*Nucleic Acids Research*, 2025, gkaf1062
https://doi.org/10.1093/nar/gkaf1062
**Database issue**

# ENSURE: the encyclopedia of suppressor tRNA with an AI assistant

**Zhuo Ouyang[1,†], Yifeng Zhang[1,†], Fan Feng[1], Xudong Zeng[1,2], Qiuhui Wu[1], Abdul Hafeez[1], Wenkai Teng[1], Yixin Kong[1], Xuan Bu[1], Yang Sun[1], Bin Li[1], Yanzi Wen[1], Zhao-Rong Lun[1], Lianghu Qu[1], Xiao Feng[1,*], Lingling Zheng[1,*]**

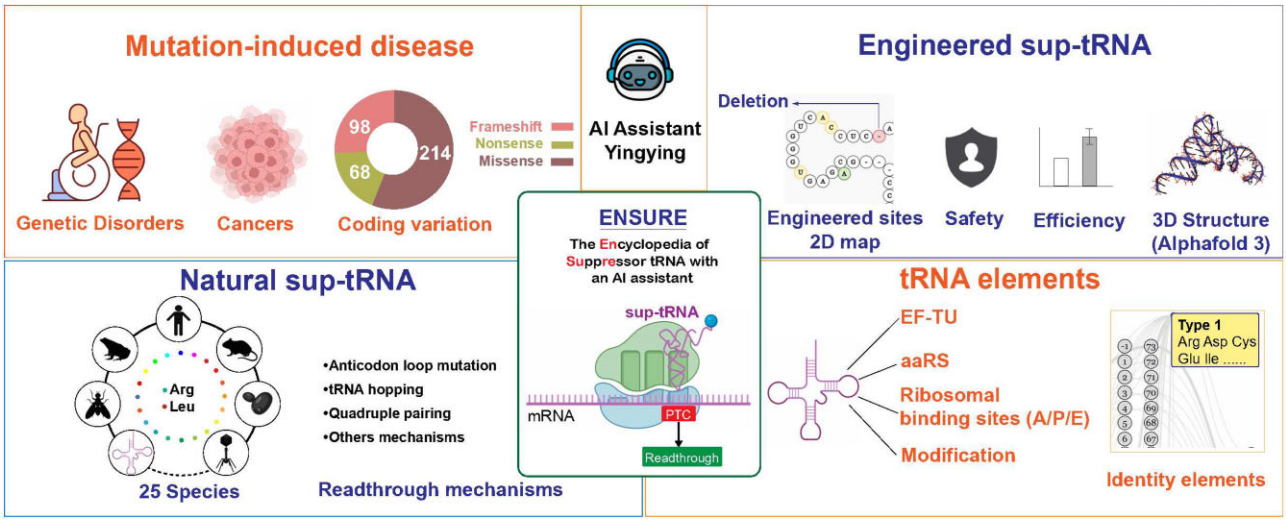
[1]MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, Innovation Center for Evolutionary Synthetic Biology, Guangdong Provincial Key Laboratory of Plant Stress Biology, School of Agriculture and Biotechnology, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China
[2]Tsinghua-Peking Center for Life Sciences, Beijing 100084, China

*To whom correspondence should be addressed. Email: zhengll33@mail.sysu.edu.cn.
Correspondence may also be addressed to Xiao Feng. Email: fengx83@mail.sysu.edu.cn
†The first two authors should be regarded as Joint First Authors.

## Abstract

Suppressor transfer RNAs (sup-tRNAs) offer a promising strategy for rescuing proteins truncated by premature termination codons via translational readthrough. Despite recent advances in genetic code expansion and RNA therapeutics that have facilitated sup-tRNA engineering, progress remains constrained by the lack of a dedicated, integrative data platform. Here, we present ENSURE (https://trna.lumoxuan.cn/), a comprehensive knowledge base that aggregates 2152 disease- and cancer-associated nonsense, missense, and frameshift variants; 86 experimentally validated natural sup-tRNAs, 1108 tRNA engineering strategies; and 487 curated tRNA element records. All sequences undergo multiple sequence alignment, secondary structure prediction, and AlphaFold 3 modeling accompanied by interactive 2D/3D visualization. ENSURE supports keyword and BLAST searches, as well as bulk downloads. A key feature is an AI assistant based on a retrieval-augmented generation architecture: ∼123.7k tokens of database pages, literature abstracts, structural annotations, and other resources are chunked and encoded with Sentence-BERT; user queries are processed similarly, matched to relevant chunks, concatenated with the query, and passed to a large language model to generate answers with illustrative resources inserted automatically. By combining curated data, structural models, and an interactive AI assistant, ENSURE provides a powerful platform to accelerate sup-tRNA research and translational applications.

## Graphical abstract




**Received:** August 14, 2025. **Revised:** September 16, 2025. **Accepted:** September 29, 2025

## Introduction

Premature termination codons (PTCs) disrupt translation, yielding truncated and nonfunctional proteins that underlie many genetic diseases. Suppressor transfer RNAs (sup-tRNAs) provide a direct strategy to circumvent this defect by decoding stop codons and inserting amino acids, thereby suppressing translation termination. Several studies demonstrated that engineered tRNAs can efficiently suppress nonsense mutations in both cellular and *in vivo* contexts, restoring protein function [1, 2]. This strategy holds considerable promise for tRNA therapeutics and protein engineering [3–6]. Typically, the stop codons UAG, UAA, and UGA trigger termination of protein synthesis; in contrast, sup-tRNAs use anticodon-mediated decoding to promote translational readthrough, enabling elongation to continue and producing full-length or posttranslationally modified proteins [7, 8]. In addition to sup-tRNAs, readthrough-inducing small molecules have been developed to promote stop codon readthrough. However, many of these compounds display toxicity and their effective therapeutic window varies markedly between diseases [9, 10]. These dose-limiting adverse effects have constrained the clinical use of such drugs. Consequently, combination strategies that pair therapeutic sup-tRNAs with low-dose or transient administration of readthrough drugs may offer a safer and more effective means to achieve clinically meaningful levels of stop codon readthrough, with the optimal regimen likely differing for each disease [11].

