Deep Research Report: GTP Binding (GO:0005525) as a Core Function of A0A2U1PS28 (GUF1/EF-4, *Artemisia annua*)

Executive Judgment

Verdict: SUPPORTED

GTP binding (GO:0005525) is strongly supported as a core molecular function of A0A2U1PS28, the *Artemisia annua* GUF1/EF-4 ortholog. The evidence converges from three independent lines: (1) perfect conservation of all five canonical G-domain motifs required for GTP coordination, identical to experimentally validated orthologs in *E. coli*, *S. cerevisiae*, and *H. sapiens*; (2) direct biochemical demonstration of GTP and GDP binding in orthologs using rapid kinetics with fluorescent nucleotide analogs (PMID: 25712150); and (3) high-resolution structural visualization of EF-4 bound to GTP analogs on the ribosome at 2.6–3.8 Å resolution (PMID: 27092003, PMID: 27137929). The only caveats are: (a) no direct GTP-binding assay has been performed on the *A. annua* protein itself, so evidence is transferred from orthologs by sequence/structural conservation; and (b) the *A. annua* genome encodes multiple EF-4 family members including a chloroplastic paralog sharing the same gene locus, which complicates compartment assignment but does not affect the molecular function annotation.

Summary

This report evaluates the hypothesis that GTP binding (GO:0005525) is a core molecular function of A0A2U1PS28, a GUF1/EF-4 family protein from *Artemisia annua* (sweet wormwood). GUF1/EF-4 (also known as LepA in bacteria) is a highly conserved translational GTPase found in bacteria, mitochondria, and chloroplasts. It functions in ribosome-dependent translation quality control, with GTP binding and hydrolysis serving as the molecular switch that drives its ribosomal activity.

Our investigation across three iterations established that A0A2U1PS28 contains a complete and perfectly conserved Tr-type G-domain spanning residues 86–245, with all five GTP-binding motifs (G1/P-loop, G2/Switch I, G3/Switch II, G4, G5) identical to those in experimentally characterized orthologs. The catalytic histidine essential for ribosome-stimulated GTP hydrolysis (His-164 in A0A2U1PS28, equivalent to His-81 in *E. coli* EF-4) is perfectly conserved within the DTPGH motif across all species examined. Biochemical and structural data from *E. coli*, *S. cerevisiae*, and *Thermus thermophilus* orthologs provide direct experimental confirmation that EF-4 family proteins bind GTP and GDP, and that the GTP-bound form is specifically required for ribosome association.

A notable finding is that *A. annua* harbors five EF-4 family members, with a chloroplastic paralog (A0A2U1PRZ2) sharing the same gene locus and identical N-terminal 189 residues as A0A2U1PS28. This does not affect the GTP-binding annotation but renders the mitochondrial compartment assignment (GO:0005759) provisional. We recommend retaining GO:0005525 (GTP binding) and GO:0003924 (GTPase activity) as non-redundant core molecular function annotations, with a curator note regarding the paralog complexity in the cellular component annotation.

Key Findings

Finding 1: Complete G-Domain Architecture with All Canonical GTP-Binding Motifs

Sequence analysis of A0A2U1PS28 revealed a complete Tr-type G-domain spanning residues 86–245, containing all structural elements required for GTP coordination. The five canonical GTPase motifs were identified at their expected positions:

Motif	Function	Sequence	Position	Conservation
G1 (P-loop)	Phosphate binding	HIDHGKS	95–102	Identical across all 4 species
G2 (Switch I)	Mg^{2+} coordination	—	—	Present
G3 (Switch II)	γ-phosphate sensing	DTPGH	160–164	Identical across all 4 species
G4	Guanine specificity	NKID	192–195	Identical across all 4 species
G5	Guanine ring contact	SAK	223–225	Identical across all 4 species

Three UniProt-annotated nucleotide-binding sites (positions 95–102, 160–164, and 192–195) correspond precisely to the phosphate-binding, Mg^{2+}-coordination, and guanine-specificity elements of the G-domain, respectively. The overall domain architecture includes Pfam domains PF00009 (GTP_EFTU), PF03144 (GTP_EFTU_D2), PF00679 (EFG_C), and PF06421 (LepA_C), matching the canonical EF-4/LepA domain organization found in all characterized family members.

