistance (Chandrasekaran et al., 2016; Pyott et al., 2016). However, the applicability of this suppression approach is limited by the fact that proviral factors carry out "normal" functions in the host, and their suppression may therefore compromise host fitness and survival.

Pathogens and their hosts co-evolve. Proviral factors can be seen as proteins hijacked by viruses that are indeed losing the arms race (Daugherty and Malik, 2012; Demogines et al., 2013; Ng et al., 2015), as they have not yet adapted to evade viral hijacking. Therefore, it stands to reason that if a proviral factor in a host were exchanged for a functional analog, the fitness of the host might be only mildly affected, whereas the fitness of the virus could be substantially impaired. We hypothesize that a means to engineer this exchange is to take an ancestral version of the proviral host protein. Because the ancestral version would have co-evolved with a very different set of pathogens, it would be evolutionarily less simple for contemporary pathogens to swap to use the ancestral protein. That is, evasion and virus resistance can be achieved by using resurrected ancestral proviral proteins that are expected to maintain sufficient functionalities while being distant enough that they should represent a difficult challenge for the virus in terms of recruitment and adaptation. One can thus relatively preserve host fitness but impair viral fitness. Here, we present a test/demonstration of this approach using the infection of *Escherichia coli* by the bacteriophage T7 as a simple model. As explained below in some detail, the molecular features of the relevant intermolecular interactions make this system particularly attractive as a proof of concept of our proposal. Furthermore, in the Discussion section, we will speculate as to the greater relevance of the proposed approach in engineering viral resistance, and specifically, we will outline how it could be applied to engineering plant viral resistance.

Bacteriophage T7 recruits *E. coli* thioredoxin to be a part of its "minimalist" (four proteins) replisome (Hamdan and Richardson,

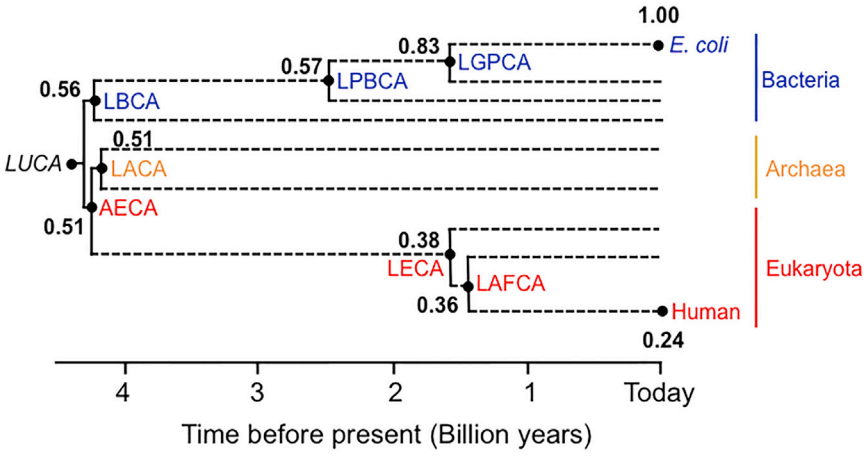



**Figure 1. Ancestral and Modern Thioredoxins Studied in This Work**

Schematic phylogenetic tree showing the Precambrian phylogenetic nodes targeted in this work and their geological ages (Perez-Jimenez et al., 2011). The nodes targeted for reconstruction are: LBCA (last bacterial common ancestor), LPBCA (last common ancestor of cyanobacterial, deinococcus, and thermus groups), LGPCA (the last common ancestor of γ-proteobacteria), AECA (archaeal-eukaryotic common ancestor), LACA (last archaeal common ancestor), LECA (last eukaryotic common ancestor), and LAFCA (last common ancestor of fungi and animals). The numbers alongside the nodes stand for the values of the identity of the reconstructed sequences with the sequence of *E. coli* thioredoxin 1. The reconstructed sequences are given in Figure 1C in Ingles-Prieto et al. (2013).

2009). When recruited, thioredoxin functions as a processivity factor for the gp5 DNA polymerase. It binds strongly to gp5 and increases its processivity from 1–15 to several hundred nucleotides per binding event, likely by suppressing hoping on and off the DNA (Etson et al., 2010). Binding to gp5 is mediated by specific interactions with a unique 76-residue fragment known as the thioredoxin-binding domain (for instance, see Figure 1 in Akabayov et al. [2010]). Such specific interaction appears to reconfigure gp5 to enhance contact with DNA and create docking sites for the other proteins of the replisome (Hamdan and Richardson, 2009; Lee and Richardson, 2011). The thioredoxin-binding domain of the phage gp5 is not found in other members of this polymerase family and is likely the result of the evolutionary adaptation of bacteriophage T7 for efficient propagation in *E. coli*. In fact, with a value for the dissociation constant of 5 nM, the interaction between *E. coli* thioredoxin and the thioredoxin-binding domain in gp5 is actually one of the highest-affinity interactions involving replication proteins (Hamdan and Richardson, 2009). The very high specificity of this interaction makes it plausible that at least some ancestral forms of thioredoxin may not be recruited by bacteriophage T7.

On the other hand, the "normal" role of thioredoxin in all known cells is to serve as a general oxido-reductase (Holmgren, 1995). As such, thioredoxin is involved in a diversity of cellular processes, and a large number of potential substrates for thioredoxin have been identified using a variety of approaches (Collet and Messens, 2010; Kumar et al., 2004). This wide substrate scope is likely achieved through a molecular mechanism that does not impose highly specific intermolecular interactions. In fact, the disulfide bridge of the conserved CGPC active site motif in thioredoxins is solvent-exposed and protrudes somewhat from the structure of the protein (thioredoxin has sometimes been referred to as a "male enzyme") (Holmgren et al., 1975). This structural feature should promote the processing of a diversity of substrates. Certainly, any change in a complex system involving interacting parts or processes is likely to generate a less fit system (Kirschner and Gerhart, 2005) and, consequently, replacing a modern protein in a modern organism with a representation of one of its ancestors is expected to bring about a fitness cost in many instances (Hobbs et al., 2015). Still, given that the most common oxido-reductase function of thioredoxin involves molecular interactions that are not highly specific, it would not be unreasonable to expect ancestral thioredoxins to show some degree of functionality within an *E. coli* cell.

Overall, it seems plausible that at least some ancient thioredoxins could functionally replace the modern thioredoxin within *E. coli* and, in addition, make *E. coli* resistant to the propagation of bacteriophage T7. Certainly, an experimental test of this possibility would seem to face an insurmountable barrier, i.e., the fact that ancient proteins no longer exist. However, phylogenetic analyses of modern protein sequences can be used to derive plausible approximations to the sequences of proteins in extinct organisms (Pauling and Zuckerkandl, 1963). Indeed, proteins encoded by such reconstructed ancestral sequences (i.e., "resurrected" ancestral proteins, to use the common jargon of the field) have been extensively used in the last ~20 years to address important problems in molecular evolution (for reviews, see Benner et al., 2007; Liberles, 2007; Harms and Thornton, 2010, 2013; Risso et al., 2014; Merkl and Sterner, 2016). Here, we specifically use several resurrected Precambrian thioredoxins that actually span the ~4 billion years of evolution of life on Earth (see Figure 1 for a description of the corresponding phylogenetic nodes and the values of the identity of the reconstructed sequences with the sequence of *E. coli* thioredoxin). These putative Precambrian thioredoxins have been previously prepared and exhaustively characterized (Perez-Jimenez et al., 2011; Ingles-Prieto et al., 2013; Romero-Romero et al., 2016). They are properly folded and share the thioredoxin fold. They are also stable, enzymatically active, and their in vitro properties define convincing evolutionary narratives as we have previously discussed in detail (Perez-Jimenez et al., 2011; Ingles-Prieto et al., 2013; Romero-Romero et al., 2016). For comparison, *E. coli* thioredoxin and human thioredoxin are also included in this study.