Naturally occurring sup-tRNAs have evolved across diverse organisms, from prokaryotes to eukaryotes. They possess unique sequence and structural features that enable stop codon recognition and amino acid delivery. For example, in *Tetrahymena*, an anticodon mutation in a glutamine tRNA reassigned UAA/UAG codons to glutamine, producing a functional sup-tRNA. Such phenomena have been documented from prokaryotes to eukaryotes, and the underlying mechanisms are diverse [8, 12]. Recently, rational design engineered sup-tRNAs have broad applications, including treatment of mutation-induced genetic diseases [13] and neurological disorders [14], inhibition of tumor growth via suppression of oncogenic protein translation [15], and site-specific incorporation of noncanonical amino acids for genetic code expansion (GCE) [16–18]. In contrast to GCE, genetic code contraction is progressing at a breathtaking pace [19–22]; notably, researchers have designed and constructed *Escherichia coli* that use a 57-codon genetic code [19].

Despite advances in sup-tRNA research and applications in GCE and RNA therapeutics, the availability of a unified and convenient access resource would accelerate progress. To address this need, we developed ENSURE (the encyclopedia of suppressor tRNA with an AI assistant; https://trna.lumoxuan.cn/), a comprehensive knowledge base that aggregates disease-associated nonsense and frameshift variants, curates natural and engineered sup-tRNAs with extensive annotations and structural models, provides analytical tools, and features an AI assistant specifically designed for sup-tRNA research, aimed at accelerating both basic research and translational applications (Fig. 1).

## Materials and methods

### Data collection

The ENSURE database comprises four data modules—mutation-induced diseases, natural sup-tRNAs, engineered sup-tRNAs, and tRNA elements—integrated from public databases and manually curated from primary literature (Table 1). Two service modules (search/visualization and the AI assistant "Yingying") are described in the Results section. The mutation-induced disease module contains 2152 nonsense, missense, and frameshift variants, together with their associated disease and cancer information, derived from HGVdb (online version, https://hgv.figshare.com, June 2025) and COSMIC (online version, June 2025) [13]. Recorded fields comprise gene name, original/mutated codon, genomic coordinates, and disease names. For the natural sup-tRNA module, entries were integrated from RNAcentral (online version, June 2025) [23] and GtRNAdb (online version, June 2025) [24], together with 86 literature-validated natural sup-tRNAs across 25 species. Each record captures species of origin, tRNA mutation sites, carried amino acid, and recognized stop codon types. Moreover, the engineered sup-tRNA module collates data from 60 important PubMed publications, encompassing 1108 engineered sup-tRNAs with documented mutation sites, target genes, readthrough efficiencies, experimental systems, and model organisms [2, 25, 26]. The tRNA elements module contains 487 curated records spanning four categories: 101 aminoacyl-tRNA synthetase (aaRS) identity/recognition entries, 60 elongation-factor recognition sites (e.g. EF-Tu or eukaryotic EF-1$\alpha$), 58 tRNA–ribosome interaction sites (A/P/E), and 268 functional/modification entries, collectively supported by 249 literature references [27, 28]. All database imports and literature curations were harmonized to a unified schema to enable cross-module querying and downstream annotation.

### Data processing and structure modeling

Only sup-tRNA sequences supported by literature or experimental validation were retained, while candidates derived solely from computational predictions or genome-wide screens were excluded. Following the workflow as shown in Fig. 1, entries were cross-validated, near-duplicate sequences were removed at a 95% identity threshold, names were normalized, and manual quality control was applied. Field standardization included matching of species names to NCBI Taxonomy Latin binomials, harmonization of disease names to a controlled vocabulary, a unified "original→mutated" notation for codon changes, and matching of modification names to MODOMICS terminology/codes. Natural entries preserved their published secondary structure schematics, whereas engineered entries were annotated with nucleotide positions using the Sprinzl numbering system [29].

For structure annotation, all sequences were subjected to multiple sequence alignment with MAFFT [30], and secondary structures were predicted with tRNAscan-SE [31]. Sequences were then modeled with a locally deployed Alpha Fold 3 (AF3) pipeline [32]. These models were ranked by average predicted Local Distance Difference Test (pLDDT) score to facilitate user selection [33]. Functional elements were curated and standardized as follows. Chemical modifications adopt MODOMICS names, for example, $m^1G37$ and $\Psi 55$ [34–36]. Literature-supported contact sites for aminoacyl-tRNA synthetases (aaRS), EF-Tu (or eukaryotic EF-1$\alpha$), and ribosomal binding sites (A/P/E) were extracted, and only positions explicitly reported in source texts or figures were retained [27, 37]. Each reported position was mapped to Sprinzl coordinates and assigned to a secondary structure domain (acceptor stem, D-arm, anticodon arm, or T-arm). Base-position