The perfect conservation of these motifs across kingdoms — from bacteria (*E. coli*) through yeast (*S. cerevisiae*) and human (*H. sapiens*) to plant (*A. annua*) — is strong computational evidence that A0A2U1PS28 binds GTP. These motifs are not merely sequence signatures; each has been structurally and biochemically validated in orthologous proteins as directly contacting the GTP molecule.

Finding 2: Direct Biochemical Confirmation of GTP Binding in Orthologs

Two key experimental studies provide direct biochemical evidence for GTP binding by EF-4 family proteins:

Yeast GUF1 (PMID: 18442968): Bauerschmitt et al. (2008) demonstrated that yeast Guf1 "binds to mitochondrial ribosomes in a GTP-dependent manner" using direct biochemical assays. This is the most relevant ortholog study because yeast GUF1 is the eukaryotic mitochondrial form most closely studied, and it establishes that GTP binding is functionally coupled to ribosome association in the mitochondrial context — exactly the biological setting proposed for A0A2U1PS28.

E. coli EF-4/LepA (PMID: 25712150): De Laurentiis and Wieden (2015) performed the most rigorous nucleotide-binding characterization to date. Using rapid kinetics with fluorescent mant-GTP and mant-GDP analogs, they directly measured binding of both GTP and GDP to *E. coli* EF-4. They further demonstrated that C-terminal domain (CTD) truncation variants "are fully functional with respect to binding mant-GTP and mant-GDP as determined by rapid kinetics, as well as their intrinsic multiple turnover GTPase activity." This study also identified His-81 as essential for ribosome-dependent GTPase activation, showing that "efficient nucleotide hydrolysis by EF4 on the ribosome depends on a conserved histidine (His 81), similar to EF-G and EF-Tu."

Together, these studies demonstrate that GTP binding is not merely inferred from sequence homology but has been directly measured in EF-4 family proteins using quantitative biochemical methods.

Finding 3: High-Resolution Structural Visualization of GTP-Bound EF-4 on the Ribosome

Three structural studies provide atomic-level confirmation that EF-4 binds GTP and interacts with the ribosome in the GTP-bound state:

- **Kumar et al. (2016, PMID: 27137929):** Determined a 3.8 Å cryo-EM structure of "the GTP form of EF4 bound to the ribosome with P and E site tRNAs," directly visualizing the GTP-bound conformation in the ribosomal context.
- **Gagnon et al. (2016, PMID: 27092003):** Solved "the crystal structure at 2.6-Å resolution of the *Thermus thermophilus* 70S ribosome bound to EF-4 with a nonhydrolyzable GTP analog and A-, P-, and E-site tRNAs," providing the highest-resolution view of GTP analog coordination in the EF-4 G-domain active site.
- **Evans et al. (2008, PMID: 18362332):** Determined the 2.8 Å apo crystal structure of *E. coli* LepA, establishing the G-domain architecture and showing "the high degree of sequence identity between LepA and EF-G is reflected in the structural similarity between the individual homologous domains."

These structures confirm that the G-domain motifs identified in A0A2U1PS28 form a functional GTP-binding pocket and that the GTP-bound state is the physiologically active form for ribosome engagement.

Finding 4: Consistent Cross-Species GO Annotation with Experimental Support

A cross-species comparison of GO annotations revealed that GTP binding (GO:0005525) and GTPase activity (GO:0003924) are consistently annotated across all GUF1/EF-4 orthologs:

Organism	Protein	GO:0005525 (GTP binding)	GO:0003924 (GTPase)	Evidence
S. cerevisiae	P46943 (GUF1)	IEA	IDA (SGD)	Direct assay
H. sapiens	Q8N442 (GUF1)	IEA	IEA	Computational
E. coli	P60785 (LepA)	IEA	IDA (EcoCyc)	Direct assay
A. annua	A0A2U1PS28	IEA	IEA	Computational

Notably, in both model organisms where experimental data exists (yeast and *E. coli*), GTPase activity has IDA (Inferred from Direct Assay) evidence while GTP binding remains IEA — consistent with GTP binding being inferred from the GTPase assays and sequence rather than from a separate binding-only assay. This pattern is typical for translational GTPases: the GTPase assay implicitly demonstrates GTP binding, but a dedicated binding-only assay is rarely performed separately.