## RESULTS

### *E. coli* Strains Used in This Work

The cytoplasm of *Escherichia coli* actually contains two thioredoxins (Miranda-Vizuete et al., 1997; El Hajjaji et al., 2009): thioredoxin 1, which we have referred to in the Introduction (and shall

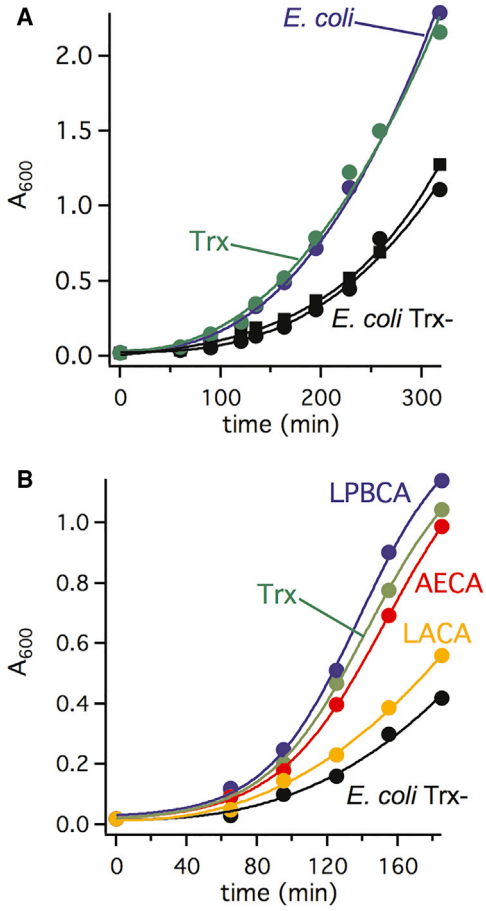



**Figure 2. Growth Curves for *E. coli* Strains**

(A) Representative growth curves of DHB4 (wild-type), FA41 (thioredoxin minus), FA41(DE3), and FA41(DE3) (pET30a(+)::*trxA*) (thioredoxin minus complemented with thioredoxin 1 in a plasmid). Overnight cultures were diluted 1/200 in fresh LB medium, and absorbance at 600 nm was monitored with time.

(B) Representative growth curves of FA41(DE3) and of this strain complemented with either thioredoxin 1, or the ancestral nodes LPBCA, AECA, and LACA in plasmids.

continue to refer to) simply as *E. coli* thioredoxin, and thioredoxin 2, which has a zinc-binding domain and is induced under oxidative stress conditions. The studies described here employ the receptor strain FA41, an *E. coli* strain deficient in the two thioredoxins described in this organism. FA41 is able to grow, albeit at a significantly slower rate than the "wild type" strain DHB4. This capability to grow slowly in the absence of thioredoxins is likely due, at least in part, to the glutaredoxin pathway, which can substitute for some of the functions associated to the thioredoxin pathway (absent in FA41). The slower-growth phenotype is in fact consistent with that of other thioredoxin minus strains described in the literature (Holmgren et al., 1978; Lim et al., 1985). From now on, we shall refer to the FA41 and DHB4 *E. coli* strains as *E. coli* $Trx^{-}$ and *E. coli* $Trx^{+}$, respectively. All studies reported here are based on the complementation of the *E. coli* $Trx^{-}$ strain with plasmids containing thioredoxin genes. We decided to use this complementation approach, rather than allelic replacement, to avoid potential effects on fitness linked to regulatory changes triggered by stress conditions in some of the fitness experiments. Expression of thioredoxins was under a T7 promoter in pET derivative plasmids. This system is "leaky," and even under non-inducing conditions there is a basal expression level from the T7 promoter (see Experimental Procedures for details), which is sufficient to compensate the growth deficiency of the *E. coli* $Trx^{-}$ strain with the plasmid encoding for the gene of *E. coli* thioredoxin 1 (Figure 2A). This suggests that the basal expression level leads to thioredoxin concentrations within the complemented *E. coli* $Trx^{-}$ strain comparable to the normal thioredoxin concentration within *E. coli* ($\sim 10^{2}$ μM) (Jacob et al., 2011).

Finally, FA41 and DHB4 are male *E. coli* strains, a feature linked to the presence of an F′ factor. Infection of male strains by the bacteriophage T7 is abortive and does not lead to subsequent virus propagation. Therefore, the *E. coli* $Trx^{-}$ and *E. coli* $Trx^{+}$ strains in our virus propagation experiments are F′ minus versions of DHB4 and FA41, in which the F′ factor has been cured following a standard protocol (see Experimental Procedures for details).

## Functionality of Precambrian Thioredoxins within *E. coli*

The resurrected Precambrian thioredoxins studied in this work have been previously shown to display redox activity in vitro, not only on the basis of regular solution assays, but also using a single-molecule atomic force spectroscopy approach that provides a high-resolution structural depiction of the catalytic process (Perez-Jimenez et al., 2011). Indeed, the resurrected Precambrian thioredoxins studied here show levels of in vitro redox activity that are similar, or even higher in some cases, than those of their modern human and *E. coli* counterparts. This in vitro activity, however, does not necessarily imply substantial functionality within a modern organism. Thioredoxin is involved in vivo in a diversity of cellular processes, and a large number of proteins have been identified as thioredoxin substrates (Collet and Messens, 2010; Kumar et al., 2004). Precambrian thioredoxins were adapted to ancestral extra- and intracellular environments and, in particular, the substrates they acted upon were likely different from the thioredoxin substrates in modern organisms.

In order to probe the functionality of the ancestral thioredoxins in vivo within a modern organism, we determined the extent to which they can rescue the slow growth phenotype of the *E. coli* $Trx^{-}$ strain. To this end, we complemented the *E. coli* $Trx^{-}$ strain with plasmids containing the genes of Precambrian thioredoxins (and also, the gene of human thioredoxin and, as a control, the gene of *E. coli* thioredoxin) and measured the generation time. Compensation of the growth deficiency was indeed observed (Figure 2B) and to an extent that roughly correlates with the "evolutionary distance" with *E. coli* thioredoxin as measured by the fraction of amino acid sequence identity (Figure 3). In fact, even the thioredoxins that are more evolutionary distant from *E. coli* thioredoxin can compensate growth deficiency to some measurable extent, as shown by generation times significantly shorter than that for the *E. coli* $Trx^{-}$ strain (Figure 3). That is, the thioredoxins from the eukaryotic branch,

