Review

The Origin and Evolution of Release Factors: Implications for Translation Termination, Ribosome Rescue, and Quality Control Pathways

A. Maxwell Burroughs and **L Aravind** *

Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA; burrough@mail.nih.gov
* Correspondence: aravind@mail.nih.gov; Tel.: +1-301-496-2475

check for updates

Abstract: The evolution of release factors catalyzing the hydrolysis of the final peptidyl-tRNA bond and the release of the polypeptide from the ribosome has been a longstanding paradox. While the components of the translation apparatus are generally well-conserved across extant life, structurally unrelated release factor peptidyl hydrolases (RF-PHs) emerged in the stems of the bacterial and archaeo-eukaryotic lineages. We analyze the diversification of RF-PH domains within the broader evolutionary framework of the translation apparatus. Thus, we reconstruct the possible state of translation termination in the Last Universal Common Ancestor with possible tRNA-like terminators. Further, evolutionary trajectories of the several auxiliary release factors in ribosome quality control (RQC) and rescue pathways point to multiple independent solutions to this problem and frequent transfers between superkingdoms including the recently characterized ArfT, which is more widely distributed across life than previously appreciated. The eukaryotic RQC system was pieced together from components with disparate provenance, which include the long-sought-after Vms1/ANKZF1 RF-PH of bacterial origin. We also uncover an under-appreciated evolutionary driver of innovation in rescue pathways: effectors deployed in biological conflicts that target the ribosome. At least three rescue pathways (centered on the prfH/RFH, baeRF-1, and C12orf65 RF-PH domains), were likely innovated in response to such conflicts.

Keywords: release factors; translation termination; peptidyl hydrolase

1. Introduction

The translation apparatus includes the most conserved cellular proteins and RNAs that provide the strongest evidence for the shared ancestry of all life forms (Figure 1). Along with the highly conserved ribosomal proteins and ribosomal RNAs, the tRNAs and their cognate amino-acyl tRNA synthetases provide the amino acid building blocks for peptide formation (Figure 1). Additionally, the actual ribosomal decoding of mRNA requires conserved protein factors observed at each of the core steps of translation: initiation, elongation, and release (Figure 1). These factors are primarily GTPases, which act as proofreading switches to ensure accurate decoding of the genetic code. Several of these GTPases display strong evidence for an ancient origin prior the Last Universal Common Ancestor (LUCA) [1].

Notably, prior to the LUCA, the GTPases involved in translation had already differentiated into several distinct clades that were since retained through the history of life on earth (Figure 1) [1]. The first of these clades is represented by EF-Tu/EF-1α, which binds the amino acylated (AA)-tRNA, protects the amino acyl group from hydrolysis, and ensures correct matching of the codon with the anticodon [2,3]. The second clade is prototyped by EF-G/EF2, which plays a role in the translocation of

the tRNA as the ribosome elongates the nascent peptide [4,5]. The GTPase IF2/eIF5B defines another clade predating the LUCA, which forms a higher order clade with EF-G/EF-2 (Figure 1). However, in this case it appears to have undergone a functional divergence after the split of the bacterial and archaeo-eukaryotic lineages from the LUCA. In bacteria, the primary member of this clade is critical for the recruitment of the Met-tRNAifMet to the ribosome [6], whereas the eukaryotic cognate plays a central role in the bringing together of the two ribosomal subunits [7]. In contrast, in eukaryotes eIF-2γ, which emerged as an offshoot of the EF-Tu/EF-1α clade, it plays the role of recruiting the Met-tRNAiMet to the ribosome to form the pre-initiation complex [8].
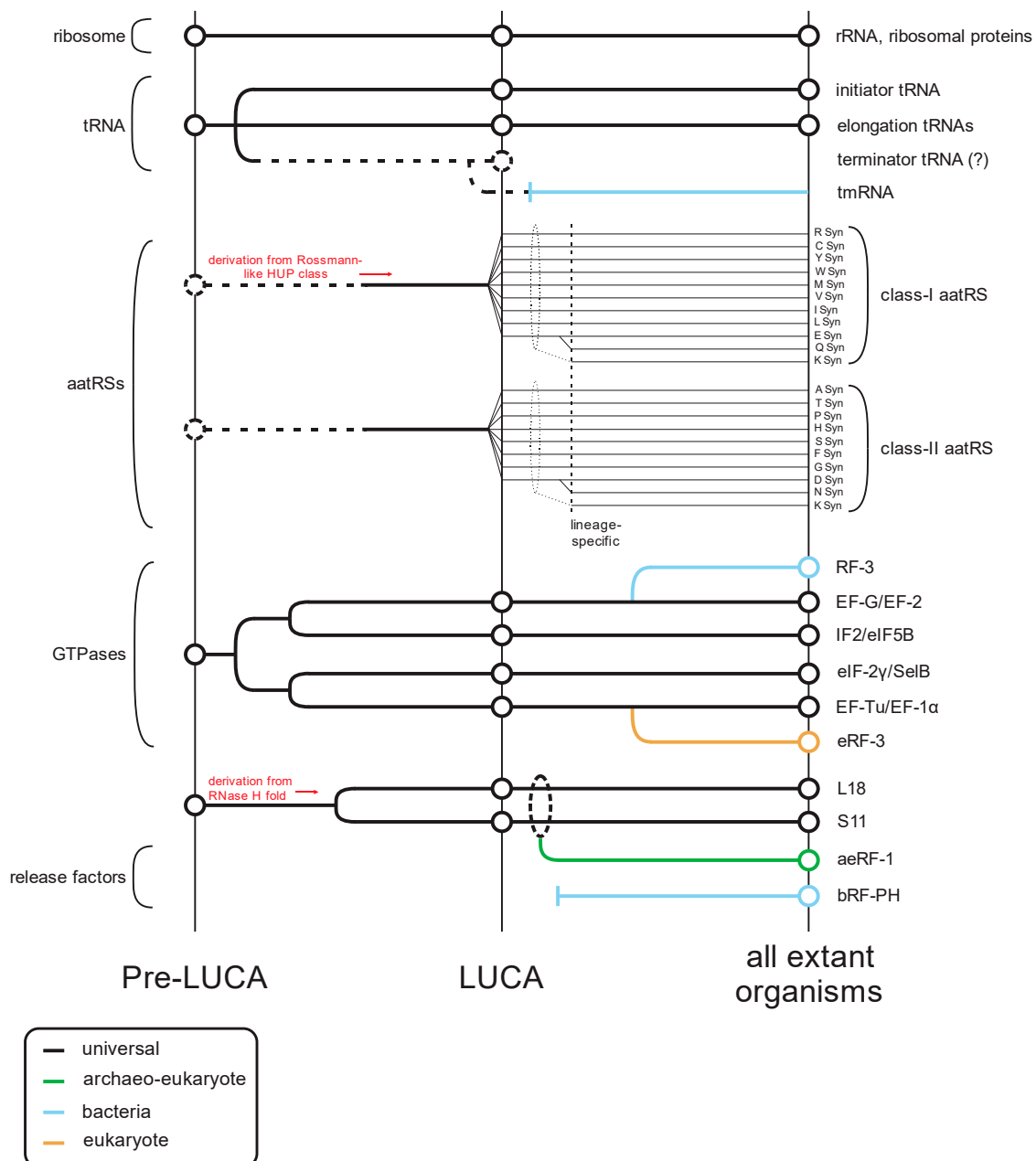

Figure 1. Overview of the evolutionary history of the translation apparatus. Key classes of protein and RNA contributors (general protein/RNA class labeled to the **left**, specific component names provided to the **right**) in translation are traced across several evolutionary epochs (labeled at the **bottom**) in this temporal diagram. Lines track evolutionary depth for components and are color-coded by phyletic distribution. Key derivation events are noted in red. Events where evolutionary depth are unclear are denoted with broken lines.

The situation with the translation release factors is more complex as there is no release factor GTPase (RF-GTPase) that can be confidently traced back to the LUCA. Rather, there are two distinct RF-GTPases (also known as the Type-II release factors): bRF-3 which emerged as an offshoot of the EF-G/EF-2 clade in bacteria; and eRF-3 which emerged as an offshoot of the EF-Tu/EF-1α clade, only in eukaryotes [1] (Figure 1). Archaea possess no such dedicated RF-GTPase and appear to use EF-1α itself [9,10]. Despite their distinct evolutionary origins, the RF-GTPases of both bacteria and eukaryotes perform a comparable biochemical function; whereas the elongation factors interact with tRNA, the RF-GTPases interact with the second release factor subunit, a peptidyl hydrolase (RF-PH, the Type-I release factors), which is believed to mimic the tRNA [11,12]. Upon recognizing a stop codon, this RF-PH protein is loaded on to the A-site of the ribosome in place of an AA-tRNA by the action of the RF-GTPase, analogous to the role played by EF-Tu/EF-1α in loading the AA-tRNA during elongation [11]. A conserved glutamine in the RF-PH protein directs a water molecule in place of the usual acceptor amino group in the A-site [13–15]. Transfer of the peptidyl chain from the tRNA at the P-site to this water molecule by the action of the peptidyl transferase catalytic center in the ribosome large subunit rRNA results in hydrolysis of the peptidyl-tRNA bond and release of the completed polypeptide [16–18].

Sequence and structure analyses of the RF-PH subunits of the release factors show an evolutionary dichotomy, which is like that of the RF-GTPases (Figure 1). Despite having a similar catalytic mechanism and a conserved active site glutamine (Figure 2A,B), the RF-PHs of the bacteria (RF-1 and RF-2) and the archaeo-eukaryotic lineage (aeRF-1) are evolutionarily unrelated (Figure 2A–D) [19,20]. In this context, it is also worth noting a comparable dichotomy in the primary ribosome recycling factors that initiate splitting of the large and small subunits after release: In the archaeo-eukaryotic lineage, this process is initiated by the ABC ATPase Rli1/ABCE1 in conjunction with aeRF-1 [21,22]. In contrast, the bacteria initiate recycling via the evolutionarily unrelated ribosome recycling factor (RRF), which instead contains a conserved domain shared with several amino-acyl tRNA synthetases [23].

These phyletic patterns of the RF-PHs, the GTPases that load them on to the ribosome, and the recycling factors, are atypical for components of the core apparatus that perform essential functions in translation. Hence, this suggests that there are key aspects of the early evolution of the translation apparatus that remain poorly understood. First, it remains unclear as to what was the mechanism of peptide release in the LUCA. Second, even though translation termination is an essential and universal process across the translation systems of extant life forms, the lack of universal conservation of its components, unlike those involved in elongation (Figure 1), raises questions about the importance of this step in the ancestral translation systems. Finally, preliminary surveys have indicated that the RFs may show a much greater tendency for paralog formation and concomitant diversification along with trans-superkingdom lateral transfers. This indicates that multiple, distinct evolutionary forces might be operating on these molecules. In this article we systematically survey the RF-PHs and bring to light several hitherto unnoticed or poorly appreciated aspects of their natural history. We synthesize this information to provide plausible explanations and hypotheses relating to the above questions.

2. Results and Discussion

2.1. The Prior State of Understanding of the Evolution of Peptidyl Hydrolase Release Factors

The general outlines of the evolutionary history of both the archaeo-eukaryotic and bacterial RF-PH (bRF-PH) have been described as follows: the catalytic domain of the aeRF-1s displays a RNase H fold domain (Figure 2A,C) whereas that of the bRF-PHs shows a distinct α + β fold with no relationship to the former (Figure 2B,D). However, both these types of RF-PHs have further convergently evolved to a three-domain architecture (Figure 2E,F). In the archaeo-eukaryotic lineage the central catalytic aeRF-1 domain is sandwiched between the N-terminal stop-codon-recognition domain of the archaeo-eukaryotic lineage and the C-terminal Pelota domain (which often contains a further zinc-ribbon domain inserted into it [24–26] (Figure 2E), a widely-distributed domain which binds

specific RNA secondary structures [27,28]. In contrast, the bRF-PH catalytic domain is inserted into the core of the bacterial stop-codon-recognition domain [20]. Together, this combined module is fused to the N-terminal α-helical GTPase-interacting domain [29,30] (Figure 2F). The convergence to a comparable three-domain architecture in the RF-PH subunits of both the bacterial and archaeo-eukaryotic lineages has been proposed to have resulted from the selective pressure to adopt structures mimicking the tRNA [24,31,32] (Figure 2E,F).
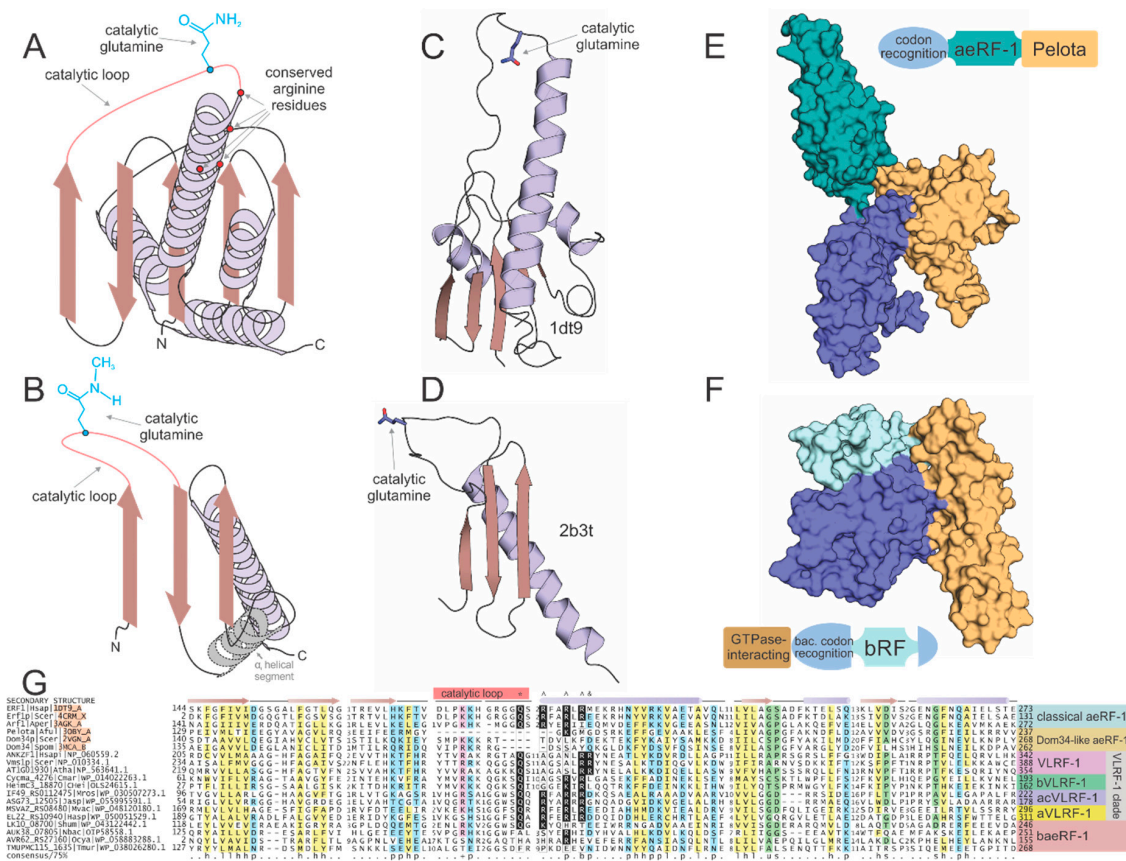