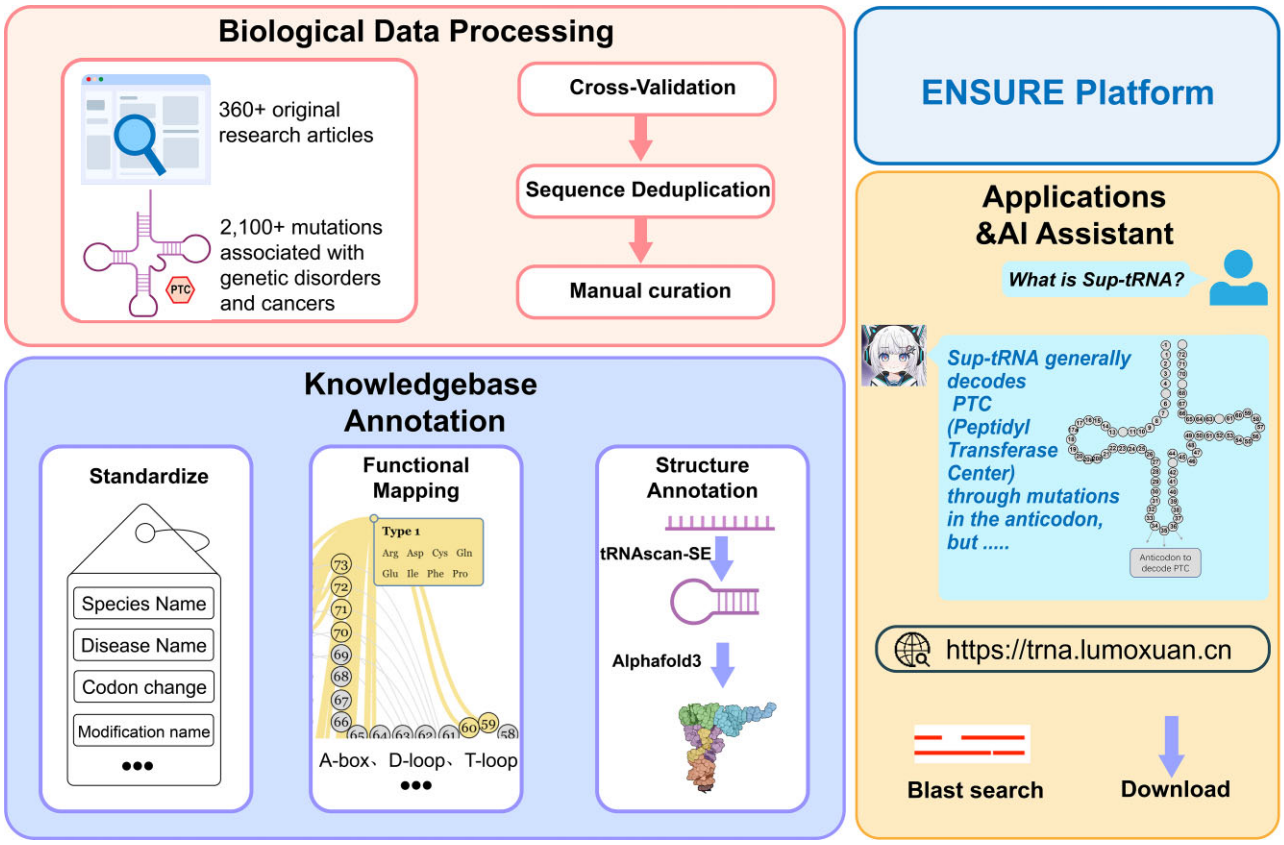



**Figure 1.** Schematic overview of the ENSURE platform. This figure illustrates ENSURE's integrated pipeline for processing biological data, annotating knowledge resources, and enabling AI-assisted applications. Biological data processing: curates sup-tRNA sequences and disease-associated mutations from over 360 research articles, encompassing >2100 mutations across various genetic disorders and cancers, followed by cross-validation, deduplication, and manual curation. Knowledge base annotation: standardizes data, maps functional attributes, and predicts secondary and tertiary structures using tRNAscan-SE and AF3. Applications: an AI assistant integrates user queries with curated knowledge as prompts for large language models, generating responses augmented with automatically inserted visual resources.

**Table 1.** Summary of data source and entry counts for ENSURE database modules

| Module | Source | Entries | Key fields |
|---|---|---|---|
| Mutation-induced disease | HGVdb, COSMIC | 2152 variants | gene name, original/mutated codon, genomic coordinates, associated disease/cancer |
| Natural sup-tRNA | RNAcentral, GtRNAdb, 58 Literatures | 86 entries | species, anticodon, amino acid, recognized stop codon |
| Engineered sup-tRNA | 60 PubMed publications | 1108 entries | Structural sites engineered in sup-tRNAs, readthrough efficiency, assay system, model organism |
| tRNA elements | 249 Literatures | 487 records | aaRS recognition sites, elongation factor /A/P/E binding sites, modification types |

tokens were deduplicated across sources, and the organism/system context was recorded. All entries include Sprinzl numbering and literature references.

## Construction and implementation of the AI assistant

The key feature of ENSURE is the AI assistant "Yingying," which was implemented using a retrieval-augmented generation (RAG) architecture [38]. Source materials, including database pages, literature abstracts, structure annotations, and other resources in the database, were tokenized into ~123.7k tokens, segmented into text chunks, and encoded into vector embeddings using Sentence-BERT [39]. Chunk embeddings are stored in a vector index to enable fast similarity response. User queries are processed in the same manner and matched to chunk vectors to retrieve the most relevant chunks. Retrieved chunks are concatenated with the user's question to form the prompt context for a large language model (LLM), such as GPT-4, which generates the reply. Additionally, all supporting materials, including figures, sequences, and citations, are automatically integrated into the response.

## Database implementation and visualization

The backend is implemented in Python using Flask and PostgreSQL for data storage and exposes RESTful endpoints for CSV export and local BLAST searches [40]. The frontend is a

component-based, statically typed web client built with Vue 3 and TypeScript, using Element Plus for UI components, Pinia for state management, and vue-router for routing. Static assets are served by Nginx with CDN delivery via Cloudflare. Interactive visualizations include secondary RNA structure rendering with Forna [41], three-dimensional (3D) model viewing of mmCIF files with NGL Viewer [42], and tRNA functional-element network diagrams generated with D3.js [43]. These visualizations support interactive annotation, zooming, and data download. The implementation emphasizes modularity, responsive performance, and reproducible programmatic access.

## Results

### Database content

ENSURE is a comprehensive knowledge base for sup-tRNAs that combines systematic literature retrieval, manual curation, and large-scale data processing to aggregate datasets and tools for the sup-tRNA community. Its core features include a mutation-disease module, a catalog of experimentally validated natural sup-tRNAs, a repository of engineered sup-tRNAs, detailed annotations of known tRNA functional elements, and an integrated AI assistant (Fig. 1). Meanwhile, this database provides BLAST-based sequence search and visualization tools alongside an LLM-powered assistant delivering expert guidance for both fundamental and applied sup-tRNA research.