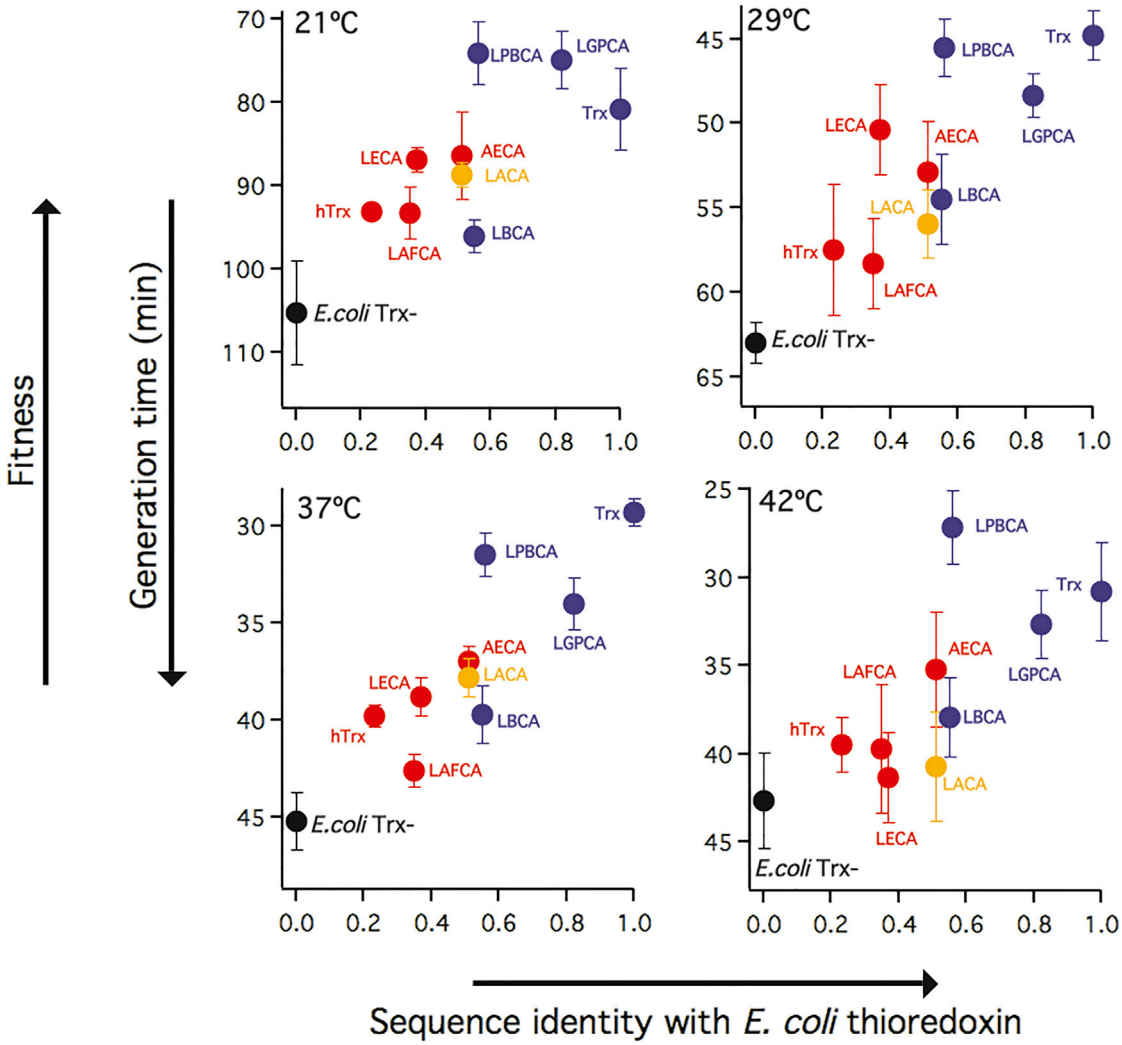



**Figure 3. Functionality of Precambrian Thioredoxins within *E. coli***

Generation times for *E. coli* Trx$^-$ strain complemented with plasmids containing the genes of laboratory resurrections of Precambrian thioredoxins, as well as the genes of human and *E. coli* thioredoxins. For comparison, generation times of the non-complemented *E. coli* Trx$^-$ strain are also shown. Experiments were performed at four different temperatures. In all cases, four to six independent determinations were carried out and the average values are shown with error bars that correspond to the calculated SDs. Data are plotted versus the evolutionary distance to the *E. coli* thioredoxin 1 as measured by sequence identity. Such calculation is of course not possible for the non-complemented *E. coli* Trx$^-$ strain. For representation purposes, a sequence identity of zero has been assigned to data for the non-complemented strain. The data labeled Trx correspond to the *E. coli* Trx$^-$ strain complemented with *E. coli* thioredoxin 1.

including human thioredoxin, as well as the thioredoxins corresponding to ~4 billion years nodes (last bacterial common ancestor [LBCA], last archaeal common ancestor [LACA], archaeal-eukaryotic common ancestor [AECA]; see Figure 1) display some degree of functionality within *Escherichia coli*.

As shown in Figure 3, we actually performed generation time determinations at different temperatures within the range 21°C–42°C in an attempt to determine whether any of the observed effects on generation time could be linked to protein stability. Resurrected Precambrian thioredoxins have been previously shown to display much enhanced stability when compared with their modern mesophilic counterparts (Perez-Jimenez et al., 2011; Romero-Romero et al., 2016). This high stability may plausibly reflect that Precambrian (in particular, Archaean) oceans were hot (Knauth, 2005; Robert and Chaussidon, 2006; Gaucher et al., 2008; Romero-Romero et al., 2016). Still, last common ancestor of cyanobacterial, deinococcus, and thermus groups (LPBCA) thioredoxin is the most stable among the studied resurrected proteins (denaturation temperature of ~123°C) (Perez-Jimenez et al., 2011) and this might specifically reflect a high temperature local environment, as the LPBCA node is the ancestor of many modern thermophilic organisms (Perez-Jimenez et al., 2011). Interestingly, LPBCA thioredoxin seems to deviate somewhat, in particular at 42°C, from the correlation between compensation of the growth deficiency and evolutionary distance from *E. coli* thioredoxin (Figure 3).

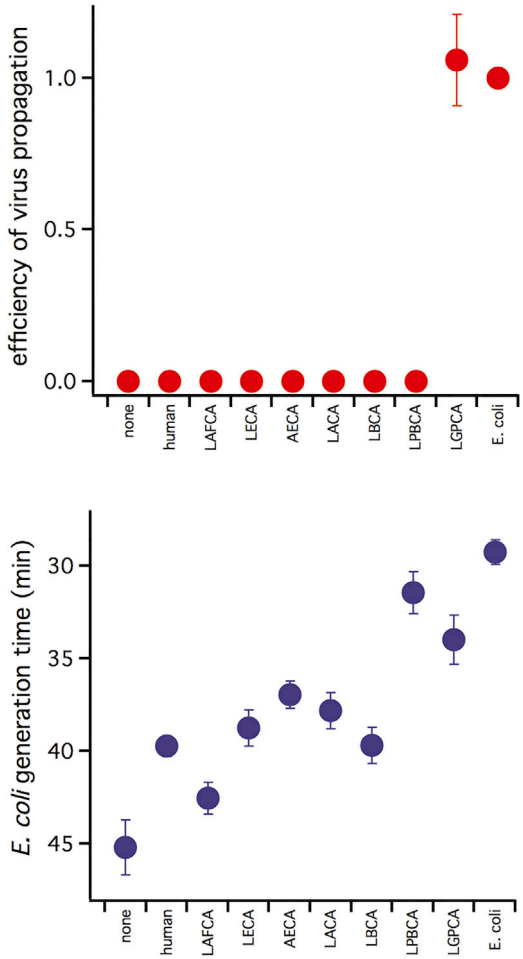



**Figure 4. Effect of Precambrian Thioredoxins on the Propagation of Bacteriophage T7**

Efficiency of bacteriophage T7 propagation in *E. coli* complemented with human and ancestral thioredoxins (top panel). Data are given as percentage of the maximal efficiency with *E. coli* thioredoxin. Experimental details are given in the Experimental Procedures. Generation times at 37°C are displayed in the bottom panel. Error bars represent the SD from the average.

The deviation is in the direction of decreasing generation time, i.e., in the direction of increasing fitness.