Figure 2. Structural and sequence overview of the catalytic domains of the two types of RF-PHs. (**A**,**B**) Topology diagrams of (**A**) aeRF-1 superfamily and (**B**) bRF-PH domain. (**C**,**D**) Cartoon rendering of solved crystal structures of known examples of (**C**) aeRF-1 superfamily and (**C**) bRF-PH domain. Protein DataBank id (PDBID) provided to right of structure. Coloring as in (**A**,**B**). Catalytic glutamine residue rendered as ball-and-stick. (**E**,**F**) Space-filled rendering of (**C**,**D**), while including full complement of domains found in the RF-PHs; positioning altered from (**C**,**D**) to emphasize proposed tRNA mimicry of the tri-domain release factors. Domain architectures provided adjacent to space-filled structural renderings, color-coded with structure to demarcate domains. (**G**) Multiple sequence alignment of aeRF-1 superfamily. Sequences are labeled to the left of the alignment and secondary structure is provided above the alignment. The catalytic loop region is marked with a pink box along secondary structure line and the catalytic glutamine residue position is marked with an asterisk above. Conserved positively-charged positions are marked with '^'. Conserved arginine position present only in eVLRF-1 family is marked with ampersand. Consensus line provided below alignment, coloring and abbreviations as follows: p, polar residue shaded in blue; '+', positively-charged residue shaded in purple; h, hydrophobic residue shaded in yellow; l, aliphatic residue shaded in yellow; u, tiny residue shaded in green; s, small residue shaded in green; b, big residue.

On the archaeo-eukaryotic side, the aeRF-1 RF-PH domain underwent a duplication early in the common ancestor of the archaeo-eukaryotic lineage, diversifying into two distinct paralogous versions (Figure 3A). The first of these, the bona fide aeRF-1 (hereinafter classical aeRF-1), serves as the peptidyl

hydrolase for polypeptide release during regular translation. The regular aeRF-1 recognizes all three stop codons via a single anti-codon binding domain [24] (Figure 3A,B). The second paralog Dom34 plays a specialized role in recycling a ribosome stalled due to impassable secondary structure in the mRNA (no-go) or a defective mRNA lacking a stop codon due to mutation or premature truncation, which can be caused by endogenous processes or introduced by the extrinsic action of RNase effectors (non-stop) [33–37]. Consistent with this function, Dom34 lacks the conserved glutamine residue and is thereby catalytically inactive. Further, in lieu of the N-terminal stop-codon-recognition domain, Dom34 has a distinct β-barrel domain belonging to the superfamily of Sm-like RNA-binding domains [38] (Figure 3C). This N-terminal domain was at one point incorrectly identified as a nuclease, however, subsequent studies have disabused this notion and indicated that it acts in binding the mRNA [33,39]. In eukaryotes, Dom34 is loaded to the A-site by the GTPase Hbs1, a paralog of the release factor GTPase eRF-3 [36,40].
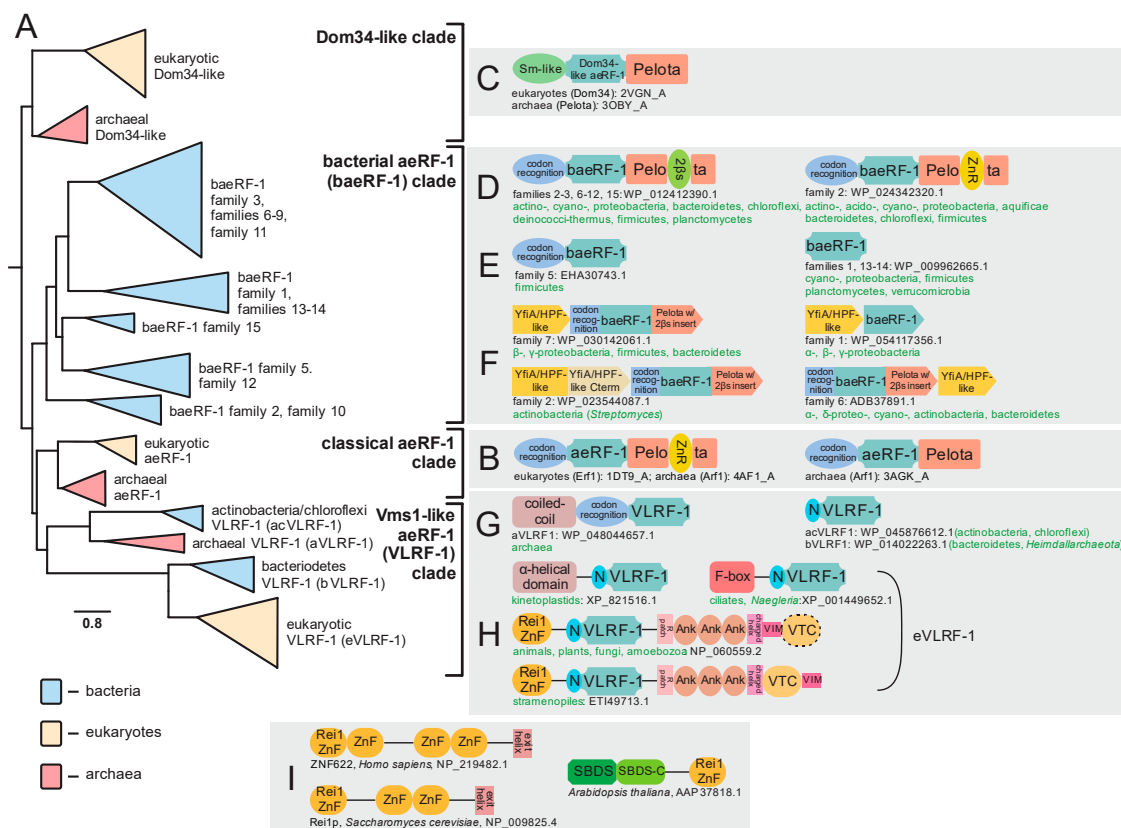

Figure 3. Evolution of aeRF-1 superfamily with known domain architectures and conserved gene neighborhoods. (**A**) Collapsed phylogenetic tree depicting relationships within the aeRF-1. (**B–I**) Notable domain architectures and conserved gene neighborhoods grouped according to clades. Individual domains of any given single protein architecture are depicted as discrete, colored shapes. In (**F**), individual genes in conserved gene neighborhoods are depicted as boxed arrows with the arrowhead pointing in the 3′ direction. Genes with multiple domains have their boxes split into distinct colors, one color for each discrete domain. Domain abbreviations: VTC, Vms1 Treble Clef; ZnR, zinc ribbon; 2βs, 2-β-strand insert domain; SBDS, Shwachman-Bodian-Diamond syndrome; N, distinctive N-terminal extension shared between certain families of the VLRF-1 clade.

Bacteria possess two paralogous RF-PHs, RF1 and RF2, which are nearly absolutely conserved across bacteria and are responsible for recognizing distinct sets of stop codons [41] (Figure 4A). Additional paralogs of these conserved versions with more specialized roles appear to have emerged subsequently in bacterial evolution. However, their evolutionary trajectories remain poorly understood

beyond the observation that eukaryotes also likely acquired this domain early in their evolution [42] (Figure 4).
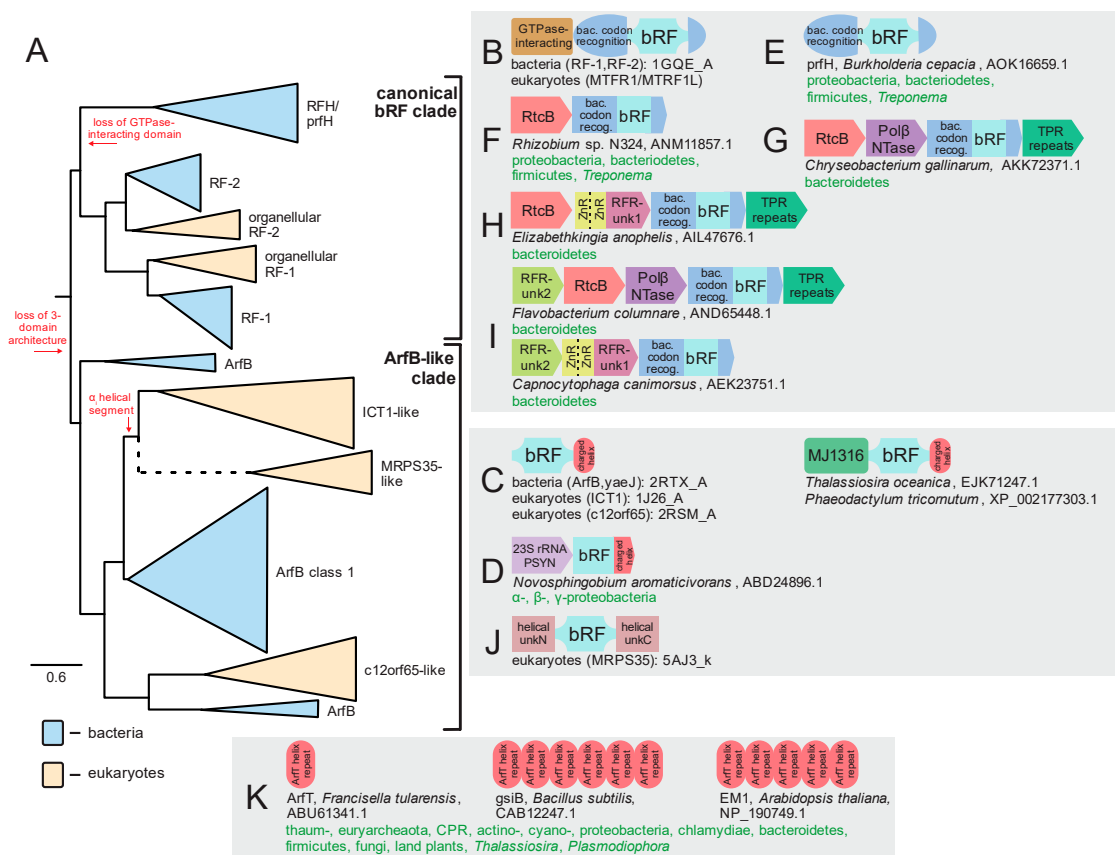

Figure 4. Evolution of bRF-PH domain with known domain architectures and conserved gene neighborhoods. (**A**) Collapsed phylogenetic tree depicting the radiation of the bRF-PH domains. Key transitions are denoted in red lettering along the tree. (**B–K**) Notable domain architectures and conserved gene neighborhoods. See Figure 3 legend for further descriptions and details. Abbreviations: PSYN, pseudouridine synthase; RFR, RFH and RtcB-associating unknown domains; bac. codon recog., bacterial codon recognition domain.

