



Error-prone protein synthesis in parasites with the smallest eukaryotic genome

Sergey V. Melnikov^a, Keith D. Rivera^b, Denis Ostapenko^a, Arthur Makarenko^b, Neil D. Sanscrainte^c, James J. Becnel^c, Mark J. Solomon^a, Catherine Texier^d, Darryl J. Pappin^b, and Dieter Söll^{a,e,1}

^aDepartment of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520; ^bCold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; ^cCenter for Medical, Agricultural, and Veterinary Entomology, Agricultural Research Service, US Department of Agriculture, Gainesville, FL 32608; ^dLaboratoire Microorganismes, Génome et Environnement, Université Clermont Auvergne, CNRS, F-63000 Clermont-Ferrand, France; and ^eDepartment of Chemistry, Yale University, New Haven, CT 06520

Contributed by Dieter Söll, May 23, 2018 (sent for review February 21, 2018; reviewed by Michael Ibba and Xiang-Lei Yang)

Microsporidia are parasitic fungi-like organisms that invade the interior of living cells and cause chronic disorders in a broad range of animals, including humans. These pathogens have the tiniest known genomes among eukaryotic species, for which they serve as a model for exploring the phenomenon of genome reduction in obligate intracellular parasites. Here we report a case study to show an apparent effect of overall genome reduction on the primary structure and activity of aminoacyl-tRNA synthetases, indispensable cellular proteins required for protein synthesis. We find that most microsporidian synthetases lack regulatory and eukaryote-specific appended domains and have a high degree of sequence variability in tRNA-binding and catalytic domains. In one synthetase, LeuRS, an apparent sequence degeneration annihilates the editing domain, a catalytic center responsible for the accurate selection of leucine for protein synthesis. Unlike accurate LeuRS synthetases from other eukaryotic species, microsporidian LeuRS is error-prone: apart from leucine, it occasionally uses its near-cognate substrates, such as norvaline, isoleucine, valine, and methionine. Mass spectrometry analysis of the microsporidium *Vavraia culicis* proteome reveals that nearly 6% of leucine residues are erroneously replaced by other amino acids. This remarkably high frequency of mistranslation is not limited to leucine codons and appears to be a general property of protein synthesis in microsporidian parasites. Taken together, our findings reveal that the microsporidian protein synthesis machinery is editing-deficient, and that the proteome of microsporidian parasites is more diverse than would be anticipated based on their genome sequences.

levels of genome reduction. In eukaryotes, the extremely small genomes are found in Microsporidia, a group of unicellular fungi-related organisms, which might be the most ancient parasites among eukaryotic species (8–12). Many microsporidian species have approximately 1,800 genes, nearly threefold fewer genes than in free-living single-cell fungi (13, 14). However, despite having such tiny genomes, at least 17 microsporidian species can efficiently infect humans and cause a wide range of infections, primarily of the gastrointestinal tract, which are especially frequent in immune-compromised individuals (13, 14). For instance, microsporidian infections are found in approximately 15% of HIV-infected humans, where it causes chronic diarrhea, abdominal pain, sinusitis, and other intestinal, lung, kidney, brain, sinus, muscle, or eye diseases (15). Also, Microsporidia infect other animals, including industrially valuable monkfish and honeybees (16, 17). These pathogenic properties, along with their extremely compact and rapidly evolving genomes, make Microsporidia a good model for studying how genome erosion affects the activities of essential proteins and nucleic acids and alters their sensitivity to drugs (13).

In the present study, we explored the protein synthesis machinery of microsporidian parasites. We focused on aminoacyl-tRNA synthetases, ubiquitous proteins that are required for protein synthesis in every living cell (18, 19). Aminoacyl-tRNA synthetases participate in protein synthesis by conjugating individual amino acids to tRNA molecules. The synthetases act with remarkable precision by accurately selecting one out of several dozen types of amino acids in the cell (19–21). Here we show

Muller's ratchet | genome erosion | Microsporidia | aminoacyl-tRNA synthetases | mistranslation

Our planet might be inhabited by more than 1 million distinct species that proliferate exclusively in the interior of another living cell (1). These organisms are known as obligate intracellular parasites.

Unlike free-living species, the obligate parasites are fully dependent on their host for survival. Typically, they live in small populations in the host cytosol, divide asexually, and undergo frequent population bottlenecks during transmission from one host cell to another. This lifestyle weakens the stringency of natural selection and causes an evolutionary tendency that is rarely observed in free-living species but is common in the overwhelming majority of intracellular parasites (2). This tendency, known as Muller's ratchet, is a process in which genomes of intracellular parasites accumulate mildly deleterious mutations (3–5). In the long term, this continuous deterioration of genomes impairs the activity of individual proteins and nucleic acids, causes gene loss, and results in genome erosion and an irreversible decline in fitness (6). Because this erosive evolution is common for most intracellular parasites, it is considered a potential basis for therapies against a broad range of parasite infections (7).

Some intracellular parasites represent particularly valuable models for exploring genetic erosion because of the extreme

Significance

Microsporidia are poorly treatable eukaryotic pathogens that threaten human health and industrially valuable insects and fish, yet we are only beginning to understand the complex biology of these emerging pathogens. Here we combine bioinformatics, biochemistry, and mass spectrometry analyses to show that Microsporidia carry an error-prone machinery of protein synthesis, and that their protein synthesis is accompanied by a remarkable number of translation errors. This finding reveals a previously unknown aspect of protein synthesis in these emerging parasites and creates a potential opportunity to use the defective protein synthesis machinery as a therapeutic target to treat microsporidia infections.

Author contributions: S.V.M., M.J.S., D.J.P., and D.S. designed research; S.V.M., K.D.R., D.O., A.M., and C.T. performed research; N.D.S., J.J.B., and C.T. contributed new reagents/analytic tools; S.V.M., K.D.R., and D.S. analyzed data; and S.V.M. and D.S. wrote the paper.

Reviewers: M.L., The Ohio State University; and X.-L.Y., The Scripps Research Institute.

The authors declare no conflict of interest.

Published under the [PNAS license](#).

¹To whom correspondence should be addressed. Email: dieter.soll@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1803208115/-DCSupplemental.

Published online June 18, 2018.

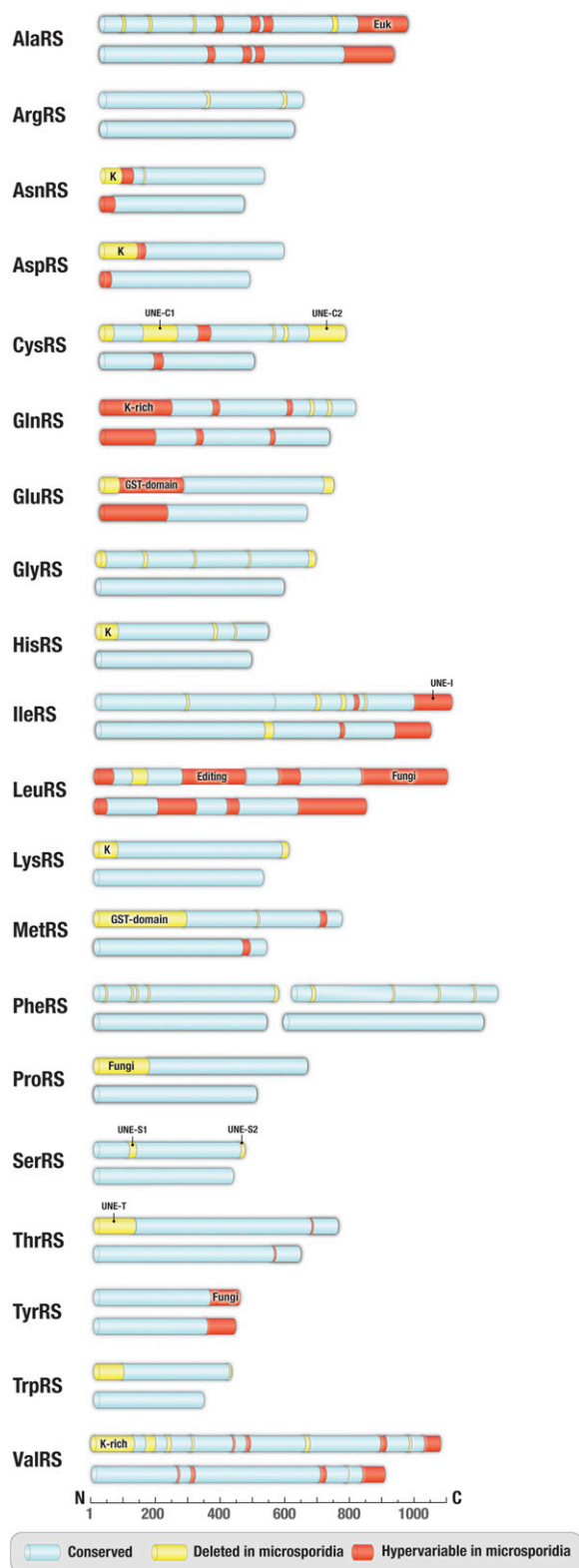


Fig. 1. Apparent reductive evolution of microsporidian aminoacyl-tRNA synthetases. Schematic comparison of primary structures of aminoacyl-tRNA synthetases from microsporidian parasites (exemplified by *V. culicis*) and nonparasitic fungi (exemplified by *S. cerevisiae*). Protein segments highlighted in blue indicate conserved regions (sequence similarity >50%), protein segments present in yeast but deleted in microsporidia are shown in yellow, and segments with highly variable sequences in microsporidia (sequence similarity <35% between *V. culicis* and *T. hominis*) but are conserved in other eukaryotes are highlighted in red. Labels indicate appended do-