The platform is organized into six main modules: (i) Mutation-induced disease module. Catalogs gene variants associated with 380 genetic disorders and 1772 cancers, providing original and mutated codons, genomic coordinates, and disease annotations to support potential therapeutic target discovery of sup-tRNAs. (ii) Natural sup-tRNA module. Aggregates experimentally validated natural sup-tRNAs from 58 publications, including species of origin, sequences, anticodon variants, and structure information. (iii) Engineered sup-tRNA module. Collects data from 60 literature reports on engineered sup-tRNAs, detailing target genes, engineered mutation sites, reported efficacy (readthrough), and safety metrics. (iv) tRNA elements module. Summarizes key recognition features of tRNA, including sequence motifs, identity elements, and modifications, to inform mechanistic interpretation and design strategies [7, 44–46]. (v) Search and visualization module. Enables BLAST-based sequence similarity searches, interactive visualization of results, and exploration of tRNA structures with target sites for engineering. (vi) AI assistant "Yingying." A RAG-driven assistant built on LLMs that returns evidence-linked answers about sup-tRNAs to user queries to accelerate hypothesis generation and experimental planning. Together, these modules deliver an integrated resource for both fundamental research and translational application of sup-tRNAs.

### Examining common pathogenic mutations across disease types

In principle, any coding frame mutation that creates a premature termination codon can be targeted by a sup-tRNA [2, 4]. When pathogenic mutations in different genes share a similar mechanistic consequence, even if they cause distinct disease phenotypes, the same type of sup-tRNA can be used as a treatment strategy. For example, a CAG-TAG nonsense mutation at position 1648 of *USP9X* causes characteristic craniofacial defects [2, 47], whereas the same codon change at position 265 of *COL7A1* causes recessive dystrophic epidermolysis bullosa (RDEB) [48]. Although these mutations occur in different genes and yield different clinical syndromes, both produce premature termination codons and therefore represent the same therapeutic target class. Thus, ENSURE's mutation-induced disease module enables users to identify such shared mechanisms precisely via keyword and field filters. For instance, filtering the "Codon change" column returns 12 pathogenic entries with the CAG-TAG nonsense mutation (Fig. 2A). Within genetic disorders, the database catalogs 380 pathogenic coding variants, comprising 214 missense, 98 frameshift, and 68 nonsense variants (Fig. 2B). Moreover, nonsense mutations account for ∼17.9% of these entries, with the CGA-TGA (from original codons to stop codons) change being the most frequent (Fig. 2C).

Oncogenic coding variants can similarly be considered for sup-tRNA strategies. Mutations in the tumor suppressor *DDX3X*, for example, are implicated in two lymphoid neoplasms, including Burkitt lymphoma and chronic lymphocytic leukemia [49–53]. ENSURE also contains an extensive cancer-associated collection ("coding variation in cancer") with 1772 records spanning five mutation classes and extensive cancer annotations. By summarizing and classifying common mutational mechanisms across genes and disease types, ENSURE supports the identification and prioritization of cross-disease targets for sup-tRNA design, thereby facilitating the development and translational evaluation of tRNA-based therapies for both genetic disorders and cancers.

### Exploring the mechanisms of natural sup-tRNAs

Under the standard genetic code, no tRNA gene usually decodes stop codons. Nevertheless, during evolution, some wild-type tRNAs have acquired mutations that enable the recognition and decoding of stop codons [7, 54]. With literature retrieval and manual curation, ENSURE currently catalogs 86 validated natural sup-tRNAs, with detailed records of their mutation sites, structural features, recognized stop codons, and carried amino acids. As an illustrative case, Kachale *et al*. reported a new unicellular parasite, *Blastocrithidia nonstop*, in which the UGA stop codon is reassigned to tryptophan (Trp) [22]. This reassignment occurs through a shortening of the Trp-tRNA anticodon stem from 5 to 4 base pairs, enabling UGA recognition and Trp insertion [22, 55]. Users can search the entry "*Blastocrithidia nonstop*" in the natural sup-tRNA module (Fig. 3A), view the basic information for this sup-tRNA (Fig. 3B), expand to view the whole sequence and annotations (Fig. 3C), and zoom the secondary structure schematic to visualize the anticodon-stem shortening (Fig. 3D). Importantly, ENSURE provides a comprehensive natural sup-tRNA collection that comprises experimentally demonstrated suppressors rather than bioinformatic predictions alone. Each entry is linked to primary evidence and includes structural annotations to facilitate mechanistic interpretation and comparative analysis.

### Demonstrating engineered sup-tRNA modification strategies

Natural sup-tRNAs often exhibit limited suppression efficiency, prompting the development of diverse engineering strategies that modify sup-tRNAs in sequence, structure, or