While the above provides an overview of the evolutionary history for RF-PH domains developed over the past three decades, recent investigations have considerably expanded the repertoire of known families in unexpected directions. In the following sections we present these findings, placing an emphasis on their evolutionary and functional implications. Additionally, these findings provide new data with a bearing on two long-standing biochemical and evolutionary problems. First, discovery of novel aeRF-1 families increases resolution of the total evolutionary history of the RF-PHs and provides an insight into translation termination events outside of routine translation and the specialized termination events mediated by Dom34. Second, these findings contribute to the continuing elucidation of distinct ribosome quality control pathways, in particular, with relevance for understanding the evolutionary origins of the eukaryotic ribosomal quality control (RQC) system. In this regard, we present a general evolutionary framework for understanding how RQC emerged in eukaryotes and what this implies for potential parallel innovations in the other superkingdoms of Life, a summary of which is provided in Figure 5.
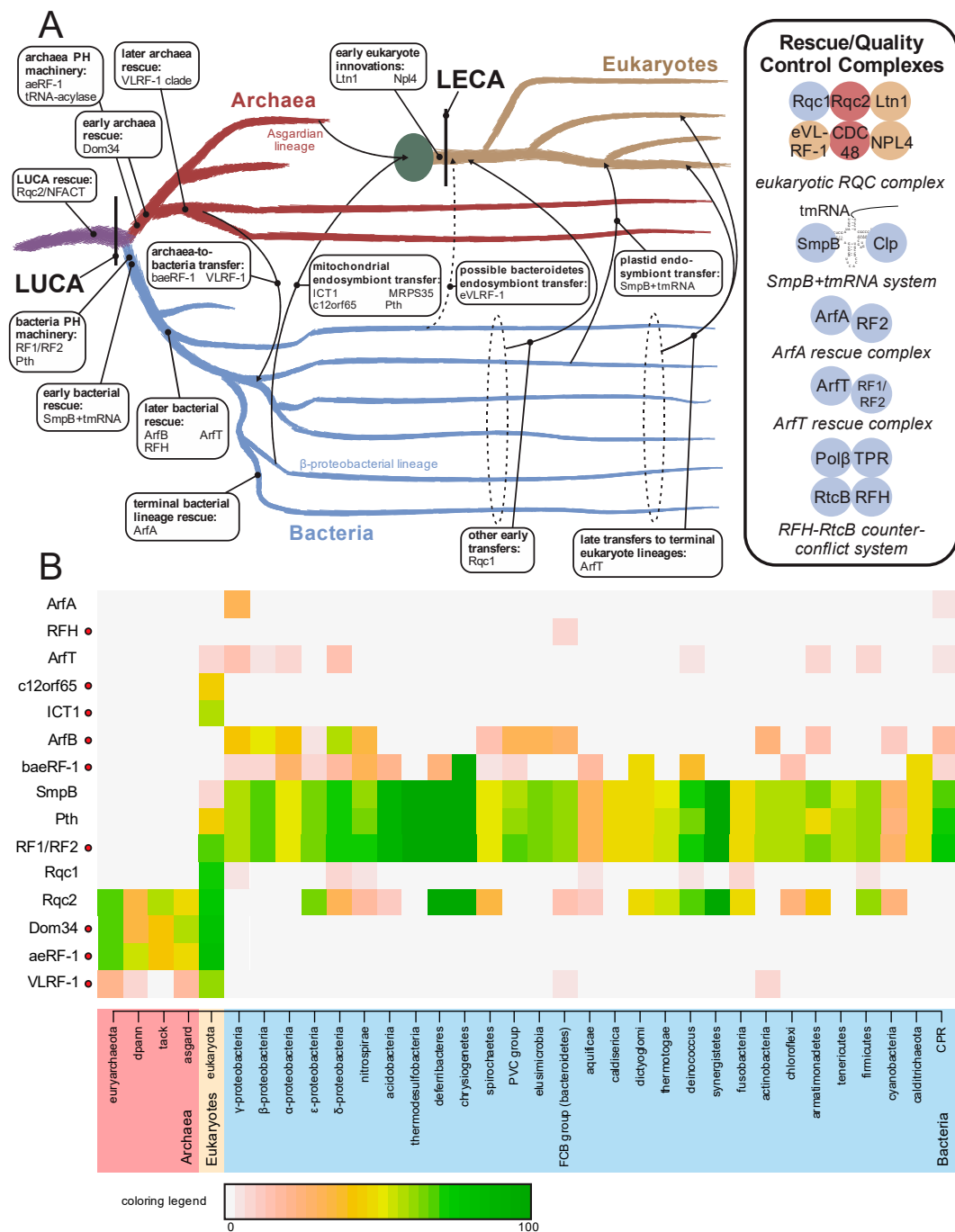
Figure 5. (**A**) Evolutionary origins and trajectories of release factors and ribosome rescue pathways, with factors listed at points of emergence/transfer. Transfer events are depicted with arrows and labeled; when timing is unclear the lines are broken. The mitochondrial endosymbiotic event is represented by a large green circle. Complexes formed by release/rescue factors are depicted to the right. Circles are color-coordinated with tree to represent the superkingdom of origin of a component. (**B**) Relative presence/absence for individual components of the ribosome quality control and rescue pathways. Presence/absence is colored by percentage of genomes within taxa which contain a given component. Names of taxa are colored by superkingdom. Components containing a RF-PH domain are marked to the right of their name by a red circle. Abbreviations: tack, Thaumarchaeota, Crenarchaeota, and Korarchaeota superphylum; dpann, Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaea superphylum; PVC group, Chlamydiae, Lentisphaerae, Planctomycetes, and Verrucomicrobia bacterial superphylum; FCB, Fibrobacteres, Bacteroidetes, and Chlorobi superphylum; CPR, Candidate Phyla Radiation.

2.2. Novel Members of the aeRF-1 Superfamily: Their Domain Architectures and Conserved Gene Neighborhood Associations

In addition to the classical aeRF-1 and Dom34 families, two novel clades of the superfamily encompassing at least 19 distinct families were identified through the application of sensitive sequence analysis methods (for complete list, see Supplementary Data). A multiple alignment of representatives of the known and newly identified clades provides an overview of their shared and distinct structure and sequence features; most notably, it allowed us to determine whether they are likely to be catalytically active as in the aeRF-1 family or inactive as in the Dom34 family (Figure 2G). Many of these families also displayed distinct domain architectures with no previous antecedents among the RF-PHs, expanding their architectural diversity beyond previously described three-domain versions (Figure 2E,F and Figure 3B,C). This multiple sequence alignment was further used to perform a phylogenetic analysis of the aeRF-1 superfamily and produce an updated natural classification of the aeRF-1 proteins (Figure 3A).

Even with the discovery of the novel families, these analyses continue to support a fundamental split in the aeRF-1 superfamily between the Dom34-like and the aeRF-1-like clade, which predated the common ancestor of the archaeo-eukaryotic lineage, as conserved representatives of both clades are found in all archaeal and eukaryotic lineages (Figure 3A). The Dom34-like clade is comprised of only the Dom34 proteins of eukaryotes and archaea, several representatives from the Asgardarchaeal lineage group closer to eukaryotic versions than to other archaeal versions, which is consistent with the recent proposal of the derivation of eukaryotes within this archaeal lineage [43].

The two newly discovered clades of aeRF-1 are more closely related to the classical aeRF-1 clade than the Dom34 clade (Figure 2B). Of these, the so-called bacterial aeRF-1 clade (baeRF-1) largely retains the standard domain architecture of the classical aeRF-1, although in many families of this clade, a distinctive insert comprised of two β-strands instead of the zinc ribbon (ZnR) domain is inserted within the Pelota domain. Given that it observed at the same insertion site as the ZnR, it is conceivable that it represents a degenerate ZnR (Figure 3D). In one distinctive baeRF-1 family, the N-terminal stop-codon-binding domain is retained, however, the C-terminal Pelota domain is lost. Three baeRF-1 families have lost both flanking domains, retaining only the core RNase H fold domain (Figure 2A,C and Figure 3E). Across representatives of several baeRF-1 families, a conserved gene neighborhood association with the ribosome hibernation factor (HPF), involved in translational dormancy, is observed suggesting a potential physical and functional association between these proteins [44,45] (Figure 3F, Supplementary Data). Phylogenetic analyses suggest that the baeRF-1 clade emerged via a single early lateral transfer from archaea to bacteria (Figure 5). This was followed by their diversification and extensive lateral transfer within the bacterial superkingdom along with loss of the flanking domains in certain clades (Figure 3A,D,E). Crucially, the baeRF-1 have never displaced the classical bacterial RF1/2 proteins, suggesting that they have been adapted for a specialized role rather than functioning as the primary RF-PH in those bacteria that possess them.

Additionally, the classical aeRF-1 clade is related to a further clade which is comprised of four families with very distinctive phyletic patterns and domain architectures (Figure 3A). This clade is prototyped by the yeast Vms1 protein, which was initially characterized as interacting with the AAA+ ATPase chaperone CDC48 [46,47]. Accordingly, this clade uniting the four families was named the Vms1-like RF-1 clade (VLRF-1) clade (Figure 2G). The eukaryotic family of this clade is referred to as the eVLRF-1 family and is found in most eukaryotes except the basal-most diplomonad and parabasalid lineages and the apicomplexa. Notably, this eukaryotic family is related to the exclusion of all other families in the VLRF-1 clade to a family found in certain bacteria, predominantly Bacteroidetes and certain heimdallarchaeota (bVLRF-1 family, Figure 3A). Additionally, the VLRF-1 clade contains two other families, which are found, respectively, in a subset of the archaea and some bacteria, primarily of the actinobacterial and chloroflexi lineages. The archaeal VLRF-1 family (aVLRF-1; Figure 2), while lacking the C-terminal Pelota domain of the classical aeRF-1 clade, has a long coiled-coil extension N-terminal to the stop-codon-binding domain (Figure 3G). The actinobacterial/chloroflexi VLRF-1

family (acVLRF-1, Figure 2) has lost the stop-codon-binding domain and is instead fused to a distinctive short N-terminal extension (Figure 3G).

Phylogenetic analysis and domain architectural patterns indicate that the VLRF-1 clade emerged within archaea and was then transferred to bacteria independently of the transfer that spawned the baeRF-1 clade (Figure 3A,G,H and Figure 5A,B). Strikingly, this analysis indicates that the eukaryotic VLRF-1 family was not directly derived from the aVLRF-1 family, to which they are only distantly-related, but rather from the bVLRF-1 family (Figures 3A and 5A). This conclusion is also supported by the sequence conservation shared by the bVLRF-1 and eVLRF-1 in the N-terminal extension that is only found in the bacterial families of the VLRF-1 clade (Figure 3A,G,H). However, members of the eukaryotic clade have considerably diversified in domain architecture. In early-branching eukaryotes, like kinetoplastids, there is an extension to the ancestral bVLRF-1 core either in the form of an α-helical N-terminal domain seen in kinetoplastids, or a F-box domain which is seen in the heterolobosean Naegleria and some ciliates (Figure 3H). In the common ancestor of stramenopiles, plants, amoebozoans, animals, and fungi, the core was embellished by the accretion of a number of other domains: (1) a N-terminal Rei1-like C2H2-Zn-finger (Rei1-ZnF); (2) a C-terminal run of 3 ankyrin repeats; (3) novel treble-clef fold Zn-binding domain which we name the a Vms1-treble-clef (VTC) domain; and (4) VIM, a motif which has been previously shown to be a diagnostic feature of proteins that interact with the AAA+ ATPase chaperon CDC48 [48] (Figure 3H). In some cases, such as the yeast Saccharomyces cerevisiae and related fungi, the VTC domain might be lost, and in stramenopiles the order of the VTC domain and the VIM are inverted (Figure 3H). Of these domains, the Rei1-ZnF is notable in that it is found in the Rei1 protein involved in eukaryotic ribosome assembly. In Rei1, the Rei1-ZnF domain likely binds the ribosome large-subunit near the eL23 protein, while a C-terminal helical region blocks the ribosome exit channel to keep it protected prior to actual use in the cytoplasm [49] (Figure 3I). The Rei1-ZnF is also found in some previously uncharacterized eukaryotic proteins, where it is fused to the C-terminus of the Shwachman-Bodian-Diamond syndrome (SBDS) domain involved in the late cytoplasmic assembly of the eukaryotic 60 ribosomal subunit by removal of eIF6 (Figure 3I) [50,51]. These features strongly suggest that the eukaryotic VLRF-1 proteins are likely to associate with the ribosome via their N-terminal Rei1-finger domain.

2.3. Members of the aeRF-1 Superfamily Display Sequence and Structure Diversity in Their Core RNase H Fold Domain

The RF-PH catalytic domain shared by all aeRF-1s belongs to the RNase H fold, which encompasses a vast assemblage of protein domains descending from a common ancestor (Figure 2A,C). These domains can be classified into multiple distinct clades, each sharing their own functional attributes. The most widely known of these clades is the titular Ribonuclease H-like superfamily, which contains various distinct families functioning as RNA and DNA endo- and exonucleases. Another distinct assemblage with this fold, typically with two tandem RNase H fold domains, possess NTPase activity [52] and act as cytoskeletal proteins (actin/FtsA) or NTP-dependent chaperones (e.g., HSP70). Related families within this fold also function as accessory scaffolding domains for other enzymes including nitrogenase, creatinase/prolidase, and DNA methyltransferases [53–55]. However, an analysis of the families of domains sharing RNase H domain reveals that the earliest branching assemblage of the RNase H fold is comprised of the catalytically inactive versions found in the conserved ribosomal proteins such as L18 and S11. A key active site residue, typically an aspartate or glutamate, found in the first strand (strand-1) of the RNase H domain separates the nuclease/NTPase versions from non-enzymatic versions [56–58]. As previously mentioned, the ribosomal versions lack the catalytic residue that plays the RNA-binding/structural role in the ribosome. The RNase H domain in the aeRF-1 superfamily is specifically related to these ribosomal proteins to the exclusion of all other RNase H fold domains. Consistent with this, the aeRF-1 RNase H domain lacks the aforesaid active site residue in the strand-1 typical of NTPases/nucleases of this fold (Figure 2C). Instead, it has an extended loop, which bears the catalytic glutamine, connecting the third of the conserved β-strands of the fold to the downstream

α-helix. The catalytic glutamine is deeply inserted into the A-site to come close to the peptidyl transferase ribozyme active site, and as stated above, directs a water molecule to this active site [24,25] (Figure 2A,C).

This structural and evolutionary framework of the RNase H fold helps to glean functional insights regarding the newly identified aeRF-1 superfamily clades. First, analysis of the multiple sequence alignments of the new aeRF-1 clades (Figure 2G) allows identification of versions which might be catalytically active like the classic aeRF-1 as opposed to being inactive like Dom34 on the basis of the catalytic glutamine in the above-mentioned loop (Figure 2A,C). Additionally, aeRF-1 superfamily members might contain one or more conserved arginines at the N-terminus of the helix following the three consecutive N-terminal β-strands of the RNase H fold (Figure 2A,G). These residues serve to bind the large subunit rRNA and have also been implicated in interaction of eRF-1 with the cognate GTPase RF (eRF-3) as a tRNA-mimic (see below) [24,59].