## Effect of Precambrian Thioredoxins on the Propagation of Bacteriophage T7 within *E. coli*

We subsequently tested whether the bacteriophage T7 could propagate within *E. coli* Trx$^-$ strains complemented with plasmids bearing ancestral thioredoxin genes. To this end, we performed experiments in which a given volume of a cell suspension was mixed with the same volume of a dilution from a stock virus solution, mixed with agar, and plated. After overnight incubation at 37°C, plaques are observed in a lawn of growing cells where virus infection and subsequent virus propagation to neighboring cells has occurred. The numbers of plaques can be easily converted into numbers of plaque-forming units per mL in the stock virus solution. Following standard protocols, we performed experiments for each strain with serial 10-fold dilutions of the virus stock solution. Furthermore, we made a substantial number (on the order of ten per dilution) of replicate plates for each dilution. For the *E. coli* Trx$^+$ strain and for an *E. coli* Trx$^-$ strain complemented with *E. coli* thioredoxin, plaques were observed even at high phage dilutions and a value of $\sim 10^9$ plaque-forming units per mL could be calculated from the plaque counting. Similar results were obtained for an *E. coli* Trx$^-$ strain complemented with the ancestral LGPCA thioredoxin, which is evolutionary close to *E. coli* thioredoxin (sequence identity 0.83: Figure 1). However, for LPBCA thioredoxin (sequence identity with *E. coli* thioredoxin of 0.57), plaques were observed only at the highest virus concentration used and, even at those concentrations, they were observed only occasionally. Specifically, out of 36 plates made with the *E. coli* Trx$^-$ strain complemented with LPBCA thioredoxin, only nine displayed plaques and, in fact, seven plates showed barely one to two plaques. These results point to $\sim 10^2$ or less plaque-forming units per mL, a value that is many orders of magnitude below the $\sim 10^9$ plaque-forming units found for *E. coli* thioredoxin. The ratio of these two numbers yields an efficiency of virus propagation for LPBCA thioredoxin of $\sim 10^{-7}$ or less (with the propagation efficiency for *E. coli* thioredoxin set to unity). We performed a more detailed analysis of the efficiency of virus propagation with LPBCA thioredoxin on the basis of an additional set of data obtained using the same phage stock under the same conditions. Of the 15 plates prepared, 9 showed no plaques, 4 showed 1 plaque, one showed 2 plaques and one showed 3 plaques. Fitting this data to the Poisson distribution yields and average number of plaques per plate of 0.56 with an associated standard error (estimated by bootstrapping) of 0.20. The amount of phage present in each plate was equivalent to 10 μL of the stock solution, leading to 56 ± 20 plaque-forming units per mL with LPBCA thioredoxin. The same stock solution had $1.3 \times 10^9$ plaque-forming units per mL when used to infect the strain complemented with *E. coli* thioredoxin. The efficiency of virus propagation with the ancestral protein is, therefore, $\sim 5 \times 10^{-8}$.

No plaques could be detected, not even at the highest virus concentration used when an *E. coli* Trx$^-$ strain was complemented with any of all the other thioredoxins studied in this work. That is, in these experiments, the efficiency of virus propagation for LBCA, LACA, AECA, last eukaryotic common ancestor (LECA), last common ancestor of fungi and animals (LAFCA), and human thioredoxins is indistinguishable from zero. Therefore, a plot of efficiency of virus propagation versus evolutionary distance to *E. coli* thioredoxin (Figure 4) calls to mind a step function with the step comparatively close to *E. coli* thioredoxin. This is in sharp contrast with the much less abrupt variation seen with the generation time (Figures 3 and 4). The implication is that ancestral thioredoxins that show substantial redox functionality within *E. coli* (as determined by their capability to rescue the slower-growth phenotype of an *E. coli* Trx$^-$ strain) do not allow (or substantially limit) phage propagation in *E. coli*.

The experiments described in the preceding paragraphs where carried out under basal thioredoxin expression levels. For some selected cases, we also performed experiments under overexpression conditions that may lead to a substantial

fraction of the total protein concentration in the cell being thioredoxin. Specifically, we used the maximum isopropyl β-D-1-thiogalactopyranoside (IPTG) concentration (10 μM) that leads to overexpression without compromising growth. We found efficiencies of virus propagation that were increased with respect to the basal-level values discussed in the preceding paragraph, but also that were still substantially smaller than the propagation efficiency for *E. coli* thioredoxin (taken as the value of reference of unity). Specifically, we found efficiencies of virus propagation of approximately $10^{-3}$, $10^{-5}$, $10^{-5}$, and $10^{-6}$ for LPBCA, LBCA, LACA, and human thioredoxins, respectively. These results seem to disfavor the possibility that the ancestral thioredoxins are efficiently recruited by the phage and the low propagation efficiency is exclusively due to very low processivity of the replisomes formed with the ancestral proteins. They also disfavor that the occasional plaques detected with LPBCA thioredoxin under basal expression and high virus concentration (preceding paragraph) are due to our virus sample containing a few mutant phages that can recruit the ancestral thioredoxin. Note that, in these two scenarios, increasing thioredoxin concentration within *E. coli* would not be expected to increase propagation efficiency. Overall, it seems more likely at this stage that the very low virus propagation efficiency found for most ancestral thioredoxins is primarily due to their association with the other molecular components of the phage replisome being highly improbable for thermodynamic or kinetic reasons.

We also explored the possibility that the occasional plaques detected with LPBCA thioredoxin were related to virus evolution during the efficiency of propagation experiments. We thus isolated phages from eight such plaques and tested in liquid culture their capability to lyse the *E. coli* $Trx^-$ strain complemented either with *E. coli* thioredoxin or with LPBCA thioredoxin. The eight virus samples were found to efficiently lyse the strain complemented with *E. coli* thioredoxin, as shown by a sharp decrease of the absorbance at 600 nm from its initial value of ~0.5. On the other hand, strong lysis was not observed with the strain complemented with the ancestral LPBCA thioredoxin. In fact, for five of the virus samples tested, absorbance at 600 nm was found to increase with time in a manner similar to the control growth curve (i.e., the absorbance versus time profile determined in the absence of virus). On the other hand, for three of the samples, absorbance at 600 nm changed comparatively little with time, indicating arrested growth and likely some degree of lysis. Virus evolution during the efficiency of propagation experiment is a plausible explanation for the results obtained with these three virus samples.

Finally, we were interested in assessing the extent to which our conclusions could be affected by the fact that natural viral populations are typically quasi-species as a result of high mutation rates. For this purpose, we used the T7 Line 1, a synthetic T7 collection of mutants obtained after a growth for 200 generations in the presence of a mutagen. T7 Line 1 has a genomic rate of four nonlethal mutations per generation, which means three orders of magnitude larger than the average value for DNA organisms such as T7 (Springman et al., 2010) (see Experimental Procedures). Particularly, T7 Line 1 comprises approximately seven missense mutations with frequencies of $\geq 0.75$ in and around the polymerase gp5 gene. We found that T7 Line 1 propagated with similar efficiency in an *E. coli* $Trx^+$ strain and in an *E. coli* $Trx^-$ strain complemented with *E. coli* thioredoxin and complemented with the ancestral LGPCA thioredoxin. The *E. coli* $Trx^-$ strain complemented with LPBCA thioredoxin produced plaques with an efficiency of six to seven orders of magnitude lower than *E. coli* thioredoxin. No progeny from T7 Line 1 was obtained with the rest of ancestral thioredoxins and with human thioredoxin. Therefore, the results obtained with the T7 wild-type (Figure 4 and first paragraph in this section) and with T7 Line 1 are essentially identical.