that one microsporidian synthetase, leucyl-tRNA synthetase (LeuRS), has a defect that impairs accurate selection of leucine for protein synthesis. Our *in vitro* assays show that, apart from using leucine, microsporidian LeuRS also occasionally mischarges tRNAs with other amino acids, such as valine, methionine, isoleucine, and norvaline. Using mass spectrometry, we found that microsporidian protein synthesis is accompanied by numerous errors in which amino acid residues in microsporidian proteins are replaced by other amino acids, causing an unanticipated diversity of the microsporidian proteome.

Results

Reductive Evolution of Microsporidian Aminoacyl-tRNA Synthetases.

We first asked whether the massive erosion of microsporidian genomes affects the structure of aminoacyl-tRNA synthetases. We anticipated that microsporidian synthetases would be smaller than their counterparts from free-living eukaryotes because on average microsporidian proteins (e.g., in *Encephalitozoon cuniculi*) are approximately 7% shorter than homologous proteins in free-living fungi, such as *Saccharomyces cerevisiae* (9).

To compare microsporidian synthetases with their homologs from other eukaryotes, we analyzed available genome sequences from 19 microsporidian species (*SI Appendix, Table S1*). Each of these genomes was shown to harbor a full set of 20 different aminoacyl-tRNA synthetases corresponding to the 20 proteinogenic amino acids, but to lack the genes coding for mitochondrial aminoacyl-tRNA synthetases due to degeneration of mitochondria in microsporidian cells (9, 16, 22–36). Moreover, microsporidian genomes lack duplicated synthetase genes or genes coding for synthetase fragments that are present in other eukaryotes (9, 16, 23–38).

We found that microsporidian synthetases are markedly smaller than their homologs in other eukaryotes. On average, the synthetases are roughly 12% shorter in microsporidians than in nonpathogenic fungi, such as *S. cerevisiae* (Fig. 1 and *SI Appendix, Table S2*). This size reduction is accomplished primarily by approximately 40- to 300-aa-long deletions of protein segments known as appended domains. These domains are missing in prokaryotes but are abundant in eukaryotes, where they regulate or enhance protein synthesis or have translation-unrelated biological activities (37, 39, 40). For example, one of the appended domains, a GST-like domain in MetRS and GluRS, triggers aminoacyl-tRNA synthetase assembly into supramolecular complexes to enhance the enzymatic activities of individual aminoacyl-tRNA synthetases (41). Other appended domains, such as lysine-rich segments in ArgRS, AsnRS, AspRS, GlnRS, HisRS, LysRS, or ValRS, enhance tRNA binding and the overall rate of aminoacyl-tRNA synthesis (41–44). Some other appended domains, such as UNE and others, have been less well studied, although they are frequently found across eukaryotes (45). All of these appended domains are missing or altered in microsporidian synthetases (Fig. 1). Apart from numerous deletions, microsporidian synthetases have a high degree of sequence variability, including variations in tRNA-binding interface (*SI Appendix, Fig. S1*).

Overall, our comparisons reveal that microsporidian synthetases are reduced to their very enzymatic cores by stripping down regulatory domains and protein segments that are unrelated to the essential function of protein synthesis. However, in one case, the synthetase degeneration propagated to a catalytic center, which is essential for the quality control of protein synthesis (Fig. 2*A*).

Atypical Variability of the Editing Domain in Microsporidian Leucyl-tRNA Synthetases.

Nearly one-third of aminoacyl-tRNA synthetases carry two catalytic centers: the aminoacylation site, which selects cognate amino acids and attaches them to tRNA, and the

main, such as eukaryote-specific (Euk), fungi-specific (Fungi), GST-domain, UNE-domain, and lysine-rich (K-rich) domains.