Inspection of the multiple sequence alignments shows that the VLRF-1 clade contains a comparable loop with a conserved Q as in the catalytically active classic aeRF-1s, as well as several conserved arginine residues in the helix downstream of the core β-strands (Figure 2G). Further, on account of their three-domain architecture, the classic aeRF-1 and Dom34 families are believed to function as tRNA mimics that interact with the cognate RF GTPase, in a similar way to the AA-tRNA interacting with EF-Tu/EF-1α [11,32]. In eukaryotes, at least, VLRF-1 RF-PH domains have acquired flanking domains that could confer a multi-domain architecture that could approximate a tRNA mimic-like structure. Hence, they could in principle interact with a cognate GTPase, however, the identity of such a GTPase remains unclear and at least some circumstantial evidence suggests that GTPase activity is not coupled to peptidyl hydrolase activity in the eukaryotic VLRF-1 family [60]. Instead, the domains fused to the eukaryotic VLRF-1 family could be involved in an interaction with the ribosomal proteins, particularly the above-mentioned Rei1-ZnF domain (Figure 3H), in addition to the VIM motif interacting with the ATPase chaperone CDC48 [48].

Recently, it has been asserted that the lack of a conserved di-glycine amino acid pair in the eukaryotic VLRF-1 family, which is an absolutely conserved, convergently acquired feature of both the classical aeRF-1 family and the bRF-PH domains, calls into question its previously demonstrated function as a peptidyl-hydrolase [61]. This reasoning was used as justification for claiming the alternative RNase activity for the domain. However, this assertion relies on the supposition that the di-glycine amino acid pair should be required for peptidyl-hydrolase enzymatic activity. This is a position unsupported by the balance of experimental evidence which has converged on an understanding that the motif supplies flexibility to the RF-PH catalytic loop to ensure proper positioning of the catalytic glutamine in the active site pocket [15,62]. The above claim also overlooks existing small residue conservation in both the eukaryotic VLRF-1 and bVLRF-1 families: a single absolutely conserved glycine residue is always coupled with a near absolutely conserved tiny (typically either a serine or alanine) residue immediately upstream of the glutamine residue to form a GxuQ motif (where "u" is a tiny residue), a conservation pattern very much comparable to known RF-PH domains. Indeed, the newly discovered families can be viewed as the continuation of the spectrum of small residue conservation seen in proximity of the catalytic glutamine residue. For example, the classical aeRF-1 domains carry a GGQ motif, the bacterial RF domains carry a uxGxGGQ motif, and both the aVLRF-1 and acVLRF-1 families carry a GGxSQ motif (Figure 2G). Additionally, the VLRF-1 family has a longer catalytic loop than its classical aeRF-1 counterpart (Figure S1), a feature that could accommodate the necessary flexibility needed to access the peptidyl transferase active site, perhaps in combination with the absence of the small subunit in the rescue context.

Shannon entropy plots for the sequence alignments of the broadly disseminated families in the baeRF-1 clade (e.g., families 1,2,6), as compared with the other families, indicate poor conservation in the region at and surrounding the catalytic loop, with punctuated exceptions in the downstream helix largely representing conservation of specific arginine residues (Figure S2). Unlike the classical aeRF-1 and VLRF-1 clades, most families in the baeRF-1 clade, particularly those versions which have lost one

or both the flanking RNA-binding domains, usually appear to lack the conserved glutamine residue in the catalytic loop. This, together with the divergence in this region, suggests that they are catalytically inactive as peptidyl hydrolases. However, a few baeRF-1 families, perhaps most notably family 10 (Figures 2G and 3A), do contain a well-conserved glutamine residue typical of the active families and could potentially function as peptidyl hydrolases. Further, an analysis of the distribution of the lengths of the catalytic loop of the baeRF-1 clade families revealed a striking lack of constraint on its length, notably in the families lacking the catalytic glutamine (Figure S1). This pattern sharply contrasts with observations in families that contain the conserved glutamine and are known to or predicted to catalyze peptidyl hydrolase reactions. While these loop lengths can be different between these families, they are confined to a narrow range within any given family. The conservation of arginines or otherwise positively charged residues in the helix downstream of the catalytic loop (Figure 2A,G), even in versions lacking fusions to the codon recognition and pelota domains (Figure 3E), suggests that representatives of the baeRF-1 clade broadly interact with the LSU rRNA and likely insert into the A-site.

These observations regarding most of the baeRF-1 families may be compared to the catalytically inactive Dom34 family and all share poor conservation in the catalytic loop and absence of the catalytic glutamine (Figure 2G). Further, the above-noted, persistent operonic association with the ribosome hibernation factor family of proteins (Figure 3F) suggests that these families might typically interact with the ribosome in conjunction with these factors. These features suggest that they might function as non-catalytic negative regulators of translation, which insert into the A-site in lieu of a tRNA to facilitate ribosomal inhibition or disassembly by inserting into the A-site. However, the few families conserving the catalytic glutamine might function as alternative RF-PHs that are used in special circumstances, much like the eukaryotic VLRFs.

2.4. Diversity and Phylogeny of bRF-PHs in the Bacteria

The second type of RF-PH, i.e., bRF-PHs, have been noted previously to exist as a range of paralogous versions, however, the totality of this diversity and its evolutionary implications has not been described. In domain architectural terms, the canonical bRF-PHs contain an N-terminal helical GTPase-interacting domain which is followed by the stop-codon-recognition domain, into which the RF-PH domain with the convergently evolved catalytic glutamine has been inserted (Figures 2F and 4B). As noted above, most bacterial genomes, excepting severely reduced genomes such as in some tenericutes, contain two copies of this version (RF1 and RF2), which specialize in different sets of stop codons during general translation termination (Figure 4A). This suggests that both RF1 and RF2 were likely present in the common ancestor of all bacteria and arose via duplication in the stem lineage of all extent bacteria.

Two additional versions of the bRF-PH, both of which retain the catalytic glutamine residue (Figure 2B,D), are widespread in bacteria. The first of these is the ArfB/yaeJ, which contains a standalone RF-PH domain connected by an extended linker to a C-terminal α-helical region (Figure 4C). This helical segment localizes to the incompletely occupied codon recognition center of ribosomes that have been "jammed" by non-stop translation events [63,64]. Thus, ArfB functions as a single-component bacterial ribosome-rescue factor. ArfB is widespread in bacteria but is sporadically absent in the bacteroidetes and the actinobacteria, and missing entirely in other lineages including the firmicutes, fusobacteria, thermatogae, aquificae, and thermus-deinococci (Figure 5B). Despite these absences, the phyletic pattern of ArfB is suggestive of a trajectory wherein ArfB emerged early in bacterial evolution and has been subsequently lost in some lineages, possibly due to the presence of other functionally overlapping rescue pathways in bacteria [65,66] (see below, Figure 5A). Being a single-component rescue pathway, ArfB does not show extensive conserved gene neighborhood association beyond a persistent association with a pseudouridine (PSU) synthase (PSYN) gene across several proteobacterial classes (Figure 4D, Supplementary Data). Strikingly, this PSYN appears most closely-related to the 23S rRNA-modifying PSYN family which includes RluC and RluD [67], suggesting a coupling of non-stop

rescue to either the PSU modification of the peptidyl transferase ribozyme active site or the mRNAs themselves [68].

The other bacterial Rf-PH that has been relatively poorly studied is the prfH or RFH release factor. This protein lacks the N-terminal GTPase-interacting domain, and in contrast to ArfB/yaeJ, retains a version of the anticodon recognition domain [69] (Figure 4E). It was originally reported as showing a conserved gene neighborhood association with the then functionally uncharacterized RtcB protein (Figure 5F) [69], which was subsequently demonstrated to function as the RNA ligase [70–72]. A subsequent analysis classified this RFH-RtcB association as a representative of a diverse and vast collection of systems which repair RNA in response to damage caused by toxin effectors deployed in the biological conflicts between invasive elements like plasmids, viruses, and conjugative elements and their host cells [73]. In some cases, as is typical of many counter-conflict systems, further accretion of accessory genes around the core two-gene system (RFH-RtcB) is observed (see Supplementary Data for complete list): (1) A gene coding for a TPR domain, which could serve as a scaffold for other components comparable to the TPR protein Rqc1 in eukaryotic ribosome rescue (see below, Figure 4G–I); (2) A gene encoding a previously uncharacterized member of the Polβ superfamily, members which catalyze nucleotidyltransferase (NTase) reactions (Figure 4G,I). Another member of the polβ superfamily, the RlaP NTase, is found across several distinct counter-RNAse repair systems [73]. While this RFH-associated NTase does not appear to be specifically related to the RlaPs, it possibly plays a biochemically equivalent role in repairing RNAs like tRNAs that may have undergone further exonucleolytic degradation after an initial endonucleolytic cleavage event [73,74]; (3) A protein combining one or two ZnR domains with a domain of unknown provenance containing a highly positively charged helical region at the C-terminus (Figure 4H,I); (4) A protein again lacking any detectable homologs but possessing highly charged helical regions (Figure 4I). These latter two proteins could interact with rRNA either via the N-terminal ZnF domains or through their various charged regions. One further possibility is that one or more of these positively charged helical regions could function as charged "tails" comparable to those found in ArfB and other bacterial ribosome rescue factors (see below).

Thus, contextual evidence from conserved gene neighborhoods strongly supports a role for prfH/RFH as a release factor that is specifically coupled to RNA-repair by the RtcB ligase. This contrasts with ArfB which functions in an "institutionalized" rescue pathway responding to stalled ribosomes arising from intrinsic transcription and translational failures. Specifically, the prfH/RFH system likely functions in rescue necessitated by impairment of the ribosome by the action of extrinsic attacks from RNase effectors [73], where it hydrolyzes the peptidyl-tRNA even as the associated RtcB ligase repairs cleaved RNAs that jam the ribosome. Accordingly, the prfH/RFH-RtcB operonic unit is sporadic in its phyletic distribution, indicating dissemination via horizontal gene transfer (HGT) (Figure 5B, Supplementary Data).

2.5. bRF-PH Domains Transferred to Eukaryotes

Just as the RF-PH domains of the aeRF-1 superfamily have been transferred to bacteria, bRF-PH domains have also been transferred to eukaryotes. The provenance and distribution of these RF-PHs has been previously studied by Huynen et al. [42] who note two release factors acquired by eukaryotes that retain the ancestral architecture, respectively, descending from the bacterial RF1 and RF2 (Figure 4A). These are well-conserved across most eukaryotes with functional mitochondria and appear to operate within a similar stop-codon-recognition dichotomy as their bacterial progenitors. They have occasionally given rise to further lineage-specific versions [42]. Additional bRF-PHs are also prevalent in eukaryotic lineages which acquired photosynthetic plastid organelles [42].

However, another bRF-PH transferred to eukaryotes, ICT1, is highly conserved in eukaryotes. Strikingly, this version is clearly related to the ArfB/yaeJ family of bacterial ribosome rescue release factors and retains the C-terminal helical region peculiar to these bRF-PHs (Figure 4C). It has been recently demonstrated to perform an equivalent role as ArfB/yaeJ with respect to the mitochondrial

ribosomes [75]. A further version, the C12orf65 protein has a parallel architecture combining the RF-PH domain with a C-terminal positively charged helical region. While this RF-PH remains experimentally uncharacterized, it has been linked to several mitochondrial deficiencies [76–79] and is speculated to play an overlapping role with ICT1 in organellar ribosomal rescue pathways [77]. Intriguingly, we identify a fusion between C12orf65-like proteins and the MJ1316 domain which recognizes RNA 2'-3' cyclic nucleotide ends [73,80] (Figure 4C). The MJ1316 domain displays diverse fusions across eukaryotic lineages to domains involved in RNA repair [73], suggesting that the release factor activity of C12orf65 might be potentially coupled with repair of organellar RNA damage as proposed above for the RFH/prfH system [73]. The relative rates of evolution of the two proteins, with ICT1 being more conserved than c12orf65, could be interpreted as supporting a scenario where it operates along with RNA repair, specifically occurring within the context of biological conflicts.

The current genomic data allows detection of C12orf65 homologs in several additional lineages beyond those previously reported [42], including the basally positioned kinetoplastids. Therefore, the current data points to more similar phyletic patterns for ICT1 and C12orf65 than was apparent before. However, we observe that the ArfB family as currently defined contains a previously unidentified structural dichotomy. Specifically, while several members hew close to the observed structure of the canonical RF1/RF2 bRF-PHs catalytic domains (Figure 2B,D), a large subset contains a helical segment (previously labeled as the αi segment [81,82]) between the final two β-strands of the core β-sheet of the bRF-PH domain (Figure 2B). Hereafter, we refer to these versions as ArfB class 1, with the more ancestral versions referred to as the classical ArfB. Strikingly, comparison of known structures to multiple sequence alignments reveals that the ICT1 RF-PH domain clearly contains this αi segment, suggesting direct descent from the ArfB class 1 while C12orf65 lacks the segment. Phylogenetic analysis further supports this distinction, clearly grouping ICT1 with ArfB class 1 and C12orf65 with classical ArfB domains (Figure 4A). This strongly supports a scenario where the eukaryotes acquired two distinct bRF-PHs early in their evolution, likely from the independent transfer of two distinct antecedents of the ArfB family (Figures 4A and 5A).