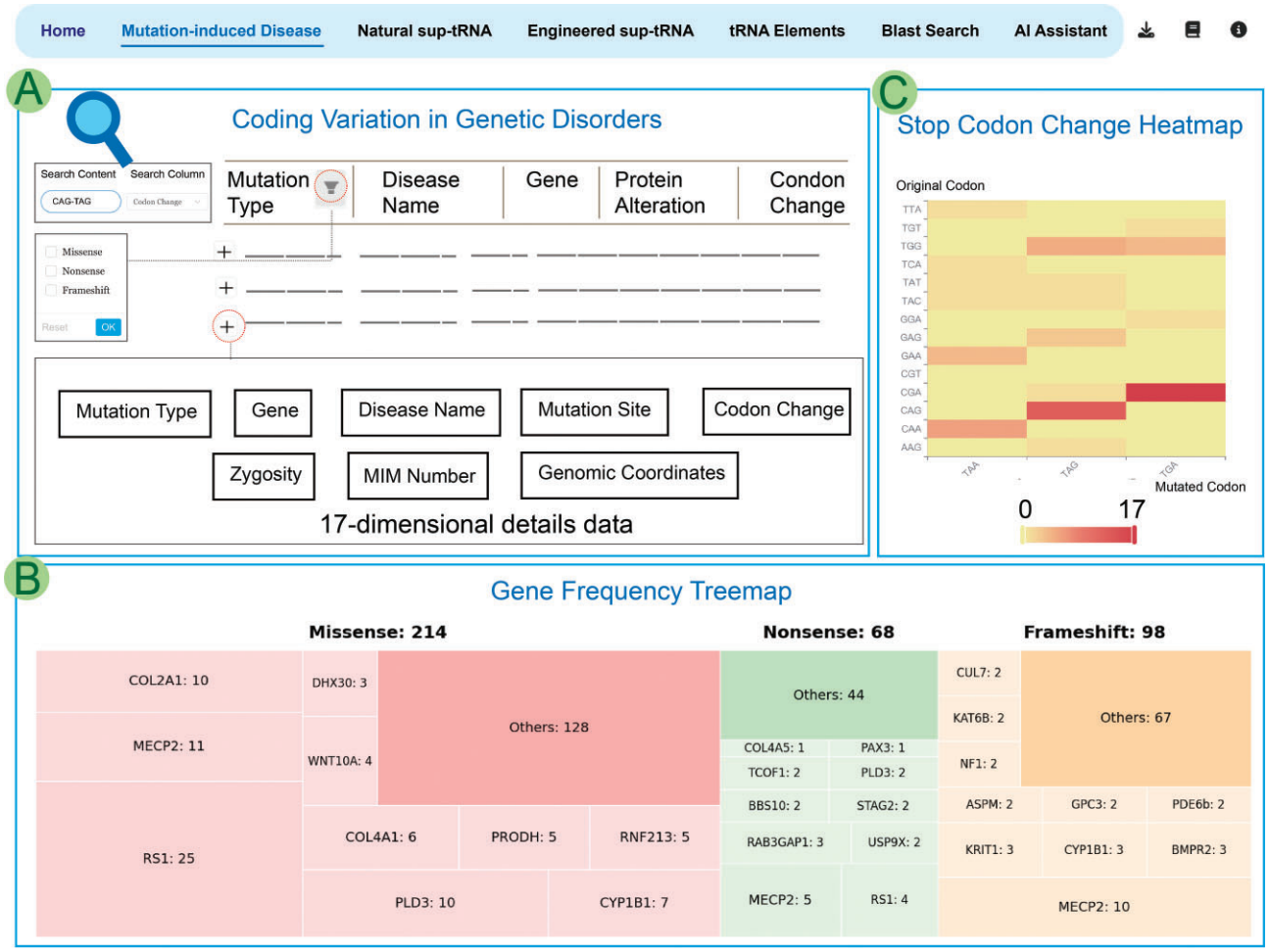



**Figure 2.** Mutation-induced disease module in the ENSURE platform. (**A**) Coding variation in genetic disorders subsection showing a CAG-TAG nonsense mutation example. Clicking "+" expands 17 detailed fields per variation (e.g. MIM number, genomic coordinates). (**B**) Gene frequency treemap depicting the distribution of mutation types (missense, nonsense, and frameshift) across genes associated with genetic disorders. Area represents occurrence frequency. (**C**) Stop codon change heatmap displaying conversion frequencies from original codons to stop codons. Color intensity scales with frequency.

posttranscriptional modification. The effects of engineered tRNA modifications are highly context-dependent, varying with the tRNA feature and cellular environment. ENSURE enables comparison of reported outcomes to inform modification strategies [56]. Users can rapidly locate and inspect engineered-tRNA studies using the platform's filters by paper information, for example, PMID: 37258671 (Fig. 4A and B). Expanding a table row with the "+" icon reveals summary metadata including the target PTC gene, introduced amino acid and anticodon, cell type, and measured readthrough efficiency (Fig. 4C). The view details page provides the mechanistic annotations for each engineered sup-tRNA, including sequence, secondary structure map, annotated modification sites, and alignments between original and engineered sequences (Fig. 4D). ENSURE also supplies AF3-predicted tertiary structure models for engineered sup-tRNAs, which can be rotated, zoomed, and inspected interactively (Fig. 4E), together with secondary structure schematics that precisely highlight modification positions (Fig. 4D). By aggregating empirical readthrough data, structural annotations, and interactive visualizations, ENSURE can help users evaluate which engineering strategies have been effective in particular systems and prioritize rational engineering approaches for new sup-tRNA designs.

## Tools provided by the ENSURE platform

The ENSURE platform offers a BLAST-based sequence search tool that enables users to compare their designed sup-tRNAs with natural and engineered entries in the database. It accelerates design and screening by identifying genetically similar sup-tRNAs, providing a strong theoretical basis for experimental validation and optimization. ENSURE also integrates comprehensive tRNA structural annotations, including binding sites for aminoacyl-tRNA synthetases (aaRS), ribosomal E-, P-, and A-sites, as well as elongation factors [37]. These features guide the rational design of engineered sup-tRNAs by pinpointing positions critical for efficient suppression. For instance, optimizing the tRNA–aaRS interaction can enhance translation fidelity, while tuning ribosome contact points can improve suppression efficacy in specific biological systems [57, 58]. These features not only enhance research resources for sup-tRNAs but also provide powerful tools for designing new tRNA molecules with enhanced suppressor functions. By integrating comprehensive structural and functional tRNA data, researchers can now design more precise sup-tRNAs tailored to specific therapeutic needs, advancing their application in gene therapy and disease intervention.