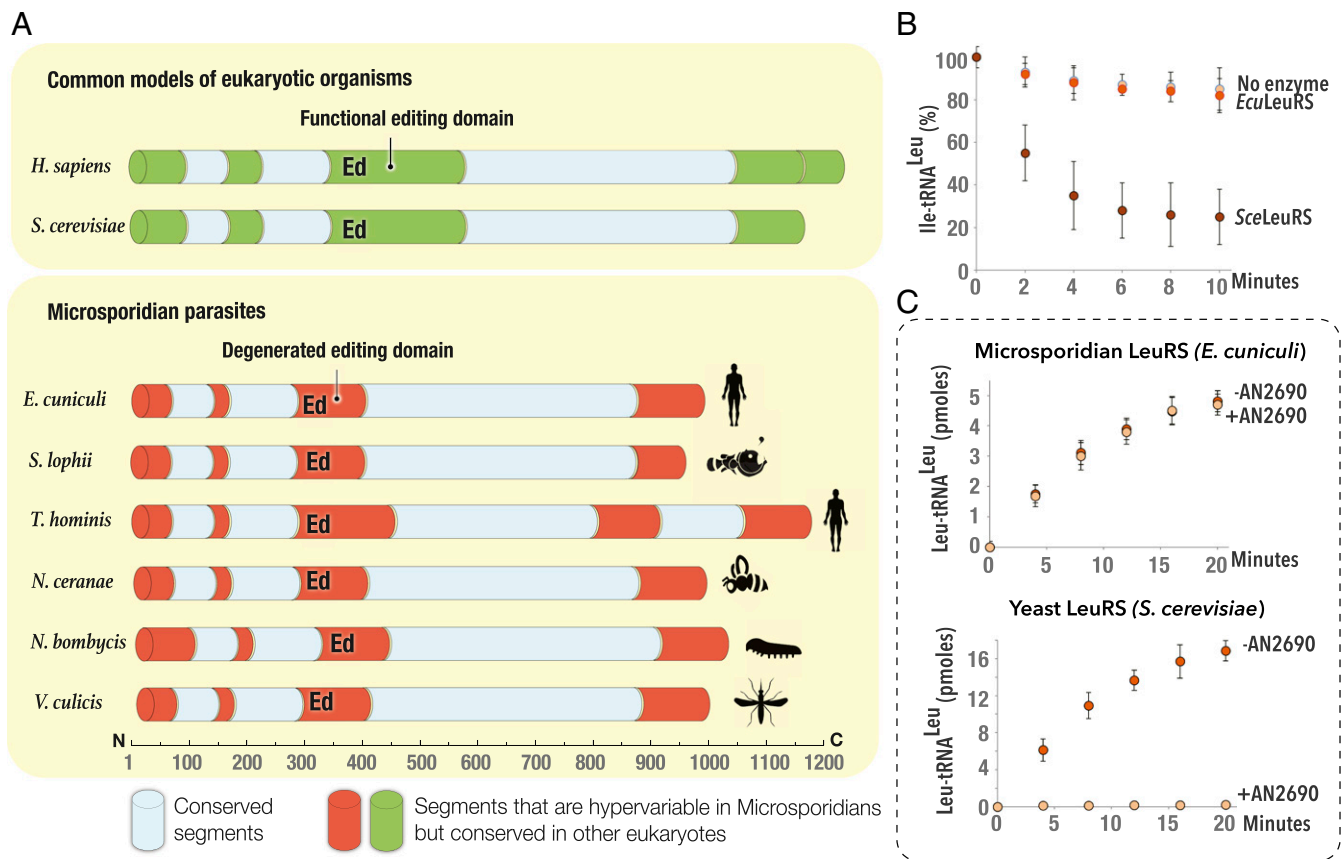


Fig. 2. Microsporidian LeuRS synthetase lacks a proofreading domain and is naturally resistant to a drug from the benzoxaborole family. (A) Schematic structures of aminoacyl-tRNA synthetase LeuRS from different species, including several species of microsporidian parasites. The structures are color-coded by conservation: conserved segments are shown in blue (>50% sequence similarity between *E. cuniculi* and *S. cerevisiae*), segments degenerated in microsporidia but conserved in other eukaryotes are in green, and hypervariable segments in microsporidian LeuRS are in red (sequence similarity <35% between *E. cuniculi* and *T. hominis*). The scale indicates LeuRS length (in amino acids), and the silhouette symbols indicate a best-studied host (human, monkfish, honey bee, and silkworm) for each microsporidian parasite. The figure illustrates that, unlike LeuRS from other eukaryotic species, microsporidian LeuRS synthetases have a massively truncated editing domain, suggesting that microsporidian LeuRS synthetases lack the editing activity. (B) Aminoacyl-tRNA hydrolysis assay to measure hydrolysis of the misaminoacylated substrate, Ile-tRNA^{Leu}, by microsporidian (*E. cuniculi*) or yeast (*S. cerevisiae*) LeuRS. The figure shows that unlike yeast LeuRS (used as a positive control), LeuRS from *E. cuniculi* is not capable of degrading the Ile-tRNA^{Leu}, illustrating a lack of the editing activity. (C) Kinetics of Leu-tRNA^{Leu} synthesis by microsporidian (*E. cuniculi*) and yeast (*S. cerevisiae*) LeuRS in the absence and in the presence of AN2690, a member of the benzoxaborole family that targets the editing domain of LeuRS synthetase. While both *S. cerevisiae* and *E. cuniculi* LeuRS synthetases are capable of producing Leu-tRNA^{Leu} in the absence of AN2690, only microsporidian LeuRS synthetase remains active in the presence of the drug, illustrating the natural resistance of *E. cuniculi* LeuRS to benzoxaboroles.

editing site used to proofread amino acid selection (46–54). For instance, the amino acylation site of LeuRS synthetase occasionally mischarges tRNA^{Leu} with norvaline, valine, methionine, and other amino acids that have chemical structures similar to leucine (55). However, mischarged tRNA^{Leu} is rapidly hydrolyzed by the editing domain, thereby preventing errors during protein synthesis (56, 57). While a typical LeuRS editing domain is conserved in virtually all eukaryotes and comprises approximately 240 residues, we found that it is reduced to approximately 100–140 residues in microsporidian species (Fig. 2A). Furthermore, these 100–140 residues show only approximately 35% sequence similarity even between microsporidian species (SI Appendix, Figs. S2 and S3). This sequence reduction and hypervariability suggest that microsporidian LeuRS lacks a functional editing domain.

Microsporidian LeuRS Synthetase Is a Proofreading-Deficient Enzyme.

We next asked whether microsporidian LeuRS synthetase is incapable of proofreading. To test this, we used LeuRS from *E. cuniculi* and measured its ability to hydrolyze an erroneously charged substrate, Ile-tRNA^{Leu}. We found that *E. cuniculi* LeuRS was incapable of hydrolyzing Ile-tRNA^{Leu}, while LeuRS from *S.*

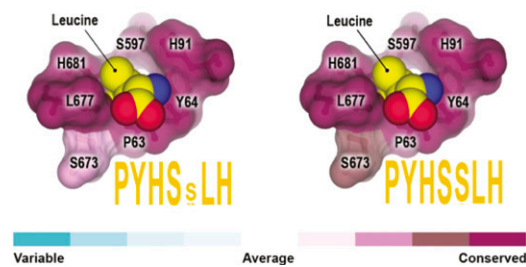
cerevisiae (used as a positive control) rapidly hydrolyzed Ile-tRNA^{Leu} (Fig. 2B). These data illustrate that microsporidian LeuRS is indeed an editing-deficient enzyme.

We next asked whether the editing deficiency alters microsporidian LeuRS affinity to drugs. The LeuRS editing domain is a target of benzoxaboroles, drugs that inhibit cell growth by forming covalent bonds with the tRNA^{Leu} in the LeuRS editing domain, thereby preventing LeuRS turnover and consequently arresting protein synthesis (58–60). We used an antifungal benzoxaborole, AN2690, and tested its ability to inhibit LeuRS-dependent synthesis of Leu-tRNA^{Leu} in vitro. We found that AN2690 (5 μg/mL) fully inhibited LeuRS synthetase from *S. cerevisiae*, but had no effect on Leu-tRNA^{Leu} synthesis by microsporidian LeuRS from *E. cuniculi* (Fig. 2C). These data show that microsporidian LeuRS is naturally resistant to a benzoxaborole, suggesting benzoxaborole tolerance by microsporidian parasites.

Microsporidian LeuRS Synthetase Is Promiscuous Toward Near-Cognate Substrates.

We next asked whether the proofreading deficiency of microsporidian LeuRS is compensated for by the increased accuracy of the LeuRS aminoacylation site. To test this, we first compared

A Conservation across Microsporidia Conservation across eukaryotes



B

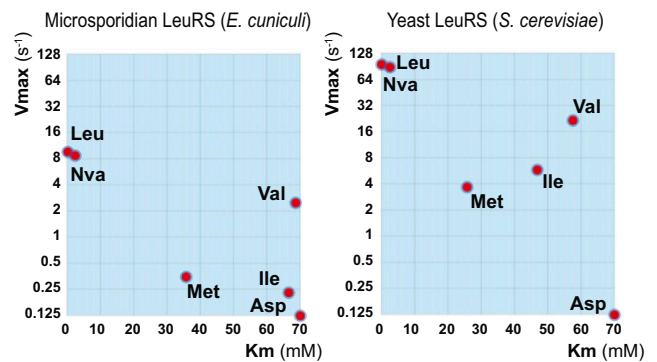


Fig. 3. Microsporidian LeuRS is promiscuous toward near-cognate substrates. (A) The panel shows a fragment of the LeuRS synthetase crystal structure (Protein Data Bank ID code 1WK8) to illustrate high conservation of the leucine-binding pocket in the LeuRS aminoacylation site across eukaryotic species. Residues that directly bind leucine molecules are numbered according to *Homo sapiens* LeuRS numbering and are colored by conservation. Motif sequences (in orange) represent consensus sequences of the leucine-binding pocket. (B) Diagrams summarizing steady-state parameters for amino acid activation by microsporidian (*E. cuniculi*) and yeast (*S. cerevisiae*) LeuRS. This panel illustrates that microsporidian LeuRS can promiscuously activate near-cognate substrates norvaline (Nva), methionine (Met), valine (Val), and isoleucine (Ile), but not the noncognate substrate aspartate (Asp; used as a negative control).

sequence conservation in the LeuRS aminoacylation site. We found that residues composing the active site are conserved across virtually all eukaryotic species, including microsporidian species (Fig. 3A). This suggests that the aminoacylation site of microsporidian LeuRS has a similar structure as in the other eukaryotes.

To experimentally estimate the accuracy of the microsporidian LeuRS aminoacylation site, we measured rates of amino acid activation by LeuRS from *E. cuniculi* and *S. cerevisiae*. For this, we used the cognate LeuRS substrate leucine; near-cognate substrates methionine, valine, isoleucine, and norvaline; and the noncognate substrate aspartate (Fig. 3B). We found that compared with *S. cerevisiae* LeuRS, *E. cuniculi* LeuRS acts at an approximately 10-fold slower maximal rate and with a comparably low accuracy (Fig. 3B and *SI Appendix*, Fig. S4 and Table S3). Both enzymes could activate not only leucine, but also near-cognate substrates, suggesting that the aminoacylation site of microsporidian LeuRS is not likely to have a compensatory increase in accuracy compared with other eukaryotes.

Error-Prone Protein Synthesis in Microsporidian Parasites. We next asked whether the proofreading-deficient LeuRS synthetase deteriorates the fidelity of protein synthesis in microsporidian cells. Based on our biochemical data and studies by others (61, 62), we anticipated that the LeuRS editing deficiency would cause mistranslation of leucine codons. For instance, inactivation of the editing domains in aminoacyl-tRNA synthetases is thought to increase translation errors from one error per roughly 3,000–10,000 codons to one error per approximately 200 codons (62).

To test whether microsporidian parasites indeed mistranslate leucine codons, we analyzed protein extracts from the microsporidium *Vavraia culicis* by mass spectrometry. *V. culicis* is a mosquito pathogen that carries a nearly identical deletion of the LeuRS editing domain as *E. cuniculi* (Fig. 2A). To estimate the rate of leucine mistranslation, we used a similar approach to that designed for *Mycoplasma* (62). We measured the total number of peptides in which leucine was erroneously replaced by valine or methionine, and divided this number by the total number of leucine residues in the experimentally observed peptides. In total, 0.71% of leucine residues were replaced by valine and 0.29% were replaced by methionine, corresponding to an apparent mistranslation rate of roughly 1:150–1:300 (Fig. 4 and *SI Appendix*, Dataset S1).