A final, previously unreported version of the bRF-PH domain in eukaryotes is prototyped by a version found in the mitochondrial ribosomal protein S35 (MRPS35). This version lacks the GTPase-interacting and codon-recognition domains, in a similar way to ICT1 and c12orf65 versions. However, MRPS35 has acquired two distinctive helical domains flanking the core bRF-PH domain (Figure 4J, Supplementary Data for complete list). Unlike ICT1 and C12orf65, MRPS35 is enzymatically inactive, as it lacks the conserved glutamine in the catalytic loop. Despite the considerable sequence divergence between MRPS35 and other characterized bRF-PHs, the presence of the αi helical segment (Figure 2B) suggests an origin via duplication of ICT1 (Figure 4A), with subsequent loss of the catalytic glutamine residue and recruitment for a structural role in the mitochondrial ribosome [83,84]. It is also possible that MRPS35 could function in some non-catalytic rescue role, perhaps approximating the role of Dom34 in the archaeo-eukaryotic lineage in triggering ribosomal subunit disassociation. Along these lines, it is interesting to note that ICT1 similarly has been observed as associating with the large subunit of the mitochondrial ribosome [85–87].

3. Functional and Evolutionary Implications

3.1. The State of Translation Termination in the LUCA and the Early History of RF-PH Domains

The observation that evolutionarily unrelated RF-PHs emerged early in the "stems" of both the bacterial and archaeo-eukaryotic lineages, together with the recruitment of distinct versions of cognate RF-GTPases, has long stood as an evolutionary puzzle. The solution to this paradox has been to either speculate that one of the types of extant RF-PHs was present in the LUCA and displaced in one or the other lineage, or that a different factor was ancestrally responsible for peptide release, with each of the two lineages independently evolving their own solution that displaced the ancestral mechanism (Figure 1).

On the surface, the demonstration in recent years that both the aeRF-1 and bRF-PH superfamilies have extensive diversity in all three superkingdoms of Life would seem to argue for the first scenario. Under this scenario these newly uncovered radiations potentially represent remnants of the ancestral RF-PH in opposing lineages. However, phylogenetic and domain architectural evidence places the ultimate provenance of these newly described release factors squarely within either the archaeo-eukaryotic (aeRF-1 RF-PHs) or the bacterial lineage (bRF-PHs). In the case of the baeRF-1 clade the diversification in the bacteria likely occurred after a single transfer of a classical aeRF-1 version from the archaeo-eukaryotic lineage to a bacterium early in their evolution followed by dissemination via extensive HGT within bacteria (Figure 5A). All characterized bRF-PHs in eukaryotes have predominantly organellar functions and are rooted within the bacterial radiation of these domains. This strongly suggests that they were all acquired by the stem during the primary endosymbiotic event that defined eukaryogenesis or from other early endosymbionts. Subsequently, they underwent diversification early in eukaryotic evolution (Figure 5A). Together, these observations instead point in the direction of the second hypothesis where no conserved RF-PH was inherited from the LUCA.

Several lines of evidence support a form of the second scenario: first, the observation that the stop signal is also a base triplet, like the codons, used to encode the amino acids suggests that in the ancestral organism the stop was treated similar to the regular codons which specify amino acids. Second, the same stop codons are used in both the archaeo-eukaryotic and bacterial divisions. This indicates that the same stop signals were already present and interpreted in the LUCA. While currently these stop codons are decoded by protein RF-PHs, in both the archaeo-eukaryotic and bacterial divisions of life, the primary versions of the RF-PHs in the two divisions of life are evolutionarily unrelated. However, they notably utilize a biochemically equivalent glutamine in conjunction with the universal peptidyl transferase active site to catalyze peptide release. Considering these observations together, a plausible scenario for the release factor in the LUCA is the presence of a modified tRNA-like molecule (Figure 1). It is conceivable that this tRNA-like molecule presented a water molecule to the peptidyl transferase active site using a glutamine-like moiety attached to it. This is consistent with the class of hypotheses which favor a larger role for RNAs in the early stages of life [88–90]. In this regard, it is also notable that the bacteria and the archaeo-eukaryotic lineages show differences in ribosome recycling which follows termination. Interestingly, the recycling factor specific to the bacterial lineage, RRF, is related to the non-catalytic tRNA-binding domain of arginyl tRNA synthetases and the catalytic editing domain of the alanyl and threonyl tRNA synthetases [91]. Indeed, RRF has been observed to bind RNA in a similar way to both domains of aatRs [92,93]. This relationship might again favor the ancestral presence of a dedicated tRNA in peptidyl chain release, with RRF being derived from an RNA-binding protein which perhaps bound such a tRNA in the ancient translation system (Figure 1).

Under the above scenario this function of the ancestral tRNA-like molecule would have been independently taken over by protein factors only after the divergence of the bacterial and the archaeo-eukaryotic lineages. This was perhaps favored due to the ability of a protein to allow unambiguous decoding of the stop codons along with resistance to misreading by other amino-acylated tRNAs that resembled the original tRNA-like stop codon decoder. Considering this proposal, it is tempting to suggest that the tmRNA from the tmRNA-smpB system (Figure 5B) in bacteria might secondarily resemble some aspects of this proposed termination mechanism in LUCA (Figure 1). It is conceivable that this mechanism survived in the context of an ancient rescue pathway in the stem of the bacterial lineage. The simultaneous coupling of this pathway to the degradation of unfinished polypeptide chains could have in turn contributed to the failure of complete prokaryotic ubiquitin systems to gain a foothold as a core process in the bacteria [94–97].

3.2. Multiple Early Recruitments of Paralogous RF-PHs to Distinct Ribosome Rescue Pathways

Another important point that our survey of the RF-PHs raises is that multiple RF-PH-like proteins can coexist alongside the primary RF-PHs in the same organism without displacing the latter. Moreover, RF-PHs from either division of life (bacterial or archaeo-eukaryotic) can function in the other. In the

case of the aeRF-1 superfamily it is apparent that RNase H fold RF-PH domain can function in the A-sites of both archaeo-eukaryotic and bacterial-type ribosomes. It remains to be seen if the same might apply for bRF-PH domains, like those of ICT1 or c12orf65. Further, this wide adoption of diverse RF-PH paralogs points to their utility in contexts that require peptide release or translation termination beyond the regular recognition of stop codons. These contexts involve either relief from endogenous ribosomal stalling events, which can harm the cell through wasted resource allocation or production of misfolded proteins that can form aggregates, or relief from "ribosome-jamming" arising from the action of effectors on the translation apparatus for killing the cell or inducing stasis. Recruitment of RF-PH paralogs to the pathways which combat these insults was thus driven by the benefits accrued from relief of ribosome stalling. Particularly in the latter contexts, RF-PHs would need repeated innovation to cope with new translation-targeting effectors that emerge as part of escalating arms races in biological conflicts [98,99].

An example of such duplication followed by fixation occurred early in the archaeo-eukaryotic lineage with the emergence of the catalytically inactive Dom34 clade, which appears to have specialized in relieving stalled ribosomes due to a variety of reasons such as the lack of stop codons in aborted or otherwise truncated transcripts, or the presence of a blocking secondary structure on the mRNA. Rather than active hydrolytic release of the nascent peptide from the tRNA, this factor simply acted to separate the ribosomal subunits and terminate translation of the defective transcript [66,100].

Comparably in bacteria, rescue solutions including the tmRNA/SmpB and ArfA systems emerged largely independently of a secondary peptidyl hydrolase activity [65,66]. These proteins harbor charged helical regions that recognize non-stop translational stalling events via interaction with exposed regions in the anticodon recognition site. In this regard it is has been recently reported that the ArfT protein in *Francisella tularensis* with a comparable charged α-helix can similarly recruit the canonical bRF-PHs, RF1/RF2, to stalled ribosomes [101]. We observed that rather than being a *Francisella*-specific novelty, the ArfT protein is a single representative of a broadly distributed family with sporadic representation in several bacterial lineages, and even certain archaea and eukaryotes such as fungi and plants (Figure 5B, Supplementary Data). Several of these newly recognized bacterial and plant ArfT proteins had been previously implicated in stress response, consistent with a role in ribosome rescue [102–104]. The ArfT α-helix implicated in direct ribosome interaction is often found in multiple copies. ArfT-like proteins were observed containing arrays between one to eight such repeats (Figure 4K, Supplementary Data). Thus, such mechanisms of recognizing stalled ribosomes could also proceed via the recruitment of general RF-PHs and serve as an alternative mechanism to rescue stalled ribosomes through direct hydrolysis of the peptidyl tRNA. This niche of ribosome rescue via catalytic action appears to have been further colonized with the emergence of dedicated active paralogs of the general bRF-PHs early in bacterial evolution in the form ArfB/YaeJ. This pathway has been retained in the mitochondrial ribosomes of eukaryotes centered on the ArfB/yaeJ homolog ICT1 [75].

A further context for the recruitment of RF-PHs is indicated by the bacterial RFH/prfH, which appears to be coupled with RNA repair. This potentially acts to release peptides from ribosomes which have been stalled due to an attack on tRNA or other RNAs by the action endoRNases that are widely deployed in biological conflicts [73]. Growing evidence points to the pervasiveness of effectors deployed in biological conflicts that "jam" the ribosome [73], i.e., cause translational stalling by targeting tRNA, rRNA, mRNA at the ribosome or in some cases ribosomal proteins [105–107]. Moreover, there is also strong support for the evolution of counter-strategies that help alleviate such blocks via repair of the damaged RNAs. Thus, the RFH/prfH family's action is potentially coupled with the repair of the damaged RNA by the linked RtcB RNA ligase (Figure 4F–I) [73]. While largely experimentally uncharacterized, another such bRF-PH, C12orf65 from eukaryotes, might play a similar role in the context of RNA damage (Figure 4C).

Considering the above, we posit that the widely distributed paralogous versions of the aeRF-1 superfamily in bacteria (baeRF-1 family) are also likely to play a role in comparable processes by relieving stalled ribosomes either via a non-catalytic mechanism like Dom34, or in some cases by

catalyzing the release of the stalled peptide. Consistent with this proposal, the baeRF-1 clade is widely disseminated by lateral transfer rather than showing a pattern of predominantly vertical inheritance that is typical of the "institutionalized" classical aeRF-1 and Dom34 versions (Figure 5B). Further, the genomic associations, which we recovered linking diverse baeRF-1 families with the ribosome hibernation factors family (YfiA/HPF) that facilitate formation of 100S and 70S inactive ribosomal aggregates [108–111], suggest a function linked to translational shutdown (Figure 3F). One possibility is that these aeRF-1 domain proteins block access to the peptidyl transferase active site at the ribosome and subsequently allow the YfiA/HPF-like proteins to lock the ribosomes as inactive aggregates during stress conditions or effector attacks. A wrinkle in this observation is that some baeRF-1 families are found as standalone RF-PH domains on the genome, in other cases they are found in the standard aeRF-1 three-domain arrangement (Figure 1). At least the versions with the full domain complement could directly recognize stop codons to halt translation, perhaps during stress conditions. In either case these mechanisms could preserve ribosomes for later resuscitation [112–115].

Further cases of cross-kingdom transfer of factors involved in rescue process are seen in the form of sporadic transfers to eukaryotic lineages of both the ArfT and the tmRNA/SmpB. While eukaryotic tmRNA homologs had been previously reported in certain stramenopiles [116–118], we identify several of their counterpart SmpB proteins (see Supplementary Data) and report a further presence in rhodophytes and the crytophyte *Guillardia*. The strongest sequence affinities for these SmpB proteins are with cyanobacteria, consistent with potential transfer during plastid endosymbiosis that gave rise to the original photosynthetic eukaryotes. While this tmRNA/SmpB system appears to have been lost in the plant lineage, it has persisted in some basal chloroplast-containing lineages and was also transferred via secondary endosymbiotic events to further lineages including the stramenopiles [119,120].

3.3. Origin of the Eukaryotic Ribosome Quality Control System

The eukaryotes are distinguished from the prokaryotic superkingdoms by possessing an elaborate and dedicated ribosome rescue system in the form of the RQC pathway which rescues stalled translation events. The release factor of this pathway remained unknown until recently [60]. Remarkably, the responsible release factor, Vms1/ANKZF, a member of the VLRF-1 clade of the aeRF-1 superfamily is most closely related specifically to those from the bacteroidetes lineage and Heimdallarchaeota (Figure 3A,G). Heimdallarchaeota are part of the Asgardarchaeal branch and are proposed to be the sister group of eukaryotes. This raises the possibility that VLRF-1 was specifically transferred from bacteroidetes to the common ancestor of Heimdallarchaeota. However, the absence of VLRF1 in basal eukaryotes suggests an alternative scenario. Multiple members of the bacteroidetes lineage are known to form endosymbiotic associations with eukaryotes [121], and early, we presented evidence for certain conserved eukaryotic proteins having their origins in bacteroidetes [99,122]. Hence, eukaryotic Vms1 could be another example of a protein with such an origin, pointing to the key role played by bacterial endosymbionts beyond the mitochondrial progenitor in the emergence of eukaryote-specific systems. In bacteria and archaea, the members of the VLRF-1 clade are standalone versions occurring independently of the other components of the RQC system. Thus, it is possible that they might function comparably to the RF-PH paralogs discussed above that function as single-component release factors in specialized contexts (Figure 5B). Thus, the core of the eukaryotic RQC systems appears to have emerged from the acquisition of such a specialized RF-PH.

The identification of Vms1 as the peptidyl-hydrolase of the RQC system follows the delineation of its interactions with other known components of the pathway, including the AAA+ ATPase CDC48 [48] and Rqc1, Rqc2, Ltn1, and Npl4 [60,123,124]. Of these, Rqc2 is particularly notable. Prior studies have suggested a role for it in the formation of template-independent C-terminal alanine-threonine (CAT) tails by the peptidyl transferase center to facilitate the extrusion of the stalled peptide from the exit channel for subsequent ubiquitination and degradation [125]. Earlier we had shown that this core player of the RQC complex goes back to the LUCA [126] (Figure 5B). Given its universal presence, a possible functional role for the protein in as-yet unknown ribosome rescue pathways in the bacteria

and archaea could be fruitful avenues of future experimental investigation. The role of the Rqc2 protein in these kingdoms could center on a process similar to CAT-tailing. Another, not mutually exclusive possibility is that these proteins could be involved in RNA modification reactions during rescue, as we had previously identified a relationship between the N-terminal domain of these proteins (NFACT-N) and the Fpg, MutM, and Nei/EndoVIII family of DNA glycosylases, and shown that several of the residues needed for catalyzing base removal were spatially shared by these two domains. The conserved gene neighborhood association observed between the ArfB and PSYN genes in some bacteria (Figure 4D, Supplementary Data) again points to potential links between RNA base modifications and ribosome rescue pathways.