## DISCUSSION

In the present work, we have provided experimental evidence that a back-to-the-ancestor replacement on a proviral factor may bypass virus adaptation for recruitment of the modern protein and confer virus resistance. Of course, for this approach to work, the ancestral protein must display functionality within the modern organism, in such a way that organismal fitness is not fatally compromised by the replacement. In the specific case studied here, we had to "travel back in time" to ~2.5 billion years ago (the age of the LPBCA node: see Figure 1) to find an ancestral thioredoxin that prevents propagation of bacteriophage T7 in *E. coli*. This long backward time-travel may reflect that the interaction of the thioredoxin-binding-domain of the virus gp5 polymerase with *E. coli* thioredoxin involves the burial of the active site cysteines (Doublié et al., 1998). Consequently, many of the mutations at the interaction surface that could easily prevent recruitment may also impair the normal redox function of thioredoxin and, therefore, will not be accepted. Note, nevertheless, that different molecular interaction scenarios will lead to different outcomes. For instance, if recruitment and "normal" functionality of the protein within the modern organism depend on interactions at different molecular surfaces, we may plausibly expect that even comparatively young ancestral forms may avoid being recruited.

From a more general viewpoint, the results reported here point to a strategy to the engineering of virus resistance. More specifically, the proof of concept we have provided here immediately suggests a reasonable approach to the engineering of virus-resistant plants. Such approach will involve the following steps:

(1) A known proviral factor in a plant is selected as a target. Obviously, this factor would be a protein that is hijacked (or suspected to be hijacked) by the virus or viruses for which we wish to engineer resistance.
(2) The known sequence of this protein is used as a search query in a sequence database.
(3) The sequences recovered from the search (belonging to homologs of the targeted protein) are aligned and the alignment is used as input for ancestral sequence reconstruction.
(4) The "modern" proviral factor is replaced by a reconstructed ancestral counterpart. Actually, ancestral sequences for many phylogenetic nodes can be derived from a single alignment of modern proteins and, therefore, the replacement could be actually performed with many

different ancestral proteins leading to many engineered plant variants.

(5) The engineered plant variants are screened for fitness under conditions of interest (normal growth conditions, for instance) and for virus resistance.

While this protocol to virus-resistant plants is obviously a proposal to be tested by future work, it is important to note that all the steps involved should be feasible. We elaborate on this in some detail below:

Viruses have a limited coding capacity (this is particularly true for the more common plant RNA viruses) and rely heavily on the recruitment of many host factors. Different approaches to identify such proviral factors have been developed in recent years (Wang, 2015; Sasvari and Nagy, 2016), and databases and search engines for virus-host interactions are currently available (e.g., http://virhostnet.prabi.fr; http://virusmentha.uniroma2.it). Clearly, therefore, the first step in the protocol (selecting a proviral host factor as target) is feasible, inasmuch as the 3D-structure of the targeted proviral factor or the mechanism of recruitment would not be required to be known.

The database search of the second step is also feasible because the very large (and exponentially growing) sequence databases available in the post-genomic era make it likely that a substantial number of homologous sequences are recovered when using the sequence of the targeted proviral factor as query. It is perhaps worth noting that the current version of the Uniprot database includes 38,327 entries from 2,001 species of Viridiplantae (http://www.uniprot.org/biocuration_project/Plants/statistics). Ancestral sequence reconstruction (step 3) is a well-known bioinformatics procedure and can be carried out using software that is widely available and that has been described in some detail in several recent reviews (for instance, see Table 1 in Gumulya and Gillam [2017]).

Recent advances in gene-editing methodologies using site-specific nucleases (Whitham and Hajimorad, 2016) make it feasible to perform the back-to-the-ancestral-form replacement of step 4. Possibly, the most direct approach at the current stage of development would be the CRISPR-Cas system with homologous recombination (HR) (Shan et al., 2014; Bortesi and Fischer, 2015; Puchta, 2016). HR involves a repair template that includes the DNA sequence to be inserted into the broken chromosome. In the context of our proposal, the repair template would include a reconstructed ancestral sequence. HR is considered as more challenging that the non-homologous end-joining (NHEJ) method used to generate knockouts. Still, targeted gene insertion using HR has been demonstrated in plants (for instance, see Shan et al. [2013] and Table 3 in Bortesi and Fischer [2015]). Furthermore, the HR methodology is likely to be extensively developed in the near future due to its many potential applications (Voytas and Gao, 2014).

Certainly, viruses have an enormous potential for rapid adaptation and can evolve to overcome natural or engineered resistance (Whitham and Hajimorad, 2016). Still, the problem of virus adaptation can be rationally addressed within the context of resistance engineering through ancestral replacement. First, the ''older'' ancestral proteins will likely have large numbers of mutational differences with their modern counterpart (the hijacked proviral factor), thus making it more difficult for the virus to adapt and recruit them. Certainly, the ''older'' ancestral proteins are more likely to display a limited functionality within the modern organism, thus compromising its fitness. However, ancestral sequences for many different phylogenetic nodes (of different age) can be derived from a single alignment of modern sequences. Plausibly, it should be possible to find ancestral proteins (perhaps of ''intermediate age'') displaying the right balance between limited effect on organismal fitness and resistance to recruitment despite virus adaptation. Second, ancestral replacement does not have to be limited to a single proviral factor. Replacement on multiple proviral factors will lead to a greater barrier to virus adaptation. As suggested by one of the anonymous reviewers of this work, maybe multiple, less divergent ancestral proteins could alleviate the negative impacts on fitness while representing an insurmountable barrier for viruses on the long term and not only on the short term.

Finally, we note that resurrected ancestral proteins typically display large sequence differences with all their modern counterparts. For instance, when the sequence of LPBCA thioredoxin is used as query, a Blast search on the non-redundant protein sequence database recovers a modern protein with only 70% identity with LPBCA-thioredoxin as the closest hit. The large sequence differences between ancestral and modern proteins suggest that back-to-the-ancestral-form replacement on a proviral factor could also provide a substantial barrier to cross-species transmission.

On a different note, this work shows that even laboratory representations of proteins near the origin of life may, at least in some cases, display some significant degree of functionality within a modern organism, despite the very large number of mutational changes between the modern and the ancestral proteins. This is a remarkable result, particularly in view of recent reports of strong negative correlation between estimated evolutionary age and the effect of the reconstructed protein on organismal fitness (Hobbs et al., 2015). However, our results point to an important role of the specificity of functional intermolecular interactions in determining the functionality of resurrected ancestral proteins within a modern organism. For proteins whose function involves molecular interactions that are not highly specific, we may expect a comparatively weak negative correlation between evolutionary age and the effect on organismal fitness. In any case, it is plausible that the fitness cost of replacing a modern protein with the laboratory resurrection of one of its ancestors could be compensated by similar modern-to-ancestor replacements on its interacting partners, thus providing an approach to study the evolution of relevant biological interactions. A few years before his death (in 2013), Emile Zuckerkandl (considered, together with Linus Pauling, as one of the founders of the Molecular Evolution field) wrote that ''the greatest interest of the reconstruction of ancestral informational macromolecules may well lie in the reconstruction of their interactions'' (Zuckerkandl, 2007). We demonstrate here that ancestral proteins may display functionality within a modern organism to an extent that is related to the specificity of their interactions, thus opening up the possibility of finally realizing Zuckerkandl's proposal.

## EXPERIMENTAL PROCEDURES

### Strains and Growth Conditions

Bacterial strains used in this study were *Escherichia coli* strains DHB3 (*araD139* Δ(*ara-leu*)*7697* Δ*lacX74 galE galK rpsL phoR* Δ(*phoA*)*Pvu*II Δ*malF3 thi*), DHB4 (DHB3 F′ [*lacpro lacI*$^{Q}$]) (Boyd et al., 1987), and FA41 (DHB4 *trxA trxC*). These strains were kindly donated by Jon Beckwith (Harvard Medical School). An F′ minus FA41 derivative susceptible of infection by T7 phage was obtained in this work (see below).