However, a closer look at the mass spectrometry data reveals that leucine-to-valine and leucine-to-methionine are not the only errors in the *V. culicis* proteome (Fig. 5). In fact, we detected all possible types of leucine codon mistranslation (Fig. 5). In total,

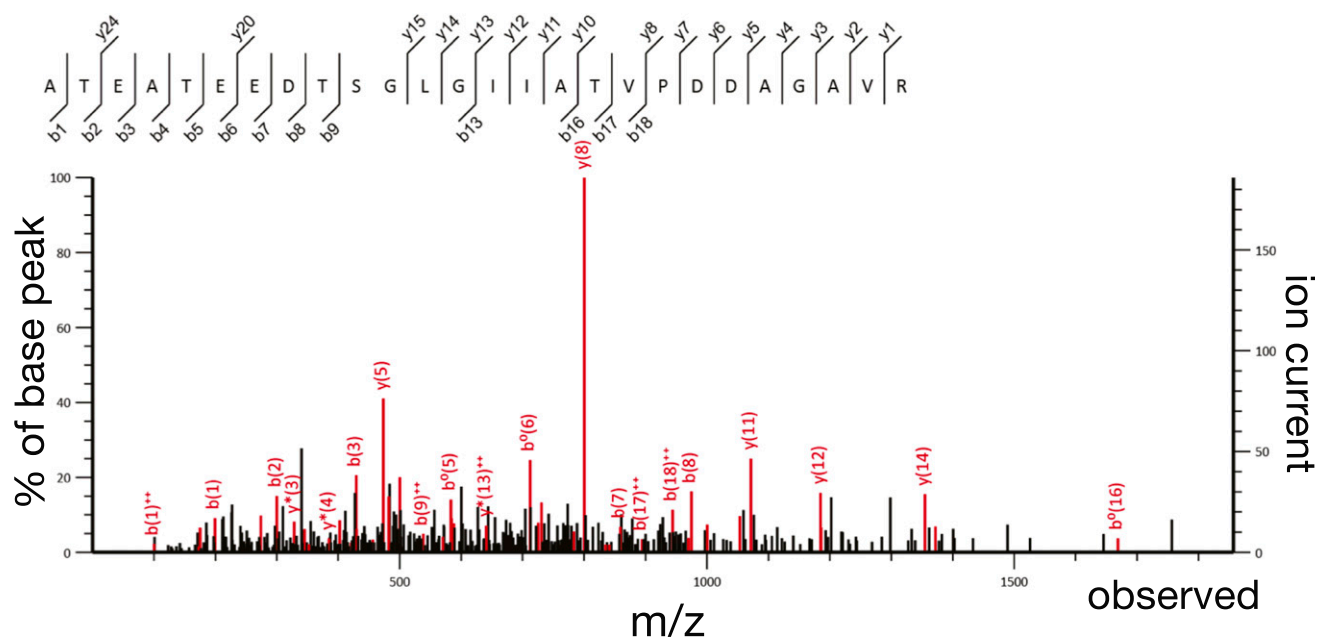
~5.9% of leucine residues in microsporidian proteins were replaced by other amino acids (Fig. 5 and *SI Appendix*, Dataset S1). Moreover, even though leucine-to-valine was the most frequent error in protein sequences, comparable error frequencies were observed for leucine substitution by other amino acids as well (Fig. 5 and *SI Appendix*, Datasets S1 and S2). Furthermore, we found that apparent translational errors are not limited to leucine codons. We calculated the translation accuracy for valine codons and found similar levels of mistranslation as for leucine codons, even though the editing domain of microsporidian valyl-tRNA synthetases was intact (Fig. 5 and *SI Appendix*, Fig. S5). This observation indicated that the actual mechanism of error generation during protein synthesis is far more complex than error-prone tRNA amino acylation by LeuRS synthetase. Thus, although mass spectrometry did not reveal the mechanism(s) of error generation in protein sequences, it uncovered remarkably high levels of mistranslated proteins in Microsporidia.

Discussion

In this work, we explore an apparent effect of the extreme compaction of parasitic genomes on the structure and activity of indispensable components of a living cell, aminoacyl-tRNA synthetases. We find that in microsporidian parasites, the synthetases lack the appended domains and have a primary structure reduced to the very catalytic cores. In one protein, LeuRS synthetases, an apparent sequence degeneration propagates to the editing domain and deteriorates the accuracy of leucine selection for protein synthesis by this enzyme. Overall, our findings show that microsporidian parasites are the only known eukaryotes that lack the editing domain in the cytosolic and essential factor of protein synthesis, LeuRS, and that protein synthesis in microsporidian parasites might be accompanied by a remarkably high rate of random errors.

Quantitative analysis of mistranslation in the microsporidium *V. culicis* suggests that the mechanism of the mistranslation is complex. This is not surprising, because protein synthesis is governed by more than 100 RNA and protein factors, including tRNA molecules, tRNA-modifying enzymes, ribosomal proteins, ribosomal RNA, and elongation factors of translation. It is possible that some of these factors act with less accuracy in microsporidian parasites. In fact, microsporidians carry remarkably small ribosomes, approximately 1 MDa smaller than ribosomes from nonparasitic fungi, with microsporidian 18S rRNA harboring only roughly 1,200 nucleotides, compared with 1,800 nucleotides in other eukaryotes (8, 63–65). These massive deletions are accompanied by high sequence variability in rRNA

A Fragment of an accurately translated protein



B Fragment of a mistranslated protein

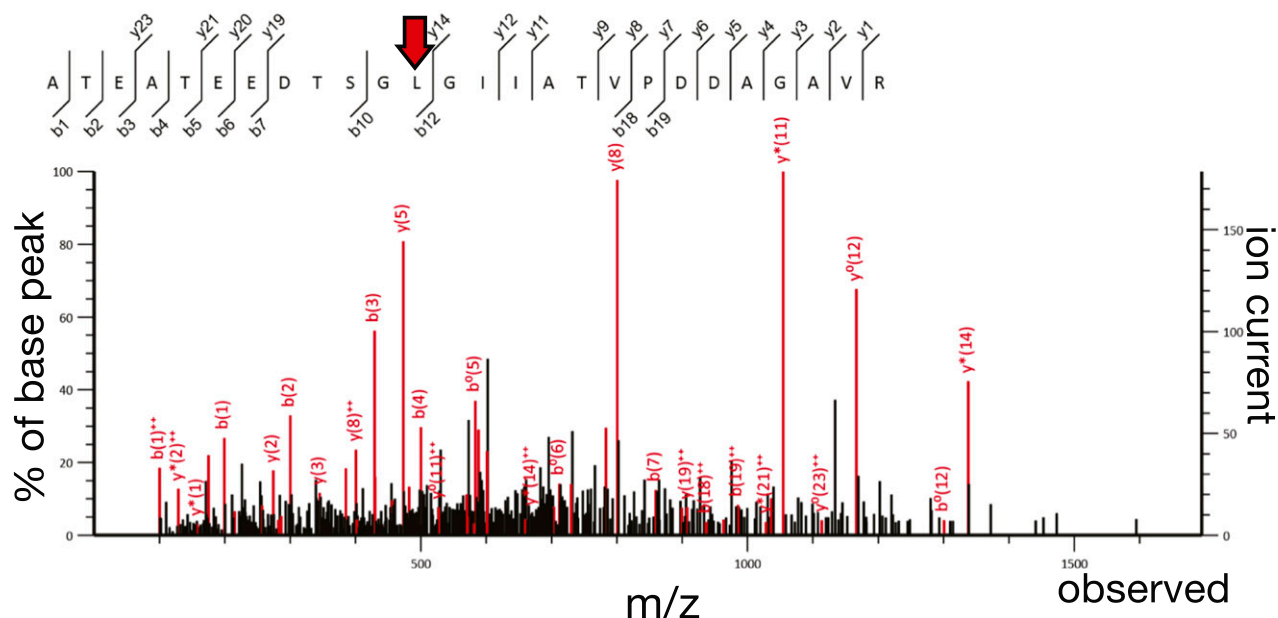


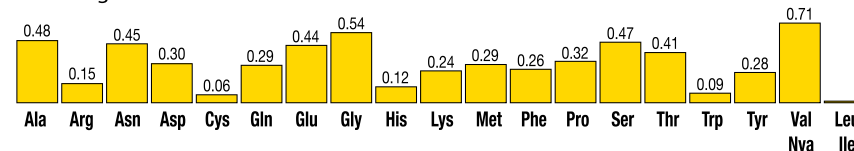
Fig. 4. Error-prone protein synthesis in microsporidian parasites observed by quantitative mass spectrometry of *V. culicis* protein extracts. Mass spectra of *V. culicis* proteome illustrating mistranslation in microsporidian parasites. Both mass spectra correspond to the same peptide that belongs to a protein component of the polar tube (a microsporidia-specific invasion organelle that allows microsporidian parasites to penetrate into the host cell). Like several hundred other peptides in *V. culicis* proteome, this protein fragment was observed in two distinct forms, the accurately translated form, in which the protein sequence corresponds to the sequence of genomic DNA (A), and a mistranslated form, in which one of the leucine residues (red arrow) is replaced with valine/norvaline (B).

and ribosomal proteins (65, 66), suggesting that microsporidian ribosomes might be less accurate in selecting cognate aminoacyl-tRNAs.