Another key aspect of the emergence of the RQC system was its coupling with the expanded ubiquitin system for protein degradation, which is typical of eukaryotes. This happened via the Ltn1 and Npl4 proteins that emerged early in eukaryotic evolution. Our analysis also suggests that the Rqc1 family of proteins, featuring a tetratricopeptide repeat toroidal module, had their origins in bacteria. Thus, the eukaryotic RQC system can now be shown to be a complex innovation involving: an ancient core inherited from the LUCA in the form of Rqc2, an archaeo-eukaryotic innovation in the form of CDC48, and several eukaryote-specific innovations relating to the coupling with the Ub-system (e.g., Npl4 and Ltn1) and key catalytic and structural components of bacterial provenance (e.g., Vms1, Rqc1) (Figure 5A).

3.4. Alternative Catalytic Mechanisms for Releasing Peptides from Peptidyl-tRNA

In addition to the previously discussed RF-PHs, there are other enzymes that catalyze the release of amino acids and peptides from tRNA. Such enzymes have emerged in at least two structurally unrelated folds: (1) the bacterial-type D-Tyr-tRNA(Tyr) deacylases with the STAS-like fold and (2) the amino acyl/peptidyl hydrolases of the peptidyl-hydrolase-phosphorylase fold. The latter fold features a key family of such enzymes, universally conserved in bacteria, prototyped by the peptidyl-tRNA hydrolase (Pth) (Figure 5B). The Pth catalyzes the hydrolysis of the peptidyl-tRNA bond in translation products that prematurely dissociate from the ribosome active sites [127–129]. It functions in conjunction with the RRF and elongation factor-G [23,130,131] to complete the release of such aborted translation products from the ribosome. The Pth was acquired by eukaryotes from bacteria and is generally concordant with the presence of the mitochondria, outside of loss in some lineages such as apicomplexa. Remarkably, human Ptrh1 was recently described as a release factor in a variant RQC pathway where the nascent peptide chains had not yet been ubiquitinated [61]. The authors report that the tRNAs attached to these chains are not "fixed" to the P site, consistent with the bacterial Pth enzymes acting on translation intermediates that dissociate from the ribosomal active sites. The Pth superfamily belongs to a vast monophyletic assemblage of enzymes, and additionally includes several families with peptidase, phosphorylase, and dioxygenase activities. Another member of this assemblage, which is highly conserved in archaea, catalyzes a similar reaction, i.e., hydrolysis of the amino acyl bond of tRNAs mischarged with D-amino acids. This observation together with the fact that Pth appears to be one of the earliest branching members of this assemblage suggests that the hydrolysis of aminoacyl/peptidyl tRNA in contexts other than translation termination might have been an ancient function of this assemblage of enzymes.

4. Conclusions

Termination is a universal step in the translation process. However, paradoxically, sequence and structure analysis indicate that unlike the factors mediating the key steps of amino acylation, peptide-bond formation and elongation, the peptidyl hydrolases and GTPases catalyzing translation termination are evolutionary distinct or unrelated in the two great divisions of life. This leads to the idea that the LUCA probably did not possess a dedicated protein-based apparatus for release of the peptidyl-tRNA during transcription termination, and that this function appears to have convergently evolved only after the bacterial and archaeo-eukaryotic lineages separated. Further,

recent analyses of the RF-PH domains have identified several novel versions. The genomic contexts and experimental evidence hint at a diversity of contexts in which such release factors might be deployed beyond conventional translation termination. These appear to be key players in processes like ribosomal quality control and repair of ribosomes "jammed" by ribosome-targeting effectors deployed by biological conflict systems. These investigations also open the door for understanding novel ribosome rescue pathways, including in the archaea, where little is known about specialized translation termination processes. Further, the RNA repair-coupled translation termination systems (e.g., prfH/RFH system) offer opportunities for developing novel biotechnological applications.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/20/8/1981/s1.

Author Contributions: Conceptualization, L.A. and A.M.B.; formal analysis, A.M.B.; project administration, L.A.; visualization, A.M.B.; writing—original draft L.A. and A.M.B.; writing—review and editing, A.M.B. and L.A.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

RF-PH	release factor-peptidyl hydrolase
LUCA	Last Universal Common Ancestor
AA	amino acylated
bRF-PH	bacterial release factor-peptidyl hydrolase
aeRF-1	archaeo-eukaryotic release factor
RQC	ribosome quality control
ZnR	zinc ribbon
VLRF-1	Vms1-like archaeo-eukaryotic release factor
aVLRF-1	archaeal Vms1-like archaeo-eukaryotic release factor
eVLRF-1	eukaryotic Vms1-like archaeo-eukaryotic release factor
acVLRF-1	actinobacterial/chloroflexi Vms1-like archaeo-eukaryotic release factor
bVLRF-1	bacteroidetes Vms1-like archaeo-eukaryotic release factor
baeRF-1	bacterial archaeo-eukaryotic release factor
VTC	Vms1 treble clef
SBDS	Shwachman-Bodian-Diamond syndrome
PSU	pseudouridine
PSYN	pseudouridine synthase
NTase	nucleotidyltransferase
Pth	peptidyl-tRNA hydrolase
aatRs	aminoacyl-tRNA synthetases

References

1. Leipe, D.D.; Wolf, Y.I.; Koonin, E.V.; Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* **2002**, *317*, 41–72. [CrossRef] [PubMed]
2. Berchtold, H.; Reshetnikova, L.; Reiser, C.O.; Schirmer, N.K.; Sprinzl, M.; Hilgenfeld, R. Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature* **1993**, *365*, 126–132. [CrossRef]
3. Kavaliauskas, D.; Nissen, P.; Knudsen, C.R. The busiest of all ribosomal assistants: Elongation factor Tu. *Biochemistry* **2012**, *51*, 2642–2651. [CrossRef] [PubMed]
4. Yamamoto, H.; Qin, Y.; Achenbach, J.; Li, C.; Kijek, J.; Spahn, C.M.; Nierhaus, K.H. EF-G and EF4: Translocation and back-translocation on the bacterial ribosome. *Nat. Rev. Microbiol.* **2014**, *12*, 89–100. [CrossRef]
5. Rodnina, M.V.; Wintermeyer, W. The ribosome as a molecular machine: The mechanism of tRNA-mRNA movement in translocation. *Biochem. Soc. Trans.* **2011**, *39*, 658–662. [CrossRef]
6. Szkaradkiewicz, K.; Zuleeg, T.; Limmer, S.; Sprinzl, M. Interaction of fMet-tRNAfMet and fMet-AMP with the C-terminal domain of Thermus thermophilus translation initiation factor 2. *Eur. J. Biochem.* **2000**, *267*, 4290–4299. [CrossRef]

7.　Pestova, T.V.; Lomakin, I.B.; Lee, J.H.; Choi, S.K.; Dever, T.E.; Hellen, C.U. The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* **2000**, *403*, 332–335. [CrossRef] [PubMed]

8.　Merrick, W.C. Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* **1992**, *56*, 291–315.

9.　Inagaki, Y.; Blouin, C.; Susko, E.; Roger, A.J. Assessing functional divergence in EF-1alpha and its paralogs in eukaryotes and archaebacteria. *Nucleic Acids Res.* **2003**, *31*, 4227–4237. [CrossRef]

10.　Kobayashi, K.; Saito, K.; Ishitani, R.; Ito, K.; Nureki, O. Structural basis for translation termination by archaeal RF1 and GTP-bound EF1alpha complex. *Nucleic Acids Res.* **2012**, *40*, 9319–9328. [CrossRef]

11.　Nakamura, Y.; Ito, K.; Isaksson, L.A. Emerging understanding of translation termination. *Cell* **1996**, *87*, 147–150. [CrossRef]

12.　Nakamura, Y.; Ito, K. Making sense of mimic in translation termination. *Trends Biochem. Sci.* **2003**, *28*, 99–105. [CrossRef]

13.　Frolova, L.Y.; Tsivkovskii, R.Y.; Sivolobova, G.F.; Oparina, N.Y.; Serpinsky, O.I.; Blinov, V.M.; Tatkov, S.I.; Kisselev, L.L. Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA* **1999**, *5*, 1014–1020. [CrossRef] [PubMed]

14.　Mora, L.; Heurgue-Hamard, V.; Champ, S.; Ehrenberg, M.; Kisselev, L.L.; Buckingham, R.H. The essential role of the invariant GGQ motif in the function and stability in vivo of bacterial release factors RF1 and RF2. *Mol. Microbiol.* **2003**, *47*, 267–275. [CrossRef]

15.　Shaw, J.J.; Green, R. Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol. Cell* **2007**, *28*, 458–467. [CrossRef]

16.　Weixlbaumer, A.; Jin, H.; Neubauer, C.; Voorhees, R.M.; Petry, S.; Kelley, A.C.; Ramakrishnan, V. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* **2008**, *322*, 953–956. [CrossRef] [PubMed]

17.　Korostelev, A.; Asahara, H.; Lancaster, L.; Laurberg, M.; Hirschi, A.; Zhu, J.; Trakhanov, S.; Scott, W.G.; Noller, H.F. Crystal structure of a translation termination complex formed with release factor RF2. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19684–19689. [CrossRef]

18.　Laurberg, M.; Asahara, H.; Korostelev, A.; Zhu, J.; Trakhanov, S.; Noller, H.F. Structural basis for translation termination on the 70S ribosome. *Nature* **2008**, *454*, 852–857. [CrossRef] [PubMed]

19.　Inagaki, Y.; Ford Doolittle, W. Evolution of the eukaryotic translation termination system: Origins of release factors. *Mol. Biol. Evol.* **2000**, *17*, 882–889. [CrossRef]

20.　Vestergaard, B.; Van, L.B.; Andersen, G.R.; Nyborg, J.; Buckingham, R.H.; Kjeldgaard, M. Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1. *Mol. Cell* **2001**, *8*, 1375–1382. [CrossRef]

21.　Pisarev, A.V.; Skabkin, M.A.; Pisareva, V.P.; Skabkina, O.V.; Rakotondrafara, A.M.; Hentze, M.W.; Hellen, C.U.; Pestova, T.V. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol. Cell* **2010**, *37*, 196–210. [CrossRef]

22.　Shoemaker, C.J.; Green, R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E1392–E1398. [CrossRef]

23.　Zavialov, A.V.; Hauryliuk, V.V.; Ehrenberg, M. Splitting of the posttermination ribosome into subunits by the concerted action of RRF and EF-G. *Mol. Cell* **2005**, *18*, 675–686. [CrossRef] [PubMed]

24.　Song, H.; Mugnier, P.; Das, A.K.; Webb, H.M.; Evans, D.R.; Tuite, M.F.; Hemmings, B.A.; Barford, D. The crystal structure of human eukaryotic release factor eRF1–mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* **2000**, *100*, 311–321. [CrossRef]

25.　Cheng, Z.; Saito, K.; Pisarev, A.V.; Wada, M.; Pisareva, V.P.; Pestova, T.V.; Gajda, M.; Round, A.; Kong, C.; Lim, M.; et al. Structural insights into eRF3 and stop codon recognition by eRF1. *Genes Dev.* **2009**, *23*, 1106–1118. [CrossRef]

26.　Preis, A.; Heuer, A.; Barrio-Garcia, C.; Hauser, A.; Eyler, D.E.; Berninghausen, O.; Green, R.; Becker, T.; Beckmann, R. Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep.* **2014**, *8*, 59–65. [CrossRef]

27.　Anantharaman, V.; Koonin, E.V.; Aravind, L. Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.* **2002**, *30*, 1427–1464. [CrossRef]

28.　Caban, K.; Kinzy, S.A.; Copeland, P.R. The L7Ae RNA binding motif is a multifunctional domain required for the ribosome-dependent Sec incorporation activity of Sec insertion sequence binding protein 2. *Mol. Cell Biol.* **2007**, *27*, 6350–6360. [CrossRef]

29.	Mora, L.; Zavialov, A.; Ehrenberg, M.; Buckingham, R.H. Stop codon recognition and interactions with peptide release factor RF3 of truncated and chimeric RF1 and RF2 from Escherichia coli. *Mol. Microbiol.* **2003**, *50*, 1467–1476. [CrossRef]

30.	Pallesen, J.; Hashem, Y.; Korkmaz, G.; Koripella, R.K.; Huang, C.; Ehrenberg, M.; Sanyal, S.; Frank, J. Cryo-EM visualization of the ribosome in termination complex with apo-RF3 and RF1. *eLife* **2013**, *2*, e00411. [CrossRef]

31.	Kristensen, O.; Laurberg, M.; Liljas, A.; Selmer, M. Is tRNA binding or tRNA mimicry mandatory for translation factors? *Curr. Protein Pept. Sci.* **2002**, *3*, 133–141. [CrossRef]

32.	Nakamura, Y.; Ito, K. tRNA mimicry in translation termination and beyond. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 647–668. [CrossRef]

33.	Passos, D.O.; Doma, M.K.; Shoemaker, C.J.; Muhlrad, D.; Green, R.; Weissman, J.; Hollien, J.; Parker, R. Analysis of Dom34 and its function in no-go decay. *Mol. Biol. Cell* **2009**, *20*, 3025–3032. [CrossRef] [PubMed]