To further support users, ENSURE includes an AI-powered assistant, "Yingying," built on an LLM specialized for the sup-

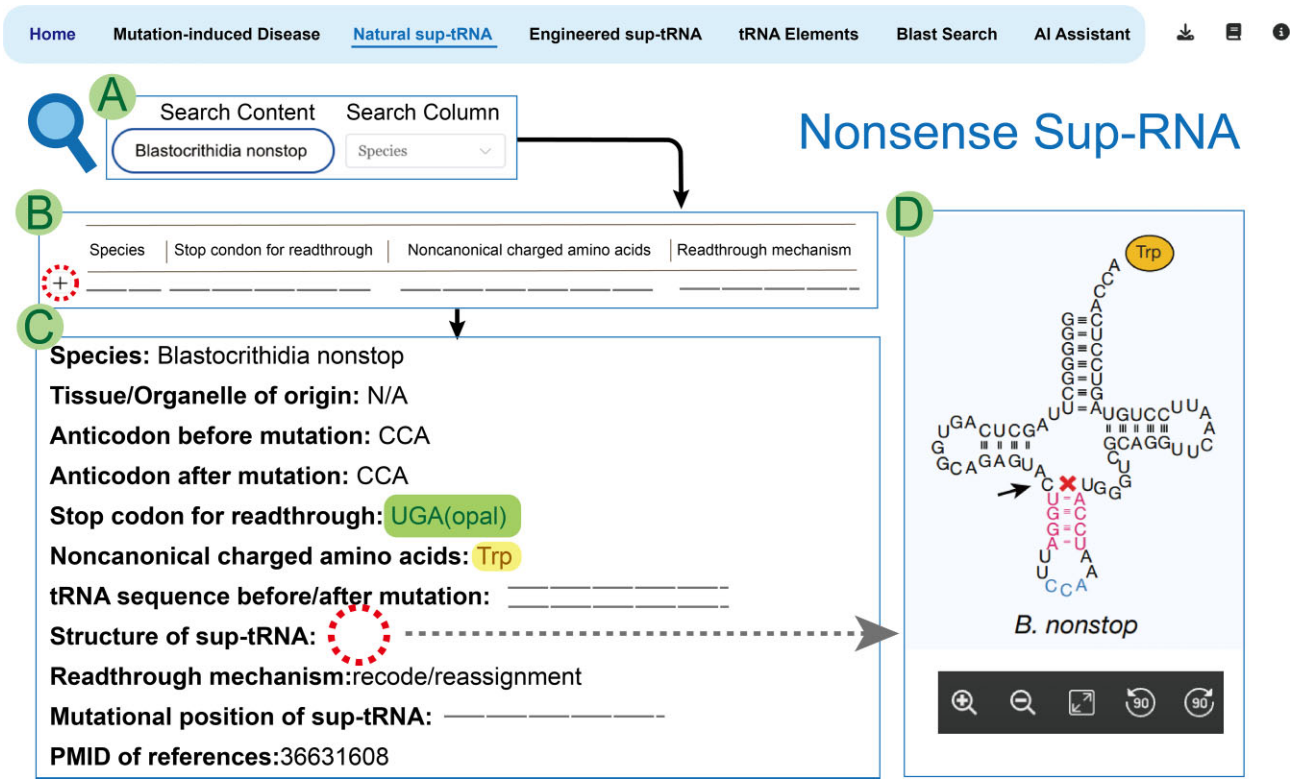



**Figure 3.** Natural sup-tRNA module in the ENSURE platform. (**A**) Search interface showing a query example using the species name "*Blastocrithidia nonstop.*" (**B**) Filterable result list. (**C**) Detail panel showing anticodon sequence (before/after mutation) and functional annotations. (**D**) Predicted secondary structure highlighting the mutation site where the anticodon stem shortened from 5 to 4 bp.

tRNA knowledge base. Drawing on 123.7k curated tokens from the ENSURE database, Yingying can answer research questions in natural language, retrieve relevant data, and embed illustrative images directly in responses. It allows users to access both textual explanations and visual context in a single, interactive exchange. Yingying is accessible via the AI Assistant icon on every page or from the homepage. It provides personalized recommendations, facilitates rapid data retrieval, and helps researchers stay up-to-date with the latest developments, significantly improving the efficiency of sup-tRNA research and applications.

## Discussion and conclusions

ENSURE is the first platform to systematically integrate information on both natural and engineered sup-tRNAs, addressing multiple gaps in this emerging field. By aggregating curated data from HGVdb, COSMIC, RNAcentral, GtRNAdb, and hundreds of peer-reviewed publications, we have developed a knowledge base comprising interconnected modules. This framework combines comprehensive sequence and functional annotations with computational 3D models generated with AF3 for illustrative visualization, enabling users to examine sup-tRNAs from sequence to structure and from molecular details to broader functional contexts.

Compared with existing resources, ENSURE offers three major advantages. First, it is dedicated exclusively to sup-tRNAs, with a clear separation between natural and engineered types, and includes only experimentally validated entries to ensure reliability. Second, it integrates detailed functional annotations, such as tRNA modifications, aminoacyl-tRNA synthetase (aaRS), and EF-Tu recognition sites, mapped precisely onto secondary structures using Sprinzl numbering and MODOMICS standards for intuitive interpretation. Third, it incorporates an AI assistant, built on an RAG framework, to deliver both textual and visual information, greatly enhancing query efficiency and user experience.

In recent years, our team has also developed specialized ncRNA resources such as tModBase, which deciphers the landscape of tRNA modifications and their dynamic changes, and starBase or ENCORI, a comprehensive platform for decoding the Encyclopedia of RNA Interactomes [59–61]. ENSURE complements these resources by systematically focusing on suppressor tRNAs, thereby extending our integrated RNA research framework and filling a critical gap in the broader effort to investigate the biological functions and mechanisms of diverse tRNA species in both fundamental and translational contexts.