Another potential source of translational errors might be related to erroneous binding of aminoacyl-tRNA synthetases to noncognate tRNAs. In particular, this scenario could explain

how microsporidian parasites mistranslate leucine codons with aspartate residues even though LeuRS synthetases do not activate aspartate and thus cannot produce Asp-tRNA^{Leu}. However, as shown in (Fig. 1), microsporidian AspRS lacks the N-terminal lysine-rich segment, which may compromise accurate tRNA selection and result in erroneous tRNA^{Leu} recognition.

Percentage of leucine codons that are mistranslated as other amino acids



Percentage of valine codons that are mistranslated as other amino acids

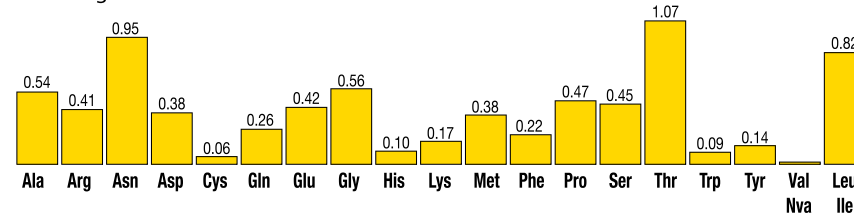


Fig. 5. Error-prone protein synthesis in microsporidian parasites observed by quantitative mass spectrometry of *V. culicis* protein extracts. Percentage of leucine codons and valine codons that are mistranslated as other amino acids (calculated as the number of mistranslated codons divided by the total number of leucine or valine residues in the experimentally detected peptides).

Moreover, our data cannot exclude the possibility that some of the errors in protein sequence might originate from error-prone transcription. The future studies combining will elucidate the nature of this spectacular increase in the error rates.

Of note, the high levels of mistranslation in microsporidia suggest another important role for the chaperone system in these and possibly other obligate intracellular parasites. Previously obligate intracellular parasites and symbionts have been shown to highly express genes coding for chaperones to buffer the effect of deleterious mutations on protein folding (67). Our observation of high mistranslation in microsporidia suggests that the high levels of chaperones could also buffer potential deterioration of protein folding by mistranslation.

Although the editing deficiency of LeuRS synthetase does not appear to be the dominant cause of mistranslation in Microsporidia, it represents a remarkable structural defect. Microsporidia now join a very limited group of species that carry editing deficiencies in otherwise highly conserved aminoacyl-tRNA synthetases. For years, it was believed that in the wild environment, a living cell is not likely to carry editing-deficient synthetases, because editing deficiency in a number of synthetases has been shown to cause amino acid toxicity, cell death, and neurologic disease in mammals (68–73). In 2005, the editing-deficient synthetase PheRS was reported in mitochondria from *S. cerevisiae* (74). Then the editing-deficient synthetases were found in mitochondria from other eukaryotic species (75–77). In 2011, the editing-deficient synthetases PheRS, LeuRS, and ThrRS were found in the cytosol of *Mycoplasma* species, bacteria that carry extremely compact genomes (<500 protein-coding genes) (62). Our finding of the editing-deficient synthetase in Microsporidia—another host-restricted species with a highly reduced genome—suggests that the editing deficiency might be a common feature of organisms with exceptionally small genomes.

It is interesting to note that, apart from LeuRS, other microsporidian aminoacyl-tRNA synthetases have seemingly intact editing domains. This is especially evident if we compare LeuRS, IleRS, and ValRS synthetases. These synthetases are thought to have common evolutionary origin, they activate small hydrophobic amino acids with similar chemical structures, and their editing domains share common architecture and belong to the ValRS/IleRS/LeuRS editing domain family (19). However, while LeuRS editing domain undergoes severe truncations and apparent degeneration in microsporidian species, IleRS and ValRS preserve both the length and sequence of their editing domains (*SI Appendix, Fig. S5*).

The next question to answer will be why did Microsporidia lose the editing domain in the essential factor of protein synthesis,

along with other deletions and sequence alterations in factors of protein synthesis? Is this loss a friend that improves parasite survival, or a foe that has no benefit and simply stems from unavoidable genome decay triggered by the host-restricted lifestyle?

Previous studies suggest that both scenarios are possible. On one hand, there are a handful of examples suggesting that Muller's ratchet deteriorates the function of individual proteins with no evident benefits for cell fitness. For instance, globular domains of proteins from the host-restricted organisms are generally less stable and more aggregation-prone than those from free-living species, which makes host-restricted species more dependent on protein chaperones (2, 67, 78). It is possible that the editing inactivation, along with other alterations of microsporidian translation machinery, has no benefits and stems from host-restricted lifestyle that favors accumulation of deleterious mutations.

On the other hand, some organisms use inaccurate protein synthesis as a meaningful adaptive strategy (79–82). For instance, pathogenic yeast *Candida albicans* translate the very same CUG codons as serine (95–99.5%) and leucine (0.5–5%) (83, 84). This ambiguous translation randomizes protein sequences and helps *C. albicans* trick the host immune system. We cannot exclude the possibility that inaccurate protein synthesis might fulfill a similar adaptive function in microsporidian parasites.

Materials and Methods

Gene Cloning, Protein Expression, and Purification. A codon-optimized gene coding for *E. cuniculi* LeuRS (UniProt ID Q8SR58) was purchased from Sigma-Aldrich and cloned into the pET15b vector between the sequence 5'-AGGAGATATACC-3' and the sequence 5'-CTAGCATAACCCC-3' using a TaKaRa cloning kit. *E. cuniculi* LeuRS was expressed in *Escherichia coli* BL31 (DE3) strain at 16 °C and purified by using Ni-affinity purification, followed by ammonium sulfate fractionation and size-exclusion chromatography on a Superdex Increase 200 10/300 GL column (Pharmacia Biotech) in buffer containing 150 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 10 mM KF, and 1 mM DTT. The protein was concentrated to 10 mg/mL using Amicon Ultracel centrifugal filters (molecular weight cutoff, 30 kDa) and used in biochemical assays. Yeast *S. cerevisiae* LeuRS was purified from the BY4741 yeast strain carrying C-terminally TAP-tagged CDC60(YPL160V) (*S. cerevisiae* LeuRS) (Dharmacon). To purify Cdc60-TAP (*S. cerevisiae* LeuRS-TAP), yeast cells were grown in YPD media to OD₆₀₀ ≈ 1. Cells from 500 mL of culture were washed with ice-cold TBS (pH 8.0), resuspended in 3 mL of Binding buffer (25 mM Hepes pH 8.0; 150 mM NaCl; 0.1% Tween; 0.5 mM DTT; 10 mg/mL of each leupeptin, chymostatin, and pepstatin; and 10% glycerol), and lysed by shaking with 2.4 g of glass beads (0.5 mm diameter; Sigma-Aldrich) in a bead beater (Biospec Products). Glass beads and cell debris were removed by centrifugation at 18,400 × g for 5 min. The supernatant was clarified by centrifugation at 30,000 rpm in a TL100.2 rotor (Beckman) for 10 min at 4 °C. Cdc60-TAP was purified from yeast extracts by incubation with 0.5 mL bed volume IgG-Sepharose (GE Healthcare). Bound proteins were eluted from the beads by cleavage with 6His-TEV

protease. The eluate was incubated with 1 mL of Ni-NTA resin to remove 6His-TEV protease and was subsequently separated by size-exclusion chromatography (Superdex Increase 200 10/300 GL column; Pharmacia Biotech) in buffer containing 150 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 10 mM KF, and 1 mM DTT. The protein was purified to approximately 1 mg/mL using Amicon Ultracel centrifugal filters (molecular weight cutoff 50 kDa) for use in biochemical assays. The enzyme concentrations were calculated based on the OD₂₈₀ measurements and predicted extinction coefficients using the ProtParam prediction tool (<https://web.expasy.org/protparam/>).

Active Site Titration to Determine *E. cuculii* LeuRS Concentration. To estimate the active concentration of microsporidian LeuRS, we monitored the initial burst of ATP consumption. The reaction was carried out at 37 °C, in 50 μ L of buffer containing 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 100 mM KCl, 2.5 mM DTT, 10 μ M ATP, γ -[³²P]-ATP (1,000 cpm/ μ L final), 0.2 U/mL yeast PPIase, 50 mM L-Leu, and 5 μ M *E. cuculii* LeuRS synthetase (added at the last moment). At each time point, ranging from 0 to 60 min, 5- μ L aliquots were taken from the reaction mixture and mixed with 50 μ L of 7% perchloric acid to quench the reaction. At the end of the experiment, 1- μ L aliquots from each of the quenched reaction solutions were applied to a TLC plate (Whatman) that had been presoaked in water and dried before use. The chromatography was run at room temperature using 0.75 M KH₂PO₄ (pH 3.5) and 1 M urea solution as a running buffer. The radioactive spots were detected and quantified by phosphorimaging.