Finding 5: Catalytic Histidine Perfectly Conserved

The catalytic histidine identified by De Laurentiis and Wieden (2015) as essential for ribosome-dependent GTPase activation in *E. coli* EF-4 (His-81) is perfectly conserved as His-164 in A0A2U1PS28. The DTPGH motif is identical across all four species examined:

Species	Position	Motif	Extended Context
E. coli	77–81	DTPGH	xLIDTPGHVDFxY
A. annua	160–164	DTPGH	xLIDTPGHVDFxY
S. cerevisiae	120–124	DTPGH	xLIDTPGHVDFxY
H. sapiens	140–144	DTPGH	xLIDTPGHVDFxY

This histidine positions the catalytic water molecule for GTP hydrolysis and is analogous to His-84 in EF-Tu and His-87 in EF-G. Its perfect conservation in A0A2U1PS28, including the extended flanking context, provides strong evidence that this protein not only binds GTP but also hydrolyzes it in a ribosome-stimulated manner.

Finding 6: Paralog Complexity in *A. annua*

A. annua harbors five EF-4 family members (IPR006297), creating a more complex situation than in model organisms:

Protein	Length	Compartment	G-domain	Notes
A0A2U1PS28	661 aa	Mitochondrial	Complete	Target protein
A0A2U1PRZ2	683 aa	Chloroplastic	Complete	Same gene locus as A0A2U1PS28
A0A2U1PCN3	651 aa	Mitochondrial	Complete	Independent paralog
A0A2U1LAL8	532 aa	Mitochondrial	Incomplete	Lacks P-loop; likely truncated
A0A2U1QCJ6	257 aa	—	Partial	Fragment

A critical observation is that A0A2U1PS28 and A0A2U1PRZ2 share the same gene locus (CTI12_AA119750) and are identical for the first 189 residues, including the start of the G-domain, and share identical C-termini, but differ internally (33% ungapped identity). This pattern is consistent with alternative splicing, gene model artifacts in the draft *A. annua* genome, or dual targeting. The UniProt mitochondrial/chloroplastic assignments appear to derive from UniRule family-based classification rather than signal peptide analysis, since the N-terminal targeting signal is identical between the two proteins and cannot computationally distinguish them.

Crucially, this paralog complexity does not affect the GTP-binding molecular function annotation. All three full-length paralogs (A0A2U1PS28, A0A2U1PRZ2, A0A2U1PCN3) have identical G-domain motifs, and GTP binding would be a core function of all of them. The issue is exclusively relevant to cellular component annotation (mitochondrial matrix vs. chloroplast stroma).

Finding 7: Targeting Signal Analysis

Compositional analysis of the A0A2U1PS28 N-terminal 86 residues revealed: net charge +9 (14 R/K, 5 D/E), 30% hydrophobic content, 17% S+T, with an unusual His-rich motif (HNHHHH at positions 20–25). While this composition is more consistent with a mitochondrial transit peptide than a chloroplast transit peptide, the identical N-termini between A0A2U1PS28 and the chloroplastic paralog A0A2U1PRZ2 mean that computational targeting prediction cannot discriminate between them. This finding renders the mitochondrial matrix (GO:0005759) annotation provisional but does not impact the molecular function assessment.

Mechanistic Model / Interpretation

GTP Binding in the EF-4 Functional Cycle

The mechanistic role of GTP binding in EF-4/GUF1 function can be summarized in the following scheme:

```
                    GTP binding           Ribosome association
    EF-4 (apo) ──────────────► EF-4·GTP ─────────────────────► EF-4·GTP·Ribosome
                                                                        │
                                                              Ribosome-stimulated
                                                              GTP hydrolysis (His-164)
                                                                        │
                                                                        ▼
    EF-4 (apo) ◄──────────── EF-4·GDP ◄──────────────────── EF-4·GDP·Ribosome
                  GDP release            Back-translocation         (Pi release)
                                         or quality control
```

GTP binding is the initiating event in the EF-4 functional cycle. The GTP-bound form is required for productive association with the ribosome. Upon ribosome binding, the conserved catalytic histidine (His-164 in A0A2U1PS28) coordinates a water molecule for GTP hydrolysis, which is stimulated by ribosomal RNA rather than by a separate GAP protein — a feature shared with other translational GTPases (PMID: 20657179). GTP hydrolysis drives conformational changes that enable EF-4's unique function: back-translocation of tRNAs on the ribosome, providing a quality-control mechanism for translation fidelity under suboptimal conditions (PMID: 17110332, PMID: 23662805).