Nevertheless, ENSURE has current limitations. To ensure experimentally validated accuracy, the database deliberately prioritizes reliability over quantity. As a result, the identification of natural sup-tRNAs still largely depends on published literature, rather than large-scale genome mining or high-throughput experimental screening. Additionally, the functional sup-tRNAs remain to be discovered. Predicted 3D structures of engineered sup-tRNAs are generated solely by AF3 and lack experimental validation, which may affect accuracy for specific variants. These AF3-derived models are provided as illustrative visualizations to aid user intuition and are not used to support quantitative or mechanistic claims. Each entry records the modeling method and version, as well as the exact input sequence. As experimentally determined structures

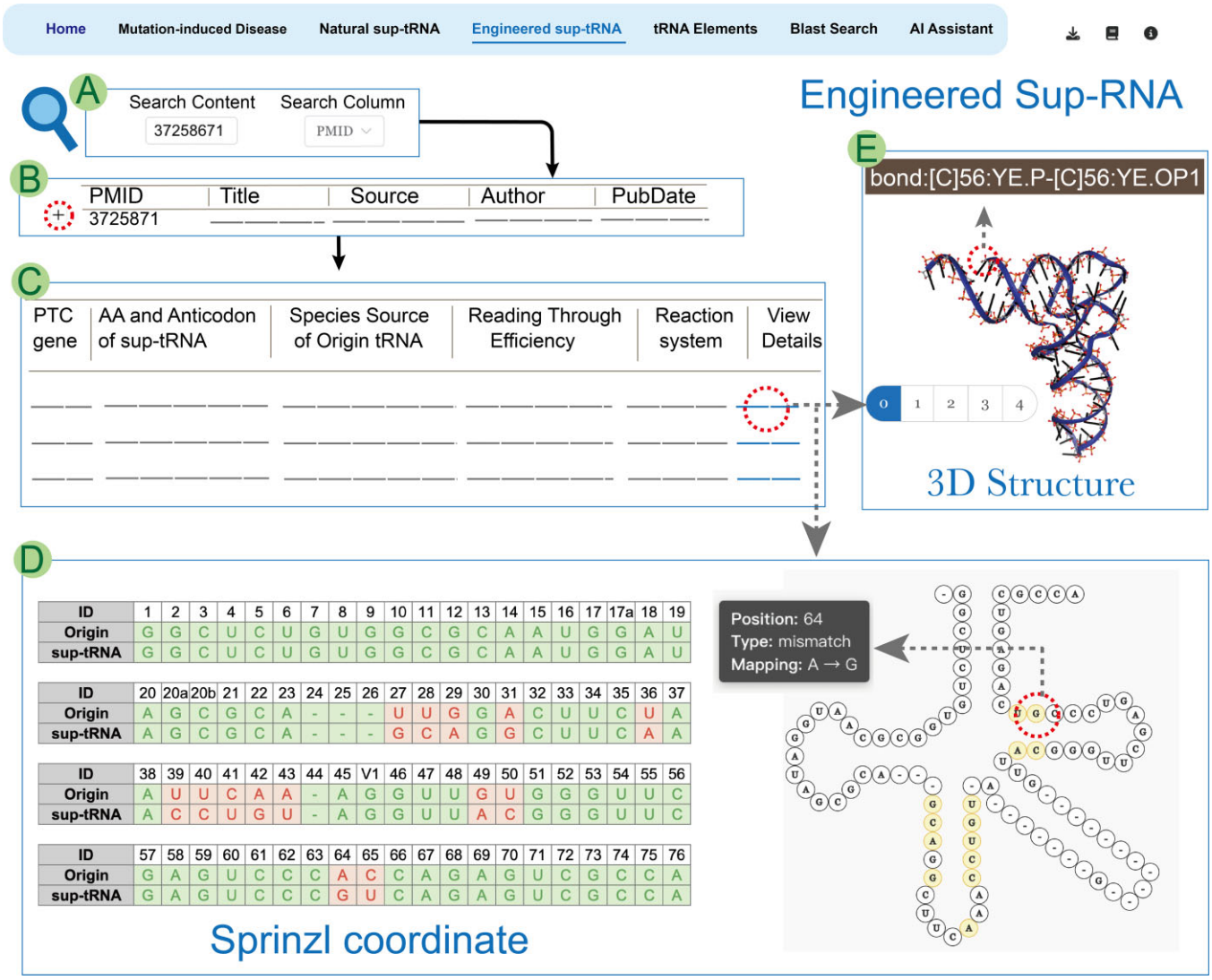



**Figure 4.** Engineered sup-tRNA module in the ENSURE platform. (**A**) Search interface showing a query example using the PMID "37258671." (**B**) Filterable result list. (**C**) Detail panel showing full features of the engineered sup-tRNA. (**D**) Secondary structure schematic highlighting precise modification sites. (**E**) 3D structural model predicted by AF3. It is rotatable and zoomable.

become available, we will integrate them and mark them as superseding AF3 predictions. Future development will focus on expanding coverage through continuous literature curation, database updates, and collaborations with experimental laboratories to incorporate high-sensitivity detection and empirical structural characterization.

In summary, ENSURE delivers a comprehensive, accurate, and user-friendly knowledge base for sup-tRNA research, bridging fundamental biology with the design needs of synthetic biology and RNA therapy. As the dataset grows and supporting technologies advance, ENSURE is positioned to become an indispensable resource for accelerating the development and application of sup-tRNA.

## Acknowledgements

We gratefully acknowledge Professor Vyacheslav Yurchenko (University of Ostrava) and Professor Jianhua Yang (Sun Yat-sen University) for their constructive suggestions regarding this research. We acknowledge computing support from the National Supercomputer Center in Guangzhou (Tianhe-2A) and from the High-performance Computing Public Platform (Shenzhen Campus) of Sun Yat-sen University.

*Author contributions*: Lingling Zheng: Conceptualization; Supervision; Project administration; Resources; Funding acquisition; Writing—review & editing. Xiao Feng: Conceptualization; Supervision; Project administration; Methodology; Writing—review & editing. Zhuo Ouyang: Software (web back-end/front-end and database implementation); Visualization (web interfaces); Writing—original draft; Writing—review & editing. Yifeng Zhang: Data curation (lead); Project administration (coordination of data collection); Investigation; Writing—review & editing. Fan Feng: Data curation; Validation; Investigation; Writing—review & editing. Xudong Zeng: Data curation; Software (analysis scripts/processing pipeline); Validation; Methodology. Qiuhui Wu: Visualization (figure design and preparation). Abdul Hafeez: Writing—review & editing (proofreading). Wenkai Teng: Data curation; Validation; Investigation. Yixin Kong: Data curation; Validation; Investigation. Xuan Bu: Data curation; Validation; Investigation. Yang Sun: Data curation; Validation; Investigation. Bin Li: Writing—review & editing. Yanzi Wen: Writing—

review & editing. Zhaorong Lun: Writing—review & editing. Lianghu Qu: Supervision; Resources; Funding acquisition; Writing—review & editing.