Measuring the Editing Activity of LeuRS Synthetases. To measure the editing activity, we first produced misacylated Ile-tRNA^{Leu}. For this purpose, we produced a transcript of *E. cuculii* tRNA^{Leu-TAA} (gene ID tRNA-Leu-TAA-1-1 from *E. cuculii*; gtrnadb.ucsc.edu/GTrnAdb2/) using in vitro transcription with T7 RNA polymerase (HiScribe T7 Kit; New England Biolabs). The tRNA^{Leu} was aminoacylated at a concentration of 2 μ M using 10 mM L-isoleucine (99%, W527602; Sigma-Aldrich), 7 μ M [³H]-labeled isoleucine (300 μ Ci/mL), and 5 μ M D345A mutant *E. coli* LeuRS in buffer containing 60 mM Hepes pH 7.5, 10 mM MgCl₂, 1 mM DTT, and 10 mM ATP at 37 °C for 1 h. The Ile-tRNA^{Leu} was purified by acid-phenol-chloroform extraction (pH 4.5; Thermo Fisher Scientific), precipitated by adding 1/10 (vol/vol) of 5 M NaCl, dissolved in 10 mM NH₄OAc and desalted using MicroSpin G25 columns (GE Healthcare). The deacylation reaction of Ile-tRNA^{Leu} by *E. cuculii* or *S. cerevisiae* LeuRS was measured by determining hydrolytic rates at 37 °C in 60 mM Hepes pH 7.5, 30 mM KCl, 10 mM MgCl₂, 1 mM DTT, and 1 μ M [³H]-labeled Ile-tRNA^{Leu} (300 μ Ci/ μ M). Reactions were initiated with enzyme diluted to 20 nM.

Measuring Amino Acid Specificity of Microsporidian LeuRS Synthetase. We estimated the ability of *E. cuculii* LeuRS to discriminate between cognate and near-cognate substrates using a pyrophosphate exchange assay. Each reaction was carried out at 37 °C in 20 μ L of buffer containing 150 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 10 mM KF, 2.5 mM DTT, 0.05% BSA (wt/vol), 1 mM sodium pyrophosphate, 2 mM ATP, γ -[³²P]-ATP (1,000 cpm/ μ L final), 0–200 mM of an L-amino acid (leucine, valine, isoleucine, methionine, norvaline, or aspartate), and 1 μ M *E. cuculii* LeuRS synthetase or LeuRS from *S. cerevisiae* (added at the last moment). Each reaction mixture was incubated at 37 °C for 1–30 min. At the end of the experiment, 1- μ L aliquots from each of the reaction solutions were applied to a TLC plate (Whatman) that had been presoaked in water and dried before use. Chromatography was done at room temperature using 0.75 M KH₂PO₄ (pH 3.5) and 1 M urea solution as a running buffer. The radioactive spots were detected and quantified by phosphorimaging. The experiments were performed three times, each time with a different sample of freshly purified LeuRS.

Assessment of the Structural Conservation of the LeuRS Active Site. To map sequence conservation on the 3D structure of LeuRS, we used the crystal structure of LeuRS from the archaeon *Pyrococcus horikoshii* (85) and the ConSurf algorithm (consurf.tau.ac.il/2016/) to calculate sequence conservation scores (86). To calculate sequence conservation scores for microsporidian LeuRS, we used LeuRS sequences from species listed in *SI Appendix, Table S1*, using one sequence per species for which more than one strain has a known genome sequence. To calculate LeuRS sequence conservation scores for non-microsporidian eukaryotic species, we used cytoplasmic LeuRS sequences retrieved from the aminoacyl-tRNA synthetase database (<https://www.uniprot.org/docs/aatrnasy>) (87). The structures were visualized using the PyMOL Molecular Graphics System, version 2.0 (Schrödinger).

Preparation of Protein Extracts from Microsporidian Spores. The microsporidium *V. culicis floridensis* was collected between 1991 and 1994 during field collections of the mosquito *Aedes albopictus* in Florida and has been

continuously propagated between *Anopheles quadrimaculatus* and the moth *Helicoverpa zea* (88). *V. culicis* spores were produced in *H. zea* as described previously (89) by exposing individually starved second or third instar larvae to 10 μ L of a spore suspension containing 3,000 sp/mL and then transferring them to a standard lepidopteran pinto bean/wheat germ diet (90). Adult *H. zea* were homogenized using a Mini-Beadbeater-24 (BioSpec) with 2.7-mm-diameter glass beads at maximum speed for 3 min. Spores were then purified on 30% Ludox HS-40 columns (Sigma-Aldrich) with 5 mM NH₄Cl, layering homogenate from four moths per column and yielding approximately 10⁷ purified spores per moth.

To extract protein, parasites were cryogenically lysed with a Retsch CryoMill using liquid nitrogen. Once parasites were lysed into a fine powder, 300 μ L of lysis buffer [5% SDS, 2 mM MgCl₂, and 50 mM triethylammonium bicarbonate (TEAB)] was added. Lysates were briefly sonicated and rotated gently at room temperature for 30 min before tryptic digestion.

Tryptic Digestion of Microsporidian Protein Extracts. Tris(2-carboxyethyl) phosphine was added to final concentration of 5 mM. Samples were then heated to 55 °C for 20 min and allowed to cool to room temperature, after which methyl methanethiosulfonate was added to a final concentration of 10 mM. Samples were incubated at room temperature for 20 min to complete the blocking of free sulfhydryl groups. An S-Trap protocol (91) was adapted and used to digest proteins. In brief, lysates were acidified with phosphoric acid to a final concentration of 1.2% and added to an S-Trap containing 6 \times lysate volume of S-Trapping buffer (90% methanol and 100 mM TEAB). The S-Trap was spun down at 4,000 \times g for 30 s to remove buffer, washed with 200 μ L of S-Trapping buffer, and then spun again to remove all buffer. Then 2 μ g of sequencing grade trypsin (Promega) in 125 μ L of 50 mM TEAB was then added to the S-Trap, followed by digestion overnight at 37 °C. After digestion, the peptides were eluted from the column with subsequent applications of 50 mM TEAB, 0.2% formic acid in water, and 0.2% formic acid in 50% acetonitrile. Peptides were dried in vacuo, and then reconstituted in 50 μ L of 0.5 M TEAB/70% ethanol and labeled with *N*-succinimidyl-2-morpholine acetate (SMA) to improve b-ion abundance (92). Then 5 μ L of 5% hydroxylamine was added to quench the reaction, and the peptides were dried down in vacuum (93, 94).

Two-Dimensional Fractionation. Peptides were fractionated using a Pierce High pH Reversed-Phase Peptide Fractionation Kit (Thermo Fisher Scientific) according to the manufacturer's instructions with slight modifications. In brief, peptides were reconstituted in 150 μ L of 0.1% TFA, loaded onto the spin column, and centrifuged at 3,000 \times g for 2 min. The column was washed with water, after which peptides were eluted with the following percentages of acetonitrile in 0.1% triethylamine: 7.5%, 12.5%, 20%, and 50%. Each of the four fractions was then separately injected into the mass spectrometer using capillary reverse-phase LC at low pH.

Mass Spectrometry. An Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific), equipped with a nano-ion spray source was coupled to an EASY-nLC 1200 System (Thermo Fisher Scientific). The LC system was configured with a self-pack PicoFrit 75- μ m analytical column with an 8- μ m emitter (New Objective) packed to 25 cm with ReproSil-Pur C18-AQ, 1.9 μ m material (Dr. Maisch GmbH). Mobile phase A consisted of 2% acetonitrile and 0.1% formic acid, and mobile phase B consisted of 90% acetonitrile and 0.1% formic acid. Peptides were then separated using the following steps: at a flow rate of 200 nL/min, 2% B to 6% B over 1 min, 6% B to 30% B over 84 min, 30% B to 60% B over 9 min, 60% B to 90% B over 1 min, held at 90% B for 5 min, 90% B to 50% B over 1 min, and then the flow rate was increased to 500 nL/min as 50% B was held for 9 min.

Eluted peptides were directly electrosprayed into the Orbitrap Fusion Lumos mass spectrometer with the application of a distal 2.3-kV spray voltage and a capillary temperature of 325 °C. A full-scan mass spectrum (Res = 60,000; 400–1,600 *m/z*) was followed by MS/MS (Res = 15,000) using the "Top Speed" method for selection with a 2-s cycle time. High-energy collisional dissociation was used with the normalized collision energy set to 35 for fragmentation, the isolation width set to 1.2, and a duration of 10 s was set for the dynamic exclusion with an exclusion mass width of 10 ppm. We used monoisotopic precursor selection for charge states of 2+ and greater, and all data were acquired in profile mode. MS/MS spectra were analyzed in the ion trap at rapid scan speed.