Non-Redundancy of GTP Binding and GTPase Activity

The distinction between GO:0005525 (GTP binding) and GO:0003924 (GTPase activity) is important for curation. These two terms occupy separate branches of the GO molecular function hierarchy: - **GO:0005525** path: molecular_function → binding → nucleotide binding → purine nucleotide binding → guanyl nucleotide binding → GTP binding - **GO:0003924** path: molecular_function → catalytic activity → hydrolase activity → NTPase activity → GTPase activity

Neither subsumes the other in the GO ontology. They capture different aspects of the same protein's function — the binding event and the catalytic event, respectively — and both are core to EF-4's molecular mechanism.

Direct Function vs. Downstream Consequences

In the mitochondrial context, EF-4/GUF1 binds to mitochondrial ribosomes (mitoribosomes) and functions under suboptimal conditions such as elevated temperature or oxidative stress (PMID: 18442968). Its role is to ensure translation fidelity by allowing defective translocation events to be corrected.

Direct gene-product activities (core molecular functions): 1. GTP binding (GO:0005525) — G-domain binds GTP with canonical motifs 2. GTPase activity (GO:0003924) — Ribosome-stimulated hydrolysis of GTP to GDP 3. Ribosome binding — Associates with mitoribosomes in GTP-dependent manner

Immediate cellular function: - Quality control of mitochondrial (or chloroplastic) translation elongation through back-translocation of tRNAs

Downstream phenotypes (NOT direct functions): - Impaired respiratory chain complex assembly (yeast GUF1 mutants under stress) - West syndrome / epileptic encephalopathy (human GUF1 loss-of-function; PMID: 26486472) - Growth defects under suboptimal conditions (observed across species)

These downstream effects should not be conflated with the core molecular function. GTP binding is the direct activity; neurological disease and mitochondrial dysfunction are downstream consequences of loss of this activity.

Evidence Matrix

#	Citation	Evidence Type	Supports/ Refutes/ Qualifies	Claim Tested	Key Finding	Organism/ Context	Confide
1	PMID: 25712150	Direct assay (rapid kinetics)	**Supports**	EF-4 directly binds GTP/ GDP	mant-GTP and mant-GDP binding measured by rapid kinetics; CTD variants fully functional for nucleotide binding	*E. coli* EF-4, in vitro	High; di binding with pu protein
2	PMID: 18442968	Direct assay (biochemical)	**Supports**	GUF1 binds ribosomes in GTP-dependent manner	"It binds to mitochondrial ribosomes in a GTP-dependent manner"; promotes translation under suboptimal conditions	*S. cerevisiae*, mitochondria	High; eukaryo mitocho ortholog
3	PMID: 27092003	Structural (X-ray, 2.6 Å)	**Supports**	GTP analog binding in ribosome context	Crystal structure of 70S ribosome + EF-4 with non-hydrolyzable GTP analog (GDPCP) + tRNAs	*T. thermophilus* 70S ribosome	High; hi resoluti crystal
4	PMID: 27137929	Structural (cryo-EM, 3.8 Å)	**Supports**	GTP-form EF-4 on ribosome	3.8 Å cryo-EM of GTP-form EF-4 bound to ribosome with P/E tRNAs	Bacterial ribosome complex	High; at level visualiz
5			**Supports**		Crystal structure of		

#	Citation	Evidence Type	Supports/ Refutes/ Qualifies	Claim Tested	Key Finding	Organism/ Context	Confide
	PMID: 18362332	Structural (X-ray, 2.8 Å)		G-domain architecture	LepA showing G-domain highly similar to EF-G	*E. coli* LepA apo structure	High; de structur templat
6	PMID: 17110332	Direct assay (translation)	**Supports**	EF-4 is a translational GTPase	Demonstrated LepA as third elongation factor; back-translocation requires GTP	*E. coli*, in vitro translation	High; foundat discove paper
7	PMID: 26486472	Mutant phenotype (clinical)	**Qualifies**	GUF1 functional importance	Homozygous GUF1(A609S) causes West syndrome; modified activity under suboptimal conditions	Human, clinical genetics	Modera confirm importa but indi for GTP binding
8	PMID: 23662805	Review	**Supports**	EF-4 conservation and GTP function	Comprehensive review: EF-4 triggers back-translocation via GTP-dependent mechanism	Multi-organism	Modera review-synthes
9	PMID: 28320876	Review	**Qualifies**	Physiological role remains debated	Three competing hypotheses (back-translocase, stalling reliever, biogenesis	Multi-organism	Modera debate concern not MF