34.	Kobayashi, K.; Kikuno, I.; Kuroha, K.; Saito, K.; Ito, K.; Ishitani, R.; Inada, T.; Nureki, O. Structural basis for mRNA surveillance by archaeal Pelota and GTP-bound EF1alpha complex. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 17575–17579. [CrossRef]

35.	Saito, S.; Hosoda, N.; Hoshino, S. The Hbs1-Dom34 protein complex functions in non-stop mRNA decay in mammalian cells. *J. Biol. Chem.* **2013**, *288*, 17832–17843. [CrossRef] [PubMed]

36.	Doma, M.K.; Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **2006**, *440*, 561–564. [CrossRef]

37.	Guydosh, N.R.; Green, R. Dom34 rescues ribosomes in 3' untranslated regions. *Cell* **2014**, *156*, 950–962. [CrossRef] [PubMed]

38.	Graille, M.; Chaillet, M.; van Tilbeurgh, H. Structure of yeast Dom34: A protein related to translation termination factor Erf1 and involved in No-Go decay. *J. Biol. Chem.* **2008**, *283*, 7145–7154. [CrossRef]

39.	Hilal, T.; Yamamoto, H.; Loerke, J.; Burger, J.; Mielke, T.; Spahn, C.M. Structural insights into ribosomal rescue by Dom34 and Hbs1 at near-atomic resolution. *Nat. Commun.* **2016**, *7*, 13521. [CrossRef]

40.	Carr-Schmid, A.; Pfund, C.; Craig, E.A.; Kinzy, T.G. Novel G-protein complex whose requirement is linked to the translational status of the cell. *Mol. Cell Biol.* **2002**, *22*, 2564–2574. [CrossRef]

41.	Scolnick, E.; Tompkins, R.; Caskey, T.; Nirenberg, M. Release factors differing in specificity for terminator codons. *Proc. Natl. Acad. Sci. USA* **1968**, *61*, 768–774. [CrossRef]

42.	Duarte, I.; Nabuurs, S.B.; Magno, R.; Huynen, M. Evolution and diversification of the organellar release factor family. *Mol. Biol. Evol.* **2012**, *29*, 3497–3512. [CrossRef]

43.	Zaremba-Niedzwiedzka, K.; Caceres, E.F.; Saw, J.H.; Backstrom, D.; Juzokaite, L.; Vancaester, E.; Seitz, K.W.; Anantharaman, K.; Starnawski, P.; Kjeldsen, K.U.; et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **2017**, *541*, 353–358. [CrossRef]

44.	Aravind, L. Guilt by association: Contextual information in genome analysis. *Genome Res.* **2000**, *10*, 1074–1077. [CrossRef]

45.	Huynen, M.; Snel, B.; Lathe, W., 3rd; Bork, P. Predicting protein function by genomic context: Quantitative evaluation and qualitative inferences. *Genome Res.* **2000**, *10*, 1204–1210. [CrossRef]

46.	Heo, J.M.; Livnat-Levanon, N.; Taylor, E.B.; Jones, K.T.; Dephoure, N.; Ring, J.; Xie, J.; Brodsky, J.L.; Madeo, F.; Gygi, S.P.; et al. A stress-responsive system for mitochondrial protein degradation. *Mol Cell* **2010**, *40*, 465–480. [CrossRef]

47.	Tran, J.R.; Brodsky, J.L. The Cdc48-Vms1 complex maintains 26S proteasome architecture. *Biochem. J.* **2014**, *458*, 459–467. [CrossRef]

48.	Stapf, C.; Cartwright, E.; Bycroft, M.; Hofmann, K.; Buchberger, A. The general definition of the p97/valosin-containing protein (VCP)-interacting motif (VIM) delineates a new family of p97 cofactors. *J. Biol. Chem.* **2011**, *286*, 38670–38678. [CrossRef]

49.	Greber, B.J.; Gerhardy, S.; Leitner, A.; Leibundgut, M.; Salem, M.; Boehringer, D.; Leulliot, N.; Aebersold, R.; Panse, V.G.; Ban, N. Insertion of the Biogenesis Factor Rei1 Probes the Ribosomal Tunnel during 60S Maturation. *Cell* **2016**, *164*, 91–102. [CrossRef]

50.	Finch, A.J.; Hilcenko, C.; Basse, N.; Drynan, L.F.; Goyenechea, B.; Menne, T.F.; Gonzalez Fernandez, A.; Simpson, P.; D'Santos, C.S.; Arends, M.J.; et al. Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes Dev.* **2011**, *25*, 917–929. [CrossRef]

51. Wessels, D.; Srikantha, T.; Yi, S.; Kuhl, S.; Aravind, L.; Soll, D.R. The Shwachman-Bodian-Diamond syndrome gene encodes an RNA-binding protein that localizes to the pseudopod of Dictyostelium amoebae during chemotaxis. *J. Cell Sci.* **2006**, *119*, 370–379. [CrossRef]

52. Andreeva, A.; Howorth, D.; Chandonia, J.M.; Brenner, S.E.; Hubbard, T.J.; Chothia, C.; Murzin, A.G. Data growth and its impact on the SCOP database: New developments. *Nucleic Acids Res.* **2008**, *36*, D419–D425. [CrossRef]

53. Dyer, D.H.; Rubio, L.M.; Thoden, J.B.; Holden, H.M.; Ludden, P.W.; Rayment, I. The three-dimensional structure of the core domain of Naf Y from Azotobacter vinelandii determined at 1.8-A resolution. *J. Biol. Chem.* **2003**, *278*, 32150–32156. [CrossRef]

54. Padmanabhan, B.; Paehler, A.; Horikoshi, M. Structure of creatine amidinohydrolase from Actinobacillus. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 1322–1328. [CrossRef]

55. Wibley, J.E.; Pegg, A.E.; Moody, P.C. Crystal structure of the human O(6)-alkylguanine-DNA alkyltransferase. *Nucleic Acids Res.* **2000**, *28*, 393–401. [CrossRef]

56. Katayanagi, K.; Miyagawa, M.; Matsushima, M.; Ishikawa, M.; Kanaya, S.; Ikehara, M.; Matsuzaki, T.; Morikawa, K. Three-dimensional structure of ribonuclease H from *E. coli*. *Nature* **1990**, *347*, 306–309. [CrossRef]

57. Yang, W.; Hendrickson, W.A.; Crouch, R.J.; Satow, Y. Structure of ribonuclease H phased at 2 A resolution by MAD analysis of the selenomethionyl protein. *Science* **1990**, *249*, 1398–1405. [CrossRef]

58. Majorek, K.A.; Dunin-Horkawicz, S.; Steczkiewicz, K.; Muszewska, A.; Nowotny, M.; Ginalski, K.; Bujnicki, J.M. The RNase H-like superfamily: New members, comparative structural analysis and evolutionary classification. *Nucleic Acids Res.* **2014**, *42*, 4160–4179. [CrossRef]

59. Seit-Nebi, A.; Frolova, L.; Justesen, J.; Kisselev, L. Class-1 translation termination factors: Invariant GGQ minidomain is essential for release activity and ribosome binding but not for stop codon recognition. *Nucleic Acids Res.* **2001**, *29*, 3982–3987. [CrossRef]

60. Verma, R.; Reichermeier, K.M.; Burroughs, A.M.; Oania, R.S.; Reitsma, J.M.; Aravind, L.; Deshaies, R.J. Vms1 and ANKZF1 peptidyl-tRNA hydrolases release nascent chains from stalled ribosomes. *Nature* **2018**, *557*, 446–451. [CrossRef]

61. Kuroha, K.; Zinoviev, A.; Hellen, C.U.T.; Pestova, T.V. Release of Ubiquitinated and Non-ubiquitinated Nascent Chains from Stalled Mammalian Ribosomal Complexes by ANKZF1 and Ptrh1. *Mol. Cell* **2018**, *72*, 286–302.e288. [CrossRef]

62. Trobro, S.; Aqvist, J. A model for how ribosomal release factors induce peptidyl-tRNA cleavage in termination of protein synthesis. *Mol. Cell* **2007**, *27*, 758–766. [CrossRef]

63. Handa, Y.; Inaho, N.; Nameki, N. YaeJ is a novel ribosome-associated protein in Escherichia coli that can hydrolyze peptidyl-tRNA on stalled ribosomes. *Nucleic Acids Res.* **2011**, *39*, 1739–1748. [CrossRef]

64. Chadani, Y.; Ono, K.; Kutsukake, K.; Abo, T. Escherichia coli YaeJ protein mediates a novel ribosome-rescue pathway distinct from SsrA- and ArfA-mediated pathways. *Mol. Microbiol.* **2011**, *80*, 772–785. [CrossRef]

65. Huter, P.; Muller, C.; Arenz, S.; Beckert, B.; Wilson, D.N. Structural Basis for Ribosome Rescue in Bacteria. *Trends Biochem. Sci.* **2017**, *42*, 669–680. [CrossRef]

66. Buskirk, A.R.; Green, R. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **2017**, *372*. [CrossRef]

67. Huang, L.; Ku, J.; Pookanjanatavip, M.; Gu, X.; Wang, D.; Greene, P.J.; Santi, D.V. Identification of two Escherichia coli pseudouridine synthases that show multisite specificity for 23S RNA. *Biochemistry* **1998**, *37*, 15951–15957. [CrossRef]

68. Hoernes, T.P.; Clementi, N.; Faserl, K.; Glasner, H.; Breuker, K.; Lindner, H.; Huttenhofer, A.; Erlacher, M.D. Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Res.* **2016**, *44*, 852–862. [CrossRef]

69. Baranov, P.V.; Vestergaard, B.; Hamelryck, T.; Gesteland, R.F.; Nyborg, J.; Atkins, J.F. Diverse bacterial genomes encode an operon of two genes, one of which is an unusual class-I release factor that potentially recognizes atypical mRNA signals other than normal stop codons. *Biol. Direct.* **2006**, *1*, 28. [CrossRef]

70. Englert, M.; Sheppard, K.; Aslanian, A.; Yates, J.R., 3rd; Soll, D. Archaeal 3′-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 1290–1295. [CrossRef]

71. Popow, J.; Englert, M.; Weitzer, S.; Schleiffer, A.; Mierzwa, B.; Mechtler, K.; Trowitzsch, S.; Will, C.L.; Luhrmann, R.; Soll, D.; et al. HSPC117 is the essential subunit of a human tRNA splicing ligase complex. *Science* **2011**, *331*, 760–764. [CrossRef]

72. Tanaka, N.; Shuman, S. RtcB is the RNA ligase component of an *Escherichia coli* RNA repair operon. *J. Biol. Chem.* **2011**, *286*, 7727–7731. [CrossRef]

73. Burroughs, A.M.; Aravind, L. RNA damage in biological conflicts and the diversity of responding RNA repair systems. *Nucleic Acids Res.* **2016**, *44*, 8525–8555. [CrossRef]

74. Klaiman, D.; Steinfels-Kohn, E.; Krutkina, E.; Davidov, E.; Kaufmann, G. The wobble nucleotide-excising anticodon nuclease RloC is governed by the zinc-hook and DNA-dependent ATPase of its Rad50-like region. *Nucleic Acids Res.* **2012**, *40*, 8568–8578. [CrossRef]

75. Feaga, H.A.; Quickel, M.D.; Hankey-Giblin, P.A.; Keiler, K.C. Human Cells Require Non-stop Ribosome Rescue Activity in Mitochondria. *PLoS Genet.* **2016**, *12*, e1005964. [CrossRef] [PubMed]

76. Wesolowska, M.; Gorman, G.S.; Alston, C.L.; Pajak, A.; Pyle, A.; He, L.; Griffin, H.; Chinnery, P.F.; Miller, J.A.; Schaefer, A.M.; et al. Adult Onset Leigh Syndrome in the Intensive Care Setting: A Novel Presentation of a C12orf65 Related Mitochondrial Disease. *J. Neuromuscul. Dis.* **2015**, *2*, 409–419. [CrossRef]

77. Antonicka, H.; Ostergaard, E.; Sasarman, F.; Weraarpachai, W.; Wibrand, F.; Pedersen, A.M.; Rodenburg, R.J.; van der Knaap, M.S.; Smeitink, J.A.; Chrzanowska-Lightowlers, Z.M.; et al. Mutations in C12orf65 in patients with encephalomyopathy and a mitochondrial translation defect. *Am. J. Hum. Genet.* **2010**, *87*, 115–122. [CrossRef] [PubMed]

78. Pyle, A.; Ramesh, V.; Bartsakoulia, M.; Boczonadi, V.; Gomez-Duran, A.; Herczegfalvi, A.; Blakely, E.L.; Smertenko, T.; Duff, J.; Eglon, G.; et al. Behr's Syndrome is Typically Associated with Disturbed Mitochondrial Translation and Mutations in the C12orf65 Gene. *J. Neuromuscul. Dis.* **2014**, *1*, 55–63. [CrossRef]

79. Nishihara, H.; Omoto, M.; Takao, M.; Higuchi, Y.; Koga, M.; Kawai, M.; Kawano, H.; Ikeda, E.; Takashima, H.; Kanda, T. Autopsy case of the C12orf65 mutation in a patient with signs of mitochondrial dysfunction. *Neurol. Genet.* **2017**, *3*, e171. [CrossRef]

80. Mazumder, R.; Iyer, L.M.; Vasudevan, S.; Aravind, L. Detection of novel members, structure-function analysis and evolutionary classification of the 2H phosphoesterase superfamily. *Nucleic Acids Res.* **2002**, *30*, 5229–5243. [CrossRef]