Cells were grown routinely on lysogeny broth (LB) medium (Bertani, 1951), except that glucose was omitted. Kanamycin 30 μg/mL was added for plasmid selection. Cultures were incubated on a rotary shaker at 180 rpm at 37°C unless otherwise indicated.

### Setting up the System to Study Fitness Effects of Thioredoxin Variants

The impacts on fitness of thioredoxin variants were studied using DHB4 and FA41, the two otherwise isogenic *E. coli* strains with or without the two thioredoxin genes identified in this organism, *trxA* and *trxC* (see above). FA41 showed a defect in growth when compared to its parental strain DHB4. This deficiency was readily apparent even at optimum growth conditions (Figure 2A). We took advantage of this phenotype to test the ability of thioredoxin variants to restore normal growth.

FA41 was adapted for expression of thioredoxin variants under a T7 promoter in pET derivative plasmids (Novagen) (see below). We were inclined to use this system instead of allelic replacement, because chromosomal expression of thioredoxin in *E. coli* is under the control of environmental factors included in the laboratory conditions used here. Expression of target genes from T7 promoters requires the presence of the specific RNA polymerase from the T7 phage. This gene was introduced in FA41 by lysogenization with λDE3 (Novagen), a modified λD69 phage that bears the T7 RNA polymerase gene under the control of the *lacUV5* promoter.

Lysogenization of FA41 was performed following manufacturer recommendations (λDE3 Lysogenization Kit, Novagen). This system is designed so that growing colonies are expected to be stable λDE3 lysogens. In effect, because λDE3 lacks the int function, it cannot integrate or be excised from the chromosome. Therefore, lysogenization of λDE3 requires coinfection with a Helper Phage providing it. In addition, a Selection Phage is also included in the infection event; this phage cannot infect λDE3 lysogens and kills non-λDE3 cells, ensuring that only lysogens survive. Specifically, the following procedure was carried out: one independent colony of the thioredoxin minus strain FA41 was grown at 37°C in LB medium supplemented with 10 mM $MgSO_4$ and 0.2% maltose (to induce expression of the lambda receptor). At an absorbance at 600 nm of 0.5, 1 μL of the culture was coinfected with $10^8$ PFU (plaque-forming units) of λDE3, $10^8$ PFU of Helper Phage, and $10^8$ PFU of Selection Phage, and the suspension was incubated for 20 min at 37°C to allow phage adsorption. Mixtures were plated on LB and incubated overnight at 37°C. Growing colonies were putative λDE3 lysogens. A number of these candidates were resistant to infection by Selection Phage, suggesting that they were in fact λDE3 lysogens. We randomly selected one of these colonies and further verified the presence of T7 RNA polymerase in it (see below).

Plasmids with thioredoxin variants where derivatives of pET30a(+)::*trxA* in which the open reading frame of thioredoxin 1 from *E. coli* followed by a stop codon was cloned as an *Nde*I-*Xho*I fragment under the control of a T7 promoter between the same sites in vector pET30a(+) (Novagen) (Godoy-Ruiz et al., 2005). The genes from human and ancestral thioredoxins were also cloned in pET30a(+) as previously described (Ingles-Prieto et al., 2013). Briefly, PCR-amplified open reading frames from the corresponding pEQ80L derivatives (Perez-Jimenez et al., 2011) were obtained using oligonucleotides that incorporated *Nde*I and *Xho*I restriction sites at their 5′ and 3′ ends, respectively. The PCR fragments where digested with *Nde*I and *Xho*I and cloned between the same sites in vector pET30a(+). Sequencing confirmed both constructions arrangements and the absence of mutations within the coding sequences of these genes.

FA41(DE3) cells were made chemically competent and transformed with the T7 derivative plasmids bearing thioredoxin variants, following the procedure by Nishimura et al. (1990).

### Validation of the System

We first analyzed the inducibility of thioredoxin gene expression under the control of the T7 promoter. The candidate FA41(DE3) was made competent and transformed with pET30a(+)::*trxA*. We incubated this clone in the presence of 1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside), and the corresponding cell extract was compared with that of non-induced cells in SDS-PAGE analyses. A protein band with the expected size of thioredoxin was obtained under inducing conditions (data not shown). This result indicated that the selected clone overexpressed thioredoxin and therefore contained the T7 RNA polymerase.

We also compared the growth of the thioredoxin-deficient strain complemented with thioredoxin with that of wild-type *E. coli*. We need to mention first that lysogenization with λDE3 had no effect on growth of FA41 (Figure 2A). Thioredoxin in pET30a(+)::*trxA* complemented the growth deficiency of FA41(DE3) even in the absence of an inducer. This presumably occurs due to the leakage of expression of T7 RNA polymerase from the *lacUV5* promoter (Studier and Moffatt, 1986). Therefore, basal expression of thioredoxin in cultures of FA41(DE3) (pET30a(+)::*trxA*) was sufficient to restore normal growth (Figure 2A).

To further verify that this complementation was also effective in all the culture conditions of this study, we measured generation times (see below) of DHB4 and FA41(DE3) with and without pET30a(+)::*trxA* at different temperatures. Doubling times of the complemented strain were similar to those of wild-type *E. coli* in all conditions (Figure 3). Therefore, we confirmed the validity of this system for the study of the in vivo performance of thioredoxin variants.

### Determination of Generation Times

The generation time, i.e., the time interval required for division of strains, was calculated by standard methods. Specifically, preinocula of the appropriate strains in LB medium were incubated overnight at 37°C and 180 rpm. Cultures were diluted 1/200 times in fresh medium and incubated at the desired temperature and 180 rpm. Samples were taken along the bacterial growth process, and absorbance at 600 nm ($A_{600}$) was measured as function of time. Generation time values were obtained from the slopes of plots of logarithm of absorbance at 600 nm versus time. Results are given as the average ± SD of four to six independent experiments.

### Determination of Efficiency of Plating by T7 Phage

#### *Curation of F′*

Growth of bacteriophage T7 in *E. coli* strains harboring F plasmids such as FA41 is abortive (reviewed in Molineux, 1991). In order to test the infectivity of T7 on *E. coli* with thioredoxin variants, we carried out the curation of F′ factor in FA41(DE3). F′ is spontaneously lost fairly frequently and clones that lack it can be easily obtained and isolated in solid media. Because F′ in FA41 carries the *lac* operon (see above), colonies without it cannot hydrolyze the chromogenic lactose analog X-gal (5-bromo-4-cholo-3-indolyl β-D-galactopyranoside) and therefore are white in a medium containing it. FA41(DE3) was streaked on LB plates with 50 μg/mL of X-gal, and white colonies were repurified several times to confirm clone homogeneity.

One well-isolated colony was selected and designed as FA41 F$^-$ (DE3). The Lac- phenotype of this clone could be due either to F′ loss or to recombination or deletion of the *lac* genes. To verify the curation of F′ in this clone, we tested its immunity to f2 phage (kindly provided by Marjorie Russell, The Rockefeller University), which infects only male strains (Loeb and Zinder, 1961). Contrary to FA41(DE3), strain FA41 F$^-$ (DE3) was not infected by f2, confirming that it was curated of the F′ factor. Finally, FA41 F$^-$ (DE3) supported the growth of T7 when transformed with pET30(a)::*trxA* (data not shown). Strain DHB3 (F$^-$ DHB4) was used as negative control strain in experiments for infection by f2 and as a positive control in those of infection by T7.