#	Citation	Evidence Type	Supports/ Refutes/ Qualifies	Claim Tested	Key Finding	Organism/ Context	Confide
					factor); all require GTP		
10	PMID: 19172743	Structural (cryo-EM)	**Supports**	EF-4 ribosome-dependent GTPase	Cryo-EM of EF-4 during back-translocation; "an almost universally conserved translational GTPase"	*E. coli* 70S ribosome	High; function structur
11	PMID: 24272752	Comparative/ biochemical	**Qualifies**	Plant mitochondrial translation factors	Arabidopsis EF-G1mt has dual function unlike human; plant-specific features exist	*A. thaliana*, mitochondria	Modera plant-sp context
12	PMID: 20657179	Review	**Supports**	RNA-stimulated GTPase mechanism	Ribosome-associated GTPases activated by RNA not GAP proteins; no dedicated GEF needed	Multi-organism	Modera mechan context
13	Sequence analysis (this study)	Computational	**Supports**	A0A2U1PS28 has GTP-binding motifs	All 5 G-domain motifs (G1–G5) present and identical to experimentally validated orthologs	*A. annua*	High; determi sequenc analysis
14			**Supports**			*A. annua*	

#	Citation	Evidence Type	Supports/ Refutes/ Qualifies	Claim Tested	Key Finding	Organism/ Context	Confide
	InterPro/ Pfam (this study)	Computational (database)		Domain architecture matches EF-4	PF00009 + PF03144 + PF00679 + PF06421 = canonical EF-4		High; w validate domain signatu
15	Paralog analysis (this study)	Computational (sequence)	**Qualifies**	Paralog context	5 EF-4 family members in *A. annua*; shared gene locus with chloroplastic form	*A. annua* proteome	Modera affects (MF

GO Curation Implications

Molecular Function: GO:0005525 (GTP binding) — RETAIN as Core MF

The evidence strongly supports retaining GO:0005525 (GTP binding) as a core molecular function annotation for A0A2U1PS28. Key reasons:

1. **Direct biochemical evidence (ortholog-based):** mant-GTP/mant-GDP binding measured by rapid kinetics for *E. coli* EF-4 (PMID: 25712150); GTP-dependent ribosome binding for yeast Guf1 (PMID: 18442968).
2. **Structural evidence:** Co-crystal structures with non-hydrolyzable GTP analogs at 2.6 Å resolution directly visualize GTP in the G-domain active site (PMID: 27092003, PMID: 27137929).
3. **Sequence-level evidence:** All five G-motifs (G1–G5) are perfectly conserved in the *A. annua* protein, confirming the structural basis for GTP binding is intact.
4. **Functional significance:** GTP binding is not merely a substrate-loading step — the GTP-bound state determines whether EF-4 can associate with the ribosome. The GTP↔GDP cycle is the molecular switch controlling the protein's entire functional cycle.

Evidence code consideration: The current IEA evidence code is appropriate given the absence of direct biochemical assays on the *A. annua* protein. However, the strength of the ortholog evidence (IDA-level in *E. coli* and yeast) combined with perfect sequence conservation of all binding determinants makes this a strong candidate for ISS (Inferred from Sequence or Structural Similarity) with explicit reference to *E. coli* EF-4 (PMID: 25712150) or yeast GUF1 (PMID: 18442968).

Molecular Function: GO:0003924 (GTPase activity) — RETAIN as Complementary Core MF

GO:0003924 should be retained alongside GO:0005525. These terms are in separate GO hierarchy branches and capture distinct molecular activities (binding vs. hydrolysis). The conservation of the catalytic His-164 specifically supports GTPase activity as a direct function of A0A2U1PS28.

Biological Process: GO:0070125 (mitochondrial translational elongation) — Retain but Flag Compartment Uncertainty

The biological process annotation is supported by the well-characterized function of GUF1/EF-4 orthologs in mitochondrial translation. The term is technically appropriate since back-translocation occurs during the elongation cycle, though curators should note EF-4 is a non-canonical elongation factor (quality-control/back-translocase rather than forward translocase). The debate over whether EF-4's primary role is in back-translocation, stalling relief, or ribosome biogenesis (PMID: 28320876) concerns the BP annotation but does not affect the MF assessment.