Database Searching to Estimate Error Rates in Protein Sequence. The total of 44,547 peptides were used in the analysis. Peak list files were generated by Mascot Distiller (Matrix Science). Protein identification and quantification were done using Mascot 2.4 (95) against the Uniprot_ UniProt_ *V. culicis* 20180327 database (2,768 sequences; 952,872 residues). Methylation of

cysteine and N-terminal and lysine SMA modifications were set as fixed modifications, methionine oxidation and mistranslation of Leu/Ala->another amino acid was set as variable modifications and each modification was searched separately. The same was repeated for Val->other amino acid substitutions. Trypsin was used as cleavage enzyme with one missed cleavages allowed. Mass tolerance was set at 3 ppm for an intact peptide mass and 0.2 Da for fragment ions. Search results were rescored to give a final 1% false discovery rate using a randomized version of the same Uniprot_V.culicis database.

ACKNOWLEDGMENTS. We thank Kyle Hoffman, Li-Tao Guo, Takahito Mukai, Hui-Si Kwok, John Beckmann, and Mark Hochstrasser (Yale Univer-

sity) for insightful discussions; Catherine Dunlovey for critically reading and editing the manuscript; Louis Weiss for providing protein samples from microsporidian spores at the initial stage of the project; and Christina Cuomo (Broad Institute) for helping establish a collaborative team that made this project happen. We also thank the reviewers whose comments helped improve our study. This work was performed with assistance from the Cold Spring Harbor Laboratory Mass Spectrometry Shared Resource, which is supported by the National Cancer Institute Cancer Center Support Grant 5P30CA045508 (to D.J.P.), and was also supported by an Established Investigator Award from the Melanoma Research Foundation (to M.J.S.) and by National Institutes of Health Grant R35 GM122560 (to D.S.).

- Dobson A, Lafferty KD, Kuris AM, Hechinger RF, Jetz W (2008) Colloquium paper: Homage to Linnaeus: How many parasites? How many hosts? *Proc Natl Acad Sci USA* 105:11482–11489.
- Moran NA (1996) Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci USA* 93:2873–2878.
- Muller HJ (1964) The relation of recombination to mutational advance. *Mutat Res* 106:2–9.
- Ota T, Kimura M (1971) On the constancy of the evolutionary rate of cistrons. *J Mol Evol* 1:18–25.
- Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics* 78:737–756.
- McCutcheon JP, Moran NA (2011) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* 10:13–26.
- Merhej V, Royer-Carenzi M, Pontarotti P, Raoult D (2009) Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biol Direct* 4:13.
- Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature* 326:411–414.
- Katinka MD, et al. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414:450–453.
- Smith JE (2009) The ecology and evolution of microsporidian parasites. *Parasitology* 136:1901–1914.
- James TY, et al. (2006) Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature* 443:818–822.
- Corsaro D, et al. (2014) Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. *Parasitol Res* 113:1909–1918.
- Corradi N (2015) Microsporidia: Eukaryotic intracellular parasites shaped by gene loss and horizontal gene transfers. *Annu Rev Microbiol* 69:167–183.
- Louis M, Weiss JJB (2014) *Microsporidia: Pathogens of Opportunity* (Wiley, Hoboken, NJ).
- Didier ES, Weiss LM (2011) Microsporidiosis: Not just in AIDS patients. *Curr Opin Infect Dis* 24:490–495.
- Campbell SE, et al. (2013) The genome of *Spraguea lophii* and the basis of host-microsporidian interactions. *PLoS Genet* 9:e1003676.
- Higes M, Meana A, Bartolomé C, Botias C, Martín-Hernández R (2013) *Nosema ceranae* (Microsporidia), a controversial 21st century honey bee pathogen. *Environ Microbiol Rep* 5:17–29.
- Koonin EV (2003) Comparative genomics, minimal gene sets and the last universal common ancestor. *Nat Rev Microbiol* 1:127–136.
- Ibba M, Soll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69:617–650.
- Perona JJ, Gruic-Sovulj I (2014) Synthetic and editing mechanisms of aminoacyl-tRNA synthetases. *Top Curr Chem* 344:1–41.
- Ling J, Reynolds N, Ibba M (2009) Aminoacyl-tRNA synthesis and translational quality control. *Annu Rev Microbiol* 63:61–78.
- Desjardins CA, et al. (2015) Contrasting host-pathogen interactions and genome evolution in two generalist and specialist microsporidian pathogens of mosquitoes. *Nat Commun* 6:7121.
- Cuomo CA, et al. (2012) Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. *Genome Res* 22:2478–2488.
- Bakowski MA, Priest M, Young S, Cuomo CA, Troemel ER (2014) Genome sequence of the microsporidian species *Nematocida* sp1 strain ERTm6 (ATCC PRA-372). *Genome Announc* 2:e00905-14.
- James TY, et al. (2013) Shared signatures of parasitism and phylogenomics unite Cryptomycota and microsporidia. *Curr Biol* 23:1548–1553.
- Heinz E, et al. (2012) The genome of the obligate intracellular parasite *Trachipleistophora hominis*: New insights into microsporidian genome dynamics and reductive evolution. *PLoS Pathog* 8:e1002979.
- Ndikumana S, et al. (2017) Genome analysis of *Pseudoloma neurophilia*: A microsporidian parasite of zebrafish (*Danio rerio*). *J Eukaryot Microbiol* 64:18–30.
- Haag KL, et al. (2014) Evolution of a morphological novelty occurred before genome compaction in a lineage of extreme parasites. *Proc Natl Acad Sci USA* 111:15480–15485.
- Mikhailov KV, Simdyanov TG, Aleoshin VV (2017) Genomic survey of a hyperparasitic microsporidian *Amphiambllys* sp. (Metchnikovellidae). *Genome Biol Evol* 9:454–467.
- Wiredu Boakye D, et al. (2017) Decay of the glycolytic pathway and adaptation to intranuclear parasitism within Enterocytozoonidae microsporidia. *Environ Microbiol* 19:2077–2089.
- Akiyoshi DE, et al. (2009) Genomic survey of the non-cultivable opportunistic human pathogen, *Enterocytozoon bieneusi*. *PLoS Pathog* 5:e1000261.
- Cornman RS, et al. (2009) Genomic analyses of the microsporidian *Nosema ceranae*, an emergent pathogen of honey bees. *PLoS Pathog* 5:e1000466.
- Chen Yp, et al. (2013) Genome sequencing and comparative genomics of honey bee microsporidia, *Nosema apis*, reveal novel insights into host-parasite interactions. *BMC Genomics* 14:451.
- Mittleider D, et al. (2002) Sequence survey of the genome of the opportunistic microsporidian pathogen, *Vittaforma corneae*. *J Eukaryot Microbiol* 49:393–401.
- Pombert JF, Haag KL, Beidas S, Ebert D, Keeling PJ (2015) The *Ordospora colligata* genome: Evolution of extreme reduction in microsporidian and host-to-parasite horizontal gene transfer. *MBio* 6:e02400-14.
- Xu J, et al. (2016) The genome of *Nosema* sp. Isolate YNPr: A comparative analysis of genome evolution within the Nosema/Vairimorpha clade. *PLoS One* 11:e0162336.
- Guo M, Yang XL, Schimmel P (2010) New functions of aminoacyl-tRNA synthetases beyond translation. *Nat Rev Mol Cell Biol* 11:668–674.
- Peyretailade E, et al. (2012) Annotation of microsporidian genomes using transcriptional signals. *Nat Commun* 3:1137.
- Lee SW, Cho BH, Park SG, Kim S (2004) Aminoacyl-tRNA synthetase complexes: Beyond translation. *J Cell Sci* 117:3725–3734.
- Pang YL, Poruri K, Martin SA (2014) tRNA synthetase: tRNA aminoacylation and beyond. *Wiley Interdiscip Rev RNA* 5:461–480.
- Quevillon S, Robinson JC, Berthonneau E, Siatecka M, Mirande M (1999) Macromolecular assemblage of aminoacyl-tRNA synthetases: Identification of protein-protein interactions and characterization of a core protein. *J Mol Biol* 285:183–195.
- Wang CC, Schimmel P (1999) Species barrier to RNA recognition overcome with nonspecific RNA binding domains. *J Biol Chem* 274:16508–16512.
- Francin M, Mirande M (2003) Functional dissection of the eukaryotic-specific tRNA-interacting factor of lysyl-tRNA synthetase. *J Biol Chem* 278:1472–1479.
- Ryckelynck M, Giegé R, Frugier M (2003) Yeast tRNA(Asp) charging accuracy is threatened by the N-terminal extension of aspartyl-tRNA synthetase. *J Biol Chem* 278:9683–9690.
- Guo M, Schimmel P, Yang XL (2010) Functional expansion of human tRNA synthetases achieved by structural inventions. *FEBS Lett* 584:434–442.
- Silvian LF, Wang J, Steitz TA (1999) Insights into editing from an ile-tRNA synthetase structure with tRNA^{Ile} and mupirocin. *Science* 285:1074–1077.
- Nureki O, et al. (1998) Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* 280:578–582.
- Fukai S, et al. (2000) Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA(Val) and valyl-tRNA synthetase. *Cell* 103:793–803.
- Palencia A, et al. (2012) Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. *Nat Struct Mol Biol* 19:677–684.
- Dock-Bregeon A, et al. (2000) Transfer RNA-mediated editing in threonyl-tRNA synthetase: The class II solution to the double discrimination problem. *Cell* 103:877–884.
- Wong FC, Beuning PJ, Nagan M, Shiba K, Musier-Forsyth K (2002) Functional role of the prokaryotic proline-tRNA synthetase insertion domain in amino acid editing. *Biochemistry* 41:7108–7115.
- Repin T, Yaremchuk A, Tkalco M, Cusack S (2006) Structures of two bacterial prolyl-tRNA synthetases with and without a cis-editing domain. *Structure* 14:1511–1525.
- Naganuma M, Sekine S, Fukunaga R, Yokoyama S (2009) Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proc Natl Acad Sci USA* 106:8489–8494.
- Sokabe M, et al. (2009) The structure of alanyl-tRNA synthetase with editing domain. *Proc Natl Acad Sci USA* 106:11028–11033.
- Martini SA, Fox GE (1997) Non-standard amino acid recognition by *Escherichia coli* leucyl-tRNA synthetase. *Nucleic Acids Symp Ser* 36:125–128.
- Chen JF, Guo NN, Li T, Wang ED, Wang YL (2000) CP1 domain in *Escherichia coli* leucyl-tRNA synthetase is crucial for its editing function. *Biochemistry* 39:6726–6731.
- Chen X, et al. (2011) Modular pathways for editing non-cognate amino acids by human cytoplasmic leucyl-tRNA synthetase. *Nucleic Acids Res* 39:235–247.
- Rock FL, et al. (2007) An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science* 316:1759–1761.
- Baker SJ, et al. (2006) Discovery of a new boron-containing antifungal agent, 5-fluoro-1,3-dihydroxy-1-hydroxy-2,1-benzoxaborole (AN2690), for the potential treatment of onychomycosis. *J Med Chem* 49:4447–4450.
- Pham JS, et al. (2013) Aminoacyl-tRNA synthetases as drug targets in eukaryotic parasites. *Int J Parasitol Drugs Drug Resist* 4:1–13.
- Karkhanis VA, Mascarenhas AP, Martin SA (2007) Amino acid toxicities of *Escherichia coli* that are prevented by leucyl-tRNA synthetase amino acid editing. *J Bacteriol* 189:8765–8768.