#### *Infection by T7 Phage*

*Escherichia coli* bacteriophage T7 (ATCC BAA-1025-B2) was used to test the efficacy of performance of human and ancestral thioredoxins in the T7 phage cycle. For that, we independently infected FA41 F$^-$ (DE3) transformed with the pET-derivative plasmids bearing *E.coli*, human, or ancestral thioredoxins. DHB3 was used as a positive control and untransformed FA41 F$^-$ (DE3) was the negative control. We carried out the bacterial infections using MOI that would render isolated and countable infection plaques. In the cases in which no plaques were obtained at this MOI, the phage load was progressively

increased up to an MOI that would still allow the formation of a lawn of growing cells.

Infection experiments were carried out as follows. Phage was serially diluted in dilution buffer (20 mM Tris-HCl pH 7.4, 100 mM NaCl, and 10 mM $MgSO_4$). Strains were grown in LB medium at 37°C to an approximate absorbance at 600 nm of 0.5. At that point, 100 μL of cells were mixed with 100 μL of the appropriate phage dilution, and the suspension was mixed with ~4 mL of molten top agar (Elbing and Brent, 2002). Plates were allowed to cool down for solidification and incubated overnight at 37°C. Plaque-forming units were then determined by visual inspection and counting. Control experiments showed that no plaques were obtained with *E. coli* $Trx^-$ strain, and the efficiency of propagation of T7 on *E. coli* $Trx^-$ with *E. coli* thioredoxin on a plasmid was similar to that of *E. coli* $Trx^+$ for any given phage suspension. Infection efficiency of *E. coli* $Trx^-$ with human and ancestral thioredoxins was calculated as the efficiency of propagation of each strain relative to that of *E. coli* $Trx^-$ with *E. coli* thioredoxin on a plasmid. Data given in Figure 4 are the average of several (at least six) independent determinations with the error bars corresponding to the SD.

The ability of ancestral and human thioredoxins to support growth of a highly mutated T7 population was tested using T7 Line 1 (kindly provided by James J. Bull, University of Texas at Austin). T7 line 1 was obtained by evolution of T7 phage growing in the susceptible *E. coli* strain IJ1133 in the presence of the mutagen N-methyl-N′-nitro-N-nitrosoguanidine. The T7 population consists of a collection of mutants obtained after ~200 generations on IJ1133 by serial transfer, with approximately four viable mutations per generation (Springman et al., 2010). Infections by T7 Line 1 were carried out following the same procedure as with wild-type T7 phage.

In the infection experiments in which expression of thioredoxins was induced by IPTG, cultures where incubated to an approximate absorbance at 600 nm of 0.15. At this point, 10 μM IPTG was added to one half of the culture and the rest was kept as un-induced control. Incubation proceeded for 2 additional hours to allow induction of thioredoxin expression over basal levels prior to phage infection. Infection experiments were conducted as detailed in the above paragraph with the exception that plates from induced samples where supplemented with 10 μM IPTG.

## AUTHOR CONTRIBUTIONS

A.D. designed and performed experiments. R.A. and B.I.-M. contributed methodological tools. J.M.S.-R. designed research. A.D. and J.M.S.-R. wrote the manuscript. All authors discussed the results and the manuscript.

## ACKNOWLEDGMENTS

This work was supported by grants BIO2012-34937, CSD2009-00088, and BIO2015-66426-R (J.M.S.-R.) from MINECO/FEFER and grant P09-CVI-5073 (B.I.-M.) from the "Junta de Andalucía" and Feder Funds. The authors express their gratitude to Dr. Jon Beckwith for donating strains *E. coli* DHB3, DHB4, and FA41; Dr. Marjorie Russel for the f2 bacteriophage; and Dr. James J. Bull for the T7 Line 1. The technical assistance of $M^a$ Carmen Tejero is gratefully acknowledged. We also thank Dr. Lynn Kamerlin, Dr. Peter Kasson, and Dr. Valeria A. Risso for useful comments on the manuscript.

Received: October 10, 2016
Revised: February 14, 2017
Accepted: April 13, 2017
Published: May 9, 2017

## REFERENCES

Akabayov, B., Akabayov, S.R., Lee, S.J., Tabor, S., Kulczyk, A.W., and Richardson, C.C. (2010). Conformational dynamics of bacteriophage T7 DNA polymerase and its processivity factor, *Escherichia coli* thioredoxin. Proc. Natl. Acad. Sci. USA *107*, 15033–15038.

Benner, S.A., Sassi, S.O., and Gaucher, E.A. (2007). Molecular paleoscience: systems biology from the past. Adv. Enzymol. Relat. Areas Mol. Biol. *75*, 1–132, xi.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. J. Bacteriol. *62*, 293–300.

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Adv. *33*, 41–52.

Boyd, D., Manoil, C., and Beckwith, J. (1987). Determinants of membrane protein topology. Proc. Natl. Acad. Sci. USA *84*, 8525–8529.

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., Sherman, A., Arazi, T., and Gal-On, A. (2016). Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. Mol. Plant Pathol. *17*, 1140–1153.

Collet, J.F., and Messens, J. (2010). Structure, function, and mechanism of thioredoxin proteins. Antioxid. Redox. Signal. *13*, 1205–1215.

Collinge, D.B., Jørgensen, H.J., Lund, O.S., and Lyngkjaer, M.F. (2010). Engineering pathogen resistance in crop plants: current trends and future prospects. Annu. Rev. Phytopathol. *48*, 269–291.

Dasgupta, I., Malathi, V.G., and Mukherjee, S.K. (2003). Genetic engineering for virus resistance. Curr. Sci. *84*, 341–354.

Daugherty, M.D., and Malik, H.S. (2012). Rules of engagement: molecular insights from host-virus arms races. Annu. Rev. Genet. *46*, 677–700.

Demogines, A., Abraham, J., Choe, H., Farzan, M., and Sawyer, S.L. (2013). Dual host-virus arms races shape an essential housekeeping protein. PLoS Biol. *11*, e1001571.

Doublié, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature *391*, 251–258.

El Hajjaji, H., Dumoulin, M., Matagne, A., Colau, D., Roos, G., Messens, J., and Collet, J.F. (2009). The zinc center influences the redox and thermodynamic properties of Escherichia coli thioredoxin 2. J. Mol. Biol. *386*, 60–71.

Elbing, K., and Brent, R. (2002). Media preparation and bacteriological tools. Curr. Protoc. Mol. Biol. *Chapter 1*, Unit 1.1.

Enard, D., Cai, L., Gwennap, C., and Petrov, D.A. (2016). Viruses are a dominant driver of protein adaptation in mammals. eLife *5*, e12469.

Etson, C.M., Hamdan, S.M., Richardson, C.C., and van Oijen, A.M. (2010). Thioredoxin suppresses microscopic hopping of T7 DNA polymerase on duplex DNA. Proc. Natl. Acad. Sci. USA *107*, 1900–1905.

Friedel, C.C., and Haas, J. (2011). Virus-host interactomes and global models of virus-infected cells. Trends Microbiol. *19*, 501–508.

Galvez, L.C., Banerjee, J., Pinar, H., and Mitra, A. (2014). Engineered plant virus resistance. Plant Sci. *228*, 11–25.

Gaucher, E.A., Govindarajan, S., and Ganesh, O.K. (2008). Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature *451*, 704–707.

Godoy-Ruiz, R., Perez-Jimenez, R., Ibarra-Molero, B., and Sanchez-Ruiz, J.M. (2005). A stability pattern of protein hydrophobic mutations that reflects evolutionary structural optimization. Biophys. J. *89*, 3320–3331.

Gumulya, Y., and Gillam, E.M.J. (2017). Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. Biochem. J. *474*, 1–19.

Hamdan, S.M., and Richardson, C.C. (2009). Motors, switches, and contacts in the replisome. Annu. Rev. Biochem. *78*, 205–243.