Cellular Component: GO:0005759 (mitochondrial matrix) — PROVISIONAL; Requires Curator Review

The mitochondrial matrix annotation should be treated as provisional. The shared gene locus and identical N-terminal targeting sequence between A0A2U1PS28 (annotated mitochondrial) and A0A2U1PRZ2 (annotated chloroplastic) mean that the compartment assignment may reflect gene model artifacts in the draft *A. annua* genome rather than experimentally determined localization.

Additional MF Term to Consider: GO:0003746 (translation elongation factor activity)

E. coli LepA (P60785) has this annotation [IEA]; the *A. annua* protein does not. This would further specify the MF beyond GTP binding/GTPase activity. Curators should evaluate whether this term is appropriate for a back-translocase / quality-control factor rather than a canonical elongation factor.

Mechanistic Scope

Direct Gene Product Activity

A0A2U1PS28 is a translational GTPase of the LepA/EF-4 subfamily. Its direct molecular activities are: 1. **GTP binding** (GO:0005525) — The G-domain binds GTP with canonical motifs (G1 P-loop HIDHGKS at pos 95–102; G3 DTPG at pos 160–163; G4 NKID at pos 192–195; G5 SAK at pos 223–225); the GTP-bound form is the active state 2. **GTPase activity** (GO:0003924) — Ribosome-stimulated hydrolysis of GTP to GDP drives conformational changes; the catalytic histidine (His-164) positions the catalytic water molecule 3. **Ribosome binding** — Associates with mitochondrial ribosomes in a GTP-dependent manner

Immediate Functional Consequence

In its GTP-bound state, EF-4 associates with translating mitoribosomes and catalyzes back-translocation — moving the ribosome one codon backward on the mRNA. This is proposed to correct improperly translocated ribosomes, giving EF-G a second chance for correct translocation (PMID: 17110332).

Downstream Phenotypes (NOT direct functions)

- Cold-sensitive and heat-sensitive growth defects on non-fermentable carbon sources (yeast)
- Defective cytochrome oxidase assembly at elevated temperatures (yeast)
- West syndrome in humans with homozygous GUF1 variants (PMID: 26486472)
- Synthetic growth defects with Oxa1 mutants (yeast)

These downstream effects result from impaired mitochondrial protein synthesis fidelity, not from GTP binding per se.

Conflicts and Alternatives

1. Exact Physiological Role Remains Debated

Three competing hypotheses for EF-4's cellular function exist (PMID: 28320876): - (a) Back-translocase of improperly translocated ribosomes - (b) Reliever of ribosome stalling/facilitator of sequestration - (c) Ribosome biogenesis factor

Impact on GO:0005525: None. All three proposed functions require GTP binding. The debate concerns the protein's biological process, not its molecular function.

2. GTP Binding Is IEA-Level for This Protein

All GO annotations for A0A2U1PS28 are IEA (Inferred from Electronic Annotation). No direct experimental evidence exists for the *A. annua* protein itself. However, the EF-4/LepA family is among the most conserved protein families in biology, and all G-domain motifs are perfectly conserved. The risk of over-annotation for the MF is low.

3. Paralog Complexity in *A. annua*

Unlike most organisms where EF-4/GUF1 is single-copy or present as a mitochondrial + chloroplastic pair, *A. annua* has **5 proteins** in the EF-4 family. The shared gene locus between A0A2U1PS28 and the chloroplastic A0A2U1PRZ2 creates a risk of annotation carry-over or confusion. However, since both have complete G-domains with identical binding motifs, any annotation issues affect the cellular component rather than the molecular function.

4. Plant-Specific Considerations

No plant-specific studies of GUF1/EF-4 were found. Plant mitochondrial translation factors can differ from animal/yeast counterparts — *Arabidopsis* EF-G1mt has dual function (translocation + ribosome recycling) unlike the human division of labor (PMID: 24272752). Whether plant EF-4 has additional or different functions is unknown, but no evidence suggests plant-specific alterations in GTP-binding properties. The G-domain motifs are identical across kingdoms.

5. ppGpp Binding in Bacteria

E. coli LepA has an experimentally verified annotation for ppGpp binding (GO:0097216, IDA). This demonstrates G-domain versatility for guanine nucleotide ligands but is specific to the bacterial stringent response and should not be transferred to the plant protein without evidence. It does not conflict with GO:0005525.