62. Li L, et al. (2011) Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in *Mycoplasma* parasites. *Proc Natl Acad Sci USA* 108:9378–9383.
63. Melnikov S, et al. (2012) One core, two shells: Bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* 19:560–567.
64. Melnikov S, Manakongtreecheep K, Söll D (2018) Revising the structural diversity of ribosomal proteins across the three domains of life. *Mol Biol Evol*, 10.1093/molbev/msy021.
65. Peyretailade E, et al. (1998) Microsporidian *Encephalitozoon cuniculi*, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. *Nucleic Acids Res* 26:3513–3520.
66. Ironside JE (2013) Diversity and recombination of dispersed ribosomal DNA and protein coding genes in microsporidia. *PLoS One* 8:e55878.
67. Fares MA, Ruiz-González MX, Moya A, Elena SF, Barrio E (2002) Endosymbiotic bacteria: groEL buffers against deleterious mutations. *Nature* 417:398.
68. Beebe K, Ribas De Pouplana L, Schimmel P (2003) Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J* 22: 668–675.
69. Antonellis A, Green ED (2008) The role of aminoacyl-tRNA synthetases in genetic diseases. *Annu Rev Genomics Hum Genet* 9:87–107.
70. Nangle LA, De Crecy Lagard V, Doring V, Schimmel P (2002) Genetic code ambiguity: Cell viability related to the severity of editing defects in mutant tRNA synthetases. *J Biol Chem* 277:45729–45733.
71. Liu Y, et al. (2014) Deficiencies in tRNA synthetase editing activity cause cardioproteinopathy. *Proc Natl Acad Sci USA* 111:17570–17575.
72. Lee JW, et al. (2006) Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature* 443:50–55.
73. Schimmel P (2008) An editing activity that prevents mistranslation and connection to disease. *J Biol Chem* 283:28777–28782.
74. Roy H, Ling J, Alfonso J, Ibba M (2005) Loss of editing activity during the evolution of mitochondrial phenylalanyl-tRNA synthetase. *J Biol Chem* 280:38186–38192.
75. Ye Q, et al. (2015) Degenerate connective polypeptide 1 (CP1) domain from human mitochondrial leucyl-tRNA synthetase. *J Biol Chem* 290:24391–24402.
76. Sarkar J, Poruri K, Boniecki MT, McTavish KK, Martinis SA (2012) Yeast mitochondrial leucyl-tRNA synthetase CP1 domain has functionally diverged to accommodate RNA splicing at expense of hydrolytic editing. *J Biol Chem* 287:14772–14781.
77. Lue SW, Kelley SO (2005) An aminoacyl-tRNA synthetase with a defunct editing site. *Biochemistry* 44:3010–3016.
78. Funk DJ, Wernegreen JJ, Moran NA (2001) Intraspecific variation in symbiont genomes: Bottlenecks and the aphid-buchnera association. *Genetics* 157:477–489.
79. Yadavalli SS, Ibba M (2012) Quality control in aminoacyl-tRNA synthesis its role in translational fidelity. *Adv Protein Chem Struct Biol* 86:1–43.
80. Schimmel P (2011) Mistranslation and its control by tRNA synthetases. *Philos Trans R Soc Lond B Biol Sci* 366:2965–2971.
81. Reynolds NM, Lazazzera BA, Ibba M (2010) Cellular mechanisms that control mistranslation. *Nat Rev Microbiol* 8:849–856.
82. Ling J, O'Donoghue P, Söll D (2015) Genetic code flexibility in microorganisms: Novel mechanisms and impact on physiology. *Nat Rev Microbiol* 13:707–721.
83. Santos M, Colthurst DR, Wills N, McLaughlin CS, Tuite MF (1990) Efficient translation of the UAG termination codon in *Candida* species. *Curr Genet* 17:487–491.
84. Miranda I, et al. (2007) A genetic code alteration is a phenotype diversity generator in the human pathogen *Candida albicans*. *PLoS One* 2:e996.
85. Fukunaga R, Yokoyama S (2005) Crystal structure of leucyl-tRNA synthetase from the archaeon *Pyrococcus horikoshii* reveals a novel editing domain orientation. *J Mol Biol* 346:57–71.
86. Ashkenazy H, et al. (2016) ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44: W344–W350.
87. Szymanski M, Deniziak MA, Barciszewski J (2001) Aminoacyl-tRNA synthetases database. *Nucleic Acids Res* 29:288–290.
88. Vávra J, Becnel JJ (2007) *Vavraia culicis* (Weiser, 1947) Weiser, 1977 revisited: Cytological characterisation of a *Vavraia culicis*-like microsporidium isolated from mosquitoes in Florida and the establishment of *Vavraia culicis* floridensis subsp. n. *Folia Parasitol (Praha)* 54:259–271.
89. Solter LF, Becnel JJ, Vávra J (2012) Research methods for entomopathogenic microsporidia and other protists. *Manual of Techniques in Invertebrate Pathology*, ed Lacey LA (Academic, New York), 2nd Ed, pp 329–371.
90. Leppla NC, Vail PV, Rye JR (1984) Mass rearing the cabbage looper, *Trichoplusia ni*. *Advances and Challenges in Insect Rearing*, eds King EG, Leppla NC (Agricultural Research Service, Southern Region, US Department of Agriculture, New Orleans), pp 248–254.
91. Zougman A, Selby PJ, Banks RE (2014) Suspension trapping (STrap) sample preparation method for bottom-up proteomics analysis. *Proteomics* 14:1006–1000.
92. Karimi-Busheri F, et al. (1999) Molecular characterization of a human DNA kinase. *J Biol Chem* 274:24187–24194.
93. Gilar M, Olivova P, Daly AE, Gebler JC (2005) Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J Sep Sci* 28:1694–1703.
94. Wang Y, et al. (2011) Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells. *Proteomics* 11: 2019–2026.
95. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567.