81. Kogure, H.; Handa, Y.; Nagata, M.; Kanai, N.; Guntert, P.; Kubota, K.; Nameki, N. Identification of residues required for stalled-ribosome rescue in the codon-independent release factor YaeJ. *Nucleic Acids Res.* **2014**, *42*, 3152–3163. [CrossRef]

82. Handa, Y.; Hikawa, Y.; Tochio, N.; Kogure, H.; Inoue, M.; Koshiba, S.; Guntert, P.; Inoue, Y.; Kigawa, T.; Yokoyama, S.; et al. Solution structure of the catalytic domain of the mitochondrial protein ICT1 that is essential for cell vitality. *J. Mol. Biol.* **2010**, *404*, 260–273. [CrossRef]

83. Kaushal, P.S.; Sharma, M.R.; Booth, T.M.; Haque, E.M.; Tung, C.S.; Sanbonmatsu, K.Y.; Spremulli, L.L.; Agrawal, R.K. Cryo-EM structure of the small subunit of the mammalian mitochondrial ribosome. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7284–7289. [CrossRef]

84. Greber, B.J.; Bieri, P.; Leibundgut, M.; Leitner, A.; Aebersold, R.; Boehringer, D.; Ban, N. The complete structure of the 55S mammalian mitochondrial ribosome. *Science* **2015**, *348*, 303–308. [CrossRef]

85. Brown, A.; Amunts, A.; Bai, X.C.; Sugimoto, Y.; Edwards, P.C.; Murshudov, G.; Scheres, S.H.W.; Ramakrishnan, V. Structure of the large ribosomal subunit from human mitochondria. *Science* **2014**, *346*, 718–722. [CrossRef]

86. Amunts, A.; Brown, A.; Toots, J.; Scheres, S.H.W.; Ramakrishnan, V. The structure of the human mitochondrial ribosome. *Science* **2015**, *348*, 95–98. [CrossRef]

87. Brown, A.; Rathore, S.; Kimanius, D.; Aibara, S.; Bai, X.C.; Rorbach, J.; Amunts, A.; Ramakrishnan, V. Structures of the human mitochondrial ribosome in native states of assembly. *Nat. Struct. Mol. Biol.* **2017**, *24*, 866–869. [CrossRef]

88. Noller, H.F. Evolution of protein synthesis from an RNA world. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*, a003681. [CrossRef]

89. Van der Gulik, P.T.; Speijer, D. How amino acids and peptides shaped the RNA world. *Life (Basel)* **2015**, *5*, 230–246. [CrossRef]

90. Caskey, C.T.; Beaudet, A.L.; Scolnick, E.M.; Rosman, M. Hydrolysis of fMet-tRNA by peptidyl transferase. *Proc. Natl. Acad. Sci. USA* **1971**, *68*, 3163–3167. [CrossRef]

91.   Wolf, Y.I.; Aravind, L.; Grishin, N.V.; Koonin, E.V. Evolution of aminoacyl-tRNA synthetases–analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **1999**, *9*, 689–710.

92.   Yokoyama, T.; Shaikh, T.R.; Iwakura, N.; Kaji, H.; Kaji, A.; Agrawal, R.K. Structural insights into initial and intermediate steps of the ribosome-recycling process. *EMBO J.* **2012**, *31*, 1836–1846. [CrossRef]

93.   Gao, N.; Zavialov, A.V.; Li, W.; Sengupta, J.; Valle, M.; Gursky, R.P.; Ehrenberg, M.; Frank, J. Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies. *Mol. Cell* **2005**, *18*, 663–674. [CrossRef] [PubMed]

94.   Iyer, L.M.; Burroughs, A.M.; Aravind, L. The prokaryotic antecedents of the ubiquitin-signaling system and the early evolution of ubiquitin-like beta-grasp domains. *Genome Biol.* **2006**, *7*, R60. [CrossRef] [PubMed]

95.   Burroughs, A.M.; Iyer, L.M.; Aravind, L. Functional diversification of the RING finger and other binuclear treble clef domains in prokaryotes and the early evolution of the ubiquitin system. *Mol. Biosyst.* **2011**, *7*, 2261–2277. [CrossRef] [PubMed]

96.   Burroughs, A.M.; Iyer, L.M.; Aravind, L. Structure and evolution of ubiquitin and ubiquitin-related domains. *Methods Mol. Biol.* **2012**, *832*, 15–63. [CrossRef] [PubMed]

97.   Burroughs, A.M.; Iyer, L.M.; Aravind, L. The natural history of ubiquitin and ubiquitin-related domains. *Front. Biosci. (Landmark Ed.)* **2012**, *17*, 1433–1460. [CrossRef]

98.   Hurst, L.D.; Atlan, A.; Bengtsson, B.O. Genetic conflicts. *Q. Rev. Biol.* **1996**, *71*, 317–364. [CrossRef]

99.   Aravind, L.; Anantharaman, V.; Zhang, D.; de Souza, R.F.; Iyer, L.M. Gene flow and biological conflict systems in the origin and evolution of eukaryotes. *Front. Cell Infect. Microbiol.* **2012**, *2*, 89. [CrossRef]

100.   Kobayashi, K.; Ishitani, R.; Nureki, O. Recent structural studies on Dom34/aPelota and Hbs1/aEF1alpha: Important factors for solving general problems of ribosomal stall in translation. *Biophysics (Nagoya-shi)* **2013**, *9*, 131–140. [CrossRef]

101.   Goralski, T.D.P.; Kirimanjeswara, G.S.; Keiler, K.C. A New Mechanism for Ribosome Rescue Can Recruit RF1 or RF2 to Nonstop Ribosomes. *MBio* **2018**, *9*. [CrossRef]

102.   Stacy, R.A.; Aalen, R.B. Identification of sequence homology between the internal hydrophilic repeated motifs of group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of Bacillus subtilis. *Planta* **1998**, *206*, 476–478. [CrossRef]

103.   Maul, B.; Volker, U.; Riethdorf, S.; Engelmann, S.; Hecker, M. sigma B-dependent regulation of gsiB in response to multiple stimuli in Bacillus subtilis. *Mol. Genes Genet.* **1995**, *248*, 114–120. [CrossRef]

104.   Asteri, I.A.; Boutou, E.; Anastasiou, R.; Pot, B.; Vorgias, C.E.; Tsakalidou, E.; Papadimitriou, K. In silico evidence for the horizontal transfer of gsiB, a sigma(B)-regulated gene in gram-positive bacteria, to lactic acid bacteria. *Appl. Environ. Microbiol.* **2011**, *77*, 3526–3531. [CrossRef]

105.   Fernandez-Puentes, C.; Vazquez, D. Effects of some proteins that inactivate the eukaryotic ribosome. *FEBS Lett.* **1977**, *78*, 143–146. [CrossRef]

106.   Winther, K.S.; Brodersen, D.E.; Brown, A.K.; Gerdes, K. VapC20 of Mycobacterium tuberculosis cleaves the sarcin-ricin loop of 23S rRNA. *Nat. Commun.* **2013**, *4*, 2796. [CrossRef] [PubMed]

107.   Neubauer, C.; Gao, Y.G.; Andersen, K.R.; Dunham, C.M.; Kelley, A.C.; Hentschel, J.; Gerdes, K.; Ramakrishnan, V.; Brodersen, D.E. The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE. *Cell* **2009**, *139*, 1084–1095. [CrossRef]

108.   Polikanov, Y.S.; Blaha, G.M.; Steitz, T.A. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science* **2012**, *336*, 915–918. [CrossRef]

109.   Ueta, M.; Ohniwa, R.L.; Yoshida, H.; Maki, Y.; Wada, C.; Wada, A. Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *J. Biochem.* **2008**, *143*, 425–433. [CrossRef]

110.   Puri, P.; Eckhardt, T.H.; Franken, L.E.; Fusetti, F.; Stuart, M.C.; Boekema, E.J.; Kuipers, O.P.; Kok, J.; Poolman, B. Lactococcus lactis YfiA is necessary and sufficient for ribosome dimerization. *Mol. Microbiol.* **2014**, *91*, 394–407. [CrossRef]

111.   Ueta, M.; Wada, C.; Wada, A. Formation of 100S ribosomes in Staphylococcus aureus by the hibernation promoting factor homolog SaHPF. *Genes Cells* **2010**, *15*, 43–58. [CrossRef]

112.   Akiyama, T.; Williamson, K.S.; Schaefer, R.; Pratt, S.; Chang, C.B.; Franklin, M.J. Resuscitation of Pseudomonas aeruginosa from dormancy requires hibernation promoting factor (PA4463) for ribosome preservation. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3204–3209. [CrossRef]

113. Basu, A.; Yap, M.N. Ribosome hibernation factor promotes Staphylococcal survival and differentially represses translation. *Nucleic Acids Res.* **2016**, *44*, 4881–4893. [CrossRef]

114. Akanuma, G.; Kazo, Y.; Tagami, K.; Hiraoka, H.; Yano, K.; Suzuki, S.; Hanai, R.; Nanamiya, H.; Kato-Yamada, Y.; Kawamura, F. Ribosome dimerization is essential for the efficient regrowth of Bacillus subtilis. *Microbiology* **2016**, *162*, 448–458. [CrossRef]

115. Basu, A.; Yap, M.N. Disassembly of the Staphylococcus aureus hibernating 100S ribosome by an evolutionarily conserved GTPase. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8165–E8173. [CrossRef]

116. Jacob, Y.; Seif, E.; Paquet, P.O.; Lang, B.F. Loss of the mRNA-like region in mitochondrial tmRNAs of jakobids. *RNA* **2004**, *10*, 605–614. [CrossRef]

117. Hafez, M.; Burger, G.; Steinberg, S.V.; Lang, B.F. A second eukaryotic group with mitochondrion-encoded tmRNA: In silico identification and experimental confirmation. *RNA Biol* **2013**, *10*, 1117–1124. [CrossRef]

118. Keiler, K.C.; Shapiro, L.; Williams, K.P. tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in Caulobacter. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7778–7783. [CrossRef]

119. Stiller, J.W.; Schreiber, J.; Yue, J.; Guo, H.; Ding, Q.; Huang, J. The evolution of photosynthesis in chromist algae through serial endosymbioses. *Nat. Commun.* **2014**, *5*, 5764. [CrossRef]

120. Archibald, J.M. Genomic perspectives on the birth and spread of plastids. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10147–10153. [CrossRef]

121. Horn, M.; Harzenetter, M.D.; Linner, T.; Schmid, E.N.; Muller, K.D.; Michel, R.; Wagner, M. Members of the Cytophaga-Flavobacterium-Bacteroides phylum as intracellular bacteria of acanthamoebae: Proposal of 'Candidatus Amoebophilus asiaticus'. *Environ. Microbiol.* **2001**, *3*, 440–449. [CrossRef]

122. Burroughs, A.M.; Kaur, G.; Zhang, D.; Aravind, L. Novel clades of the HU/IHF superfamily point to unexpected roles in the eukaryotic centrosome, chromosome partitioning, and biologic conflicts. *Cell Cycle* **2017**, *16*, 1093–1103. [CrossRef]

123. Izawa, T.; Park, S.H.; Zhao, L.; Hartl, F.U.; Neupert, W. Cytosolic Protein Vms1 Links Ribosome Quality Control to Mitochondrial and Cellular Homeostasis. *Cell* **2017**, *171*, 890–903.e818. [CrossRef]

124. Tran, J.R.; Tomsic, L.R.; Brodsky, J.L. A Cdc48p-associated factor modulates endoplasmic reticulum-associated degradation, cell stress, and ubiquitinated protein homeostasis. *J. Biol. Chem.* **2011**, *286*, 5744–5755. [CrossRef] [PubMed]

125. Shen, P.S.; Park, J.; Qin, Y.; Li, X.; Parsawar, K.; Larson, M.H.; Cox, J.; Cheng, Y.; Lambowitz, A.M.; Weissman, J.S.; et al. Protein synthesis. Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science* **2015**, *347*, 75–78. [CrossRef]

126. Burroughs, A.M.; Aravind, L. A highly conserved family of domains related to the DNA-glycosylase fold helps predict multiple novel pathways for RNA modifications. *RNA Biol* **2014**, *11*, 360–372. [CrossRef]

127. Menninger, J.R. Accumulation of peptidyl tRNA is lethal to *Escherichia coli*. *J. Bacteriol.* **1979**, *137*, 694–696.

128. Garcia-Villegas, M.R.; De La Vega, F.M.; Galindo, J.M.; Segura, M.; Buckingham, R.H.; Guarneros, G. Peptidyl-tRNA hydrolase is involved in lambda inhibition of host protein synthesis. *EMBO J.* **1991**, *10*, 3549–3555. [CrossRef] [PubMed]

129. Menez, J.; Buckingham, R.H.; de Zamaroczy, M.; Campelli, C.K. Peptidyl-tRNA hydrolase in Bacillus subtilis, encoded by spoVC, is essential to vegetative growth, whereas the homologous enzyme in Saccharomyces cerevisiae is dispensable. *Mol. Microbiol.* **2002**, *45*, 123–129. [CrossRef]

130. Heurgue-Hamard, V.; Karimi, R.; Mora, L.; MacDougall, J.; Leboeuf, C.; Grentzmann, G.; Ehrenberg, M.; Buckingham, R.H. Ribosome release factor RF4 and termination factor RF3 are involved in dissociation of peptidyl-tRNA from the ribosome. *EMBO J.* **1998**, *17*, 808–816. [CrossRef] [PubMed]

131. Karimi, R.; Pavlov, M.Y.; Heurgue-Hamard, V.; Buckingham, R.H.; Ehrenberg, M. Initiation factors IF1 and IF2 synergistically remove peptidyl-tRNAs with short polypeptides from the P-site of translating Escherichia coli ribosomes. *J. Mol. Biol.* **1998**, *281*, 241–252. [CrossRef] [PubMed]