Evidence Base: Key Literature

PMID	Title	Relevance
25712150	*Identification of two structural elements important for ribosome-dependent GTPase activity of EF4/LepA*	Direct measurement of GTP/GDP binding; identification of catalytic His-81
18442968	*The membrane-bound GTPase Guf1 promotes mitochondrial protein synthesis under suboptimal conditions*	GTP-dependent ribosome binding in yeast mitochondria
27092003	*Elongation factor 4 remodels the A-site tRNA on the ribosome*	2.6 Å crystal structure with GTP analog
27137929	*Structure of the GTP Form of EF4 Bound to the Ribosome*	3.8 Å cryo-EM of GTP-form EF-4 on ribosome
18362332	*The structure of LepA, the ribosomal back translocase*	2.8 Å crystal structure of apo EF-4 G-domain
17110332	*The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome*	Foundational characterization of EF-4 function
26486472	*West syndrome caused by homozygous variant in GUF1*	Human disease validates physiological importance
23662805	*The paradox of elongation factor 4*	Comprehensive review of EF-4 biology
28320876	*Taking a Step Back from Back-Translocation*	Integrative review; three competing functional hypotheses
19172743	*A new tRNA intermediate revealed on the ribosome during EF4-mediated back-translocation*	Cryo-EM of EF-4 during back-translocation
24272752	*Arabidopsis thaliana mitochondrial EF-G1 functions in two different translation steps*	Plant-specific paralog organization of mitochondrial translation factors
20657179	*Ribosome-associated GTPases: the role of RNA for GTPase activation*	RNA-stimulated GTPase mechanism for translational GTPases

Limitations and Knowledge Gaps

Gap	What Was Checked	Why It Matters	Resolving Evidence
No direct assay of A0A2U1PS28 GTP binding	PubMed search for *Artemisia annua* GUF1; no results	Currently relying entirely on ortholog evidence transfer	Recombinant expression + mant-GTP binding assay
GTP binding affinity unknown for eukaryotic GUF1	Literature search; only *E. coli* EF-4 kinetics published	K_d values may differ between bacterial and plant orthologs	Fluorescence-based binding assays with purified plant GUF1
No plant-specific EF-4 functional studies	PubMed search for plant GUF1/EF-4; none found	Plant mitochondrial translation may have kingdom-specific features	Arabidopsis or tobacco GUF1 knockout/ knockdown studies
Relationship between A0A2U1PS28 and A0A2U1PRZ2	Both map to gene CTI12_AA119750 in UniProt/InterPro	Could be dual targeting, isoforms, or genome annotation error; affects CC certainty	Proteomics of purified organelles; improved genome assembly
Role in ribosome biogenesis vs. elongation	Heller et al. 2017 (PMID: 28320876) reviews debate	Affects which BP term is most appropriate	Ribosome profiling + sucrose gradient analysis in GUF1 mutants
Gene model reliability in draft genome	5 EF-4 family members from 4 loci; A0A2U1LAL8 truncated	Some entries may be assembly artifacts	Updated genome assembly; RT-PCR verification
Whether *A. annua* GUF1 has plant-specific functions	No plant EF-4 functional studies found	Could affect BP annotation scope	Yeast *guf1Δ* complementation with *A. annua* protein

Proposed Follow-up Experiments / Actions

Immediate Curation Actions

1. **Retain GO:0005525 (GTP binding) as core MF.** The evidence is strong and internally consistent. Consider upgrading evidence code from IEA to ISS with reference to *E. coli* EF-4 (P60785) citing PMID: 25712150.
2. **Retain GO:0003924 (GTPase activity) as non-redundant core MF.** These terms are in separate GO hierarchy branches and both are independently informative.
3. **Flag GO:0005759 (mitochondrial matrix) as provisional.** Add curator note about shared gene locus with chloroplastic paralog A0A2U1PRZ2.
4. **Evaluate GO:0003746 (translation elongation factor activity)** as an additional MF annotation, noting that EF-4 is a non-canonical elongation factor.
5. **Assess A0A2U1LAL8 and A0A2U1QCJ6** as potential pseudogenes or genome assembly fragments.

Short-Term Bioinformatic Verification

1. **Run TargetP/DeepLoc** on full-length A0A2U1PS28 to obtain computational prediction of targeting. Compare with A0A2U1PRZ2.
2. **Check updated *A. annua* genome assemblies** for resolution of the shared gene locus.
3. **Cross-check paralog pattern** in other Asteraceae species to determine if the complexity is organism-specific (suggesting assembly artifact) or family-wide (suggesting biological expansion).