Harms, M.J., and Thornton, J.W. (2010). Analyzing protein structure and function using ancestral gene reconstruction. Curr. Opin. Struct. Biol. *20*, 360–366.

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

Hobbs, J.K., Prentice, E.J., Groussin, M., and Arcus, V.L. (2015). Reconstructed ancestral enzymes impose a fitness cost upon modern bacteria despite exhibiting favourable biochemical properties. J. Mol. Evol. *81*, 110–120.

Holmgren, A. (1995). Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. Structure *3*, 239–243.

Holmgren, A., Söderberg, B.O., Eklund, H., and Brändén, C.I. (1975). Three-dimensional structure of *Escherichia coli* thioredoxin-$S_2$ to 2.8 A resolution. Proc. Natl. Acad. Sci. USA *72*, 2305–2309.

Holmgren, A., Ohlsson, I., and Grankvist, M.L. (1978). Thiroedoxin from Escherichia coli. Radioimmunological and enzymatic determinations in wild type cells and mutants defective in phage T7 DNA replication. J. Biol. Chem. *253*, 430–436.

Ingles-Prieto, A., Ibarra-Molero, B., Delgado-Delgado, A., Perez-Jimenez, R., Fernandez, J.M., Gaucher, E.A., Sanchez-Ruiz, J.M., and Gavira, J.A. (2013). Conservation of protein structure over four billion years. Structure *21*, 1690–1697.

Jacob, C., Kriznik, A., Boschi-Muller, S., and Branlant, G. (2011). Thioredoxin 2 from Escherichia coli is not involved in vivo in the recycling process of methionine sulfoxide reductase activities. FEBS Lett. *585*, 1905–1909.

Kirschner, M.W., and Gerhart, J.C. (2005). The Plausibility of Life: Resolving Darwin's Dilemma (New Haven, CT: Yale University Press).

Knauth, L.P. (2005). Temperature and salinity history of the Precambrian ocean: implications for the course of microbial evolution. Palaeogeogr. Palaeoclimatol. Palaeoecol. *219*, 53–69.

Kumar, J.K., Tabor, S., and Richardson, C.C. (2004). Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli. Proc. Natl. Acad. Sci. USA *101*, 3759–3764.

Lee, S.J., and Richardson, C.C. (2011). Choreography of bacteriophage T7 DNA replication. Curr. Opin. Chem. Biol. *15*, 580–586.

Liberles, D. (2007). Ancestral Sequence Reconstruction (New York, NY: Oxford University Press).

Lim, C.J., Geraghty, D., and Fuchs, J.A. (1985). Cloning and nucleotide sequence of the trxA gene of Escherichia coli K-12. J. Bacteriol. *163*, 311–316.

Loeb, T., and Zinder, N.D. (1961). A bacteriophage containing RNA. Proc. Natl. Acad. Sci. USA *47*, 282–289.

Merkl, R., and Sterner, R. (2016). Ancestral protein reconstruction: techniques and applications. Biol. Chem. *397*, 1–21.

Miranda-Vizuete, A., Damdimopoulos, A.E., Gustafsson, J., and Spyrou, G. (1997). Cloning, expression, and characterization of a novel *Escherichia coli* thioredoxin. J. Biol. Chem. *272*, 30841–30847.

Molineux, I.J. (1991). Host-parasite interactions: recent developments in the genetics of abortive phage infections. New Biol. *3*, 230–236.

Ng, M., Ndungo, E., Kaczmarek, M.E., Herbert, A.S., Binger, T., Kuehne, A.I., Jangra, R.K., Hawkins, J.A., Gifford, R.J., Biswas, R., et al. (2015). Filovirus receptor NPC1 contributes to species-specific patterns of ebolavirus susceptibility in bats. eLife *4*, e11785.

Nishimura, A., Morita, M., Nishimura, Y., and Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent Escherichia coli cells. Nucleic Acids Res. *18*, 6169.

Pauling, L., and Zuckerkandl, E. (1963). Chemical paleogenetics. Molecular "restoration studies" of extinct forms of life. Acta Chem. Scand. *17*, S9–S16.

Perez-Jimenez, R., Inglés-Prieto, A., Zhao, Z.M., Sanchez-Romero, I., Alegre-Cebollada, J., Kosuri, P., Garcia-Manyes, S., Kappock, T.J., Tanokura, M., Holmgren, A., et al. (2011). Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. Nat. Struct. Mol. Biol. *18*, 592–596.

Prins, M., Laimer, M., Noris, E., Schubert, J., Wassenegger, M., and Tepfer, M. (2008). Strategies for antiviral resistance in transgenic plants. Mol. Plant Pathol. *9*, 73–83.

Puchta, H. (2016). Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. Curr. Opin. Plant Biol. *36*, 1–8.

Pyott, D.E., Sheehan, E., and Molnar, A. (2016). Engineering of CRISPR/Cas9-mediated potyvirus resistance in transgene-free Arabidopsis plants. Mol. Plant Pathol. *17*, 1276–1288.

Risso, V.A., Gavira, J.A., and Sanchez-Ruiz, J.M. (2014). Thermostable and promiscuous Precambrian proteins. Environ. Microbiol. *16*, 1485–1489.

Robert, F., and Chaussidon, M. (2006). A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. Nature *443*, 969–972.

Romero-Romero, M.L., Risso, V.A., Martinez-Rodriguez, S., Gaucher, E.A., Ibarra-Molero, B., and Sanchez-Ruiz, J.M. (2016). Selection for protein kinetic stability connects denaturation temperatures to organismal temperatures and provides clues to Archaean life. PLoS ONE *11*, e0156657.

Rybicki, E.P. (2015). A Top Ten list for economically important plant viruses. Arch. Virol. *160*, 17–20.

Sasvari, Z., and Nagy, P.D. (2016). Exploration of plant virus replication inside a surrogate host, *Saccharomyces cerevisiae*, elucidates complex and conserved mechanisms. In Current Research Topics in Plant Virology, A. Wang and X. Zhou, eds. (Cham, Switzerland: Springer International Publishing), pp. 35–65.

Savary, S., Ficke, A., Aubertot, J., and Hollier, C. (2012). Crop losses due to diseases and their implications for global food production and food security. Food Secur. *4*, 519–537.

Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L., and Gao, C. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. *31*, 686–688.

Shan, Q., Wang, Y., Li, J., and Gao, C. (2014). Genome editing in rice and wheat using the CRISPR/Cas system. Nat. Protoc. *9*, 2395–2410.

Springman, R., Keller, T., Molineux, I.J., and Bull, J.J. (2010). Evolution at a high imposed mutation rate: adaptation obscures the load in phage T7. Genetics *184*, 221–232.

Strange, R.N., and Scott, P.R. (2005). Plant disease: a threat to global food security. Annu. Rev. Phytopatol. *43*, 3.1–3.34.

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. *189*, 113–130.

Voytas, D.F., and Gao, C. (2014). Precision genome engineering and agriculture: opportunities and regulatory challenges. PLoS Biol. *12*, e1001877.

Wang, A. (2015). Dissecting the molecular network of virus-plant interactions: the complex roles of host factors. Annu. Rev. Phytopathol. *53*, 45–66.

Whitham, S.A., and Hajimorad, M.R. (2016). Plant genetic resistance to viruses. In Current Research Topics in Plant Virology, A. Wang and X. Zhou, eds. (Cham, Switzerland: Springer International Publishing), pp. 87–111.

Wilke, C.O., and Sawyer, S.L. (2016). At the mercy of viruses. eLife *5*, e16758.

Zuckerkandl, E. (2007). Introduction to the meeting in Kristineberg, Sweden. In Ancestral Sequence Reconstruction, D.A. Liberles, ed. (New York, NY: Oxford University Press).