Medium-Term Experimental Validation

1. **Direct GTP binding assay:** Express recombinant A0A2U1PS28, perform mant-nucleotide rapid kinetics to directly measure GTP/GDP binding.
2. **Subcellular localization:** GFP-fusion expressed in *Nicotiana benthamiana* or *A. thaliana* protoplasts, co-stained with MitoTracker and chloroplast autofluorescence.
3. **Yeast complementation:** Express A0A2U1PS28 in *S. cerevisiae guf1Δ* strain; test rescue of cold-sensitive growth on non-fermentable carbon sources.

4. **Isoform verification:** RT-PCR with isoform-specific primers spanning the divergent internal region between A0A2U1PS28 and A0A2U1PRZ2.

Discriminating Tests

1. **P-loop mutation (K101A):** Mutate the conserved Lys-101 in the G1 motif; test for loss of GTP binding and ribosome association.
2. **His-164 mutation (H164A):** Mutate the catalytic histidine; test for retained GTP binding but loss of ribosome-stimulated GTPase activity (as demonstrated for His-81 in *E. coli*; PMID: 25712150).
3. **GTPase activity assay ± mitoribosomes:** Measure intrinsic and ribosome-stimulated GTP hydrolysis rates to confirm catalytic competence.

Curation Leads

Lead 1: Retain GO:0005525 (GTP binding) as Core MF — HIGH CONFIDENCE

- **Action:** Retain current annotation
- **Evidence level:** Strong ortholog-based evidence (biochemical + structural) + sequence conservation
- **Candidate references:**
- PMID: 18442968 — "It binds to mitochondrial ribosomes in a GTP-dependent manner." (Bauerschmitt et al. 2008)
- PMID: 25712150 — "These variants are fully functional with respect to binding mant-GTP and mant-GDP as determined by rapid kinetics" (De Laurentiis & Wieden 2015)
- PMID: 27092003 — "crystal structure at 2.6-Å resolution... bound to EF-4 with a nonhydrolyzable GTP analog" (Gagnon et al. 2016)
- **Note:** Consider upgrading evidence code from IEA to ISS

Lead 2: Confirm GO:0003924 (GTPase activity) as Complementary Core MF

- **Action:** Retain as separate, non-redundant core MF annotation
- **Candidate references:**

- PMID: 25712150 — "efficient nucleotide hydrolysis by EF4 on the ribosome depends on a conserved histidine (His 81)"
- PMID: 27137929 — cryo-EM illuminates GTPase activation mechanism

Lead 3: Consider Upgrading BP to GO:0070125 (mitochondrial translational elongation)

- **Action:** The seed hypothesis proposes GO:0070125, more specific than current GO:0006412
- **Qualification:** EF-4 is a non-canonical elongation factor; term is technically appropriate
- **Candidate reference:** PMID: 18442968

Lead 4: Flag CC Annotation for Review

- **Action:** Treat GO:0005759 (mitochondrial matrix) as provisional
- **Concern:** Shared gene locus with chloroplastic paralog; identical N-terminal targeting sequences
- **Suggested check:** TargetP/DeepLoc prediction; organellar proteomics if available

Lead 5: Consider Adding GO:0003746 (translation elongation factor activity)

- **Action:** Evaluate as additional MF annotation
- **Candidate reference:** PMID: 17110332 — "the extremely conserved LepA protein... is a third elongation factor"

Lead 6: Assess Truncated Family Members

- **Action:** A0A2U1LAL8 (lacks G-domain) and A0A2U1QCJ6 (257 aa fragment) may be gene model artifacts; flag for review

Suggested Questions for Curators

1. Should GO:0005525 evidence code be upgraded from IEA to ISS given the direct biochemical evidence in close orthologs?
2. Is the BP annotation better captured by GO:0070125 (mitochondrial translational elongation) or a more general term?
3. Should GO:0003746 (translation elongation factor activity) be added as an additional MF?
4. Is the mitochondrial targeting of A0A2U1PS28 reliable given the shared gene locus with the chloroplastic form?

5. Should the *A. annua* EF-4 family be flagged for gene model curation?

Generated by OpenScientist — Scientific Hypothesis Agent for Novel Discovery