Zhuo Ouyang and Yifeng Zhang contributed equally to this work.

## Conflict of interest

None declared.

## Funding

This research was funded by the National Key R&D Program of China (2022YFC3400401); the National Natural Science Foundation of China (32470599, 32270604, 32500190); Guangdong Province (2021A1515010542, 2022A1515011321). Funding to pay the Open Access publication charges for this article was provided by National Key R&D Program of China.

## Data availability

ENSURE is freely available at https://trna.lumoxuan.cn/. This website is free, open to all users, and no login or password is required.

## References

1. Awawdeh A, Radecki AA, Vargas-Rodriguez O. Suppressor tRNAs at the interface of genetic code expansion and medicine. *Front Genet* 2024;15:1420331. https://doi.org/10.3389/fgene.2024.1420331
2. Albers S, Allen EC, Bharti N *et al.* Engineered tRNAs suppress nonsense mutations in cells and *in vivo*. *Nature* 2023;618:842–8. https://doi.org/10.1038/s41586-023-06133-1
3. Ogawa A. Suppressor tRNA-based biosensors for detecting analytes. *Anal Sci* 2021;37:407–14. https://doi.org/10.2116/analsci.20SCR01
4. Porter JJ, Heil CS, Lueck JD. Therapeutic promise of engineered nonsense suppressor tRNAs. *Wiley Interdiscip Rev:RNA* 2021;12:e1641. https://doi.org/10.1002/wrna.1641
5. Chen C, Yu G, Huang Y *et al.* Genetic-code-expanded cell-based therapy for treating diabetes in mice. *Nat Chem Biol* 2022;18:47–55. https://doi.org/10.1038/s41589-021-00899-z
6. Wang Y, Zhang J, Han B *et al.* Noncanonical amino acids as doubly bio-orthogonal handles for one-pot preparation of protein multiconjugates. *Nat Commun* 2023;14:974. https://doi.org/10.1038/s41467-023-36658-y
7. Hanyu N, Kuchino Y, Nishimura S *et al.* Dramatic events in ciliate evolution: alteration of UAA and UAG termination codons to glutamine codons due to anticodon mutations in two Tetrahymena tRNAsGln. *EMBO J* 1986;5:1307–11. https://doi.org/10.1002/j.1460-2075.1986.tb04360.x
8. Kao S, McClain WH. UGA suppressor of bacteriophage T4 associated with arginine transfer RNA. *J Biol Chem* 1977;252:8254–7. https://doi.org/10.1016/S0021-9258(17)40964-1
9. Kolosova O, Zgadzay Y, Stetsenko A *et al.* Mechanism of read-through enhancement by aminoglycosides and mefloquine. *Proc Natl Acad Sci USA* 2025;122:e2420261122. https://doi.org/10.1073/pnas.2420261122
10. Welch EM, Barton ER, Zhuo J *et al.* PTC124 targets genetic disorders caused by nonsense mutations. *Nature* 2007;447:87–91. https://doi.org/10.1038/nature05756
11. Ko W, Porter JJ, Spelier S *et al.* ACE-tRNAs are a platform technology for suppressing nonsense mutations that cause cystic fibrosis. *Nucleic Acids Res* 2025;53:gkaf675. https://doi.org/10.1093/nar/gkaf675
12. Diamond A, Dudock B, Hatfield D. Structure and properties of a bovine liver UGA suppressor serine tRNA with a tryptophan anticodon. *Cell* 1981;25:497–506. https://doi.org/10.1016/0092-8674(81)90068-4
13. Tate JG, Bamford S, Jubb HC *et al.* COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res* 2019;47:D941–7. https://doi.org/10.1093/nar/gky1015
14. Mort M, Ivanov D, Cooper DN *et al.* A meta-analysis of nonsense mutations causing human genetic disease. *Hum Mutat* 2008;29:1037–47. https://doi.org/10.1002/humu.20763
15. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res* 2001;11:863–74. https://doi.org/10.1101/gr.176601
16. Costello A, Peterson AA, Lanster DL *et al.* Efficient genetic code expansion without host genome modifications. *Nat Biotechnol* 2025;43:1116–27. https://doi.org/10.1038/s41587-024-02385-y
17. Spinck M, Guppy A, Chin JW. Automated orthogonal tRNA generation. *Nat Chem Biol* 2025;21:657–67. https://doi.org/10.1038/s41589-024-01782-3
18. Hermann A, Hiller E, Hubel P *et al.* Genetic code expansion for controlled surfactin production in a high cell-density *Bacillus subtilis* strain. *Microorganisms* 2025;13:353. https://doi.org/10.3390/microorganisms13020353
19. Robertson WE, Rehm FBH, Spinck M *et al.* *Escherichia coli* with a 57-codon genetic code. *Science* 2025;0:eady4368. https://doi.org/10.1126/science.ady4368
20. Lajoie MJ, Rovner AJ, Goodman DB *et al.* Genomically recoded organisms expand biological functions. *Science* 2013;342:357–60. https://doi.org/10.1126/science.1241459
21. Grome MW, Nguyen MTA, Moonan DW *et al.* Engineering a genomically recoded organism with one stop codon. *Nature* 2025;639:512–21. https://doi.org/10.1038/s41586-024-08501-x
22. Kachale A, Pavlíková Z, Nenarokova A *et al.* Short tRNA anticodon stem and mutant eRF1 allow stop codon reassignment. *Nature* 2023;613:751–8. https://doi.org/10.1038/s41586-022-05584-2
23. The RNAcentral Consortium, Petrov AI, Kay SJE *et al.* RNAcentral: a comprehensive database of non-coding RNA sequences. *Nucleic Acids Res* 2017;45:D128–34. https://doi.org/10.1093/nar/gkw1008
24. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* 2016;44:D184–9. https://doi.org/10.1093/nar/gkv1309
25. Wang J, Zhang Y, Mendonca CA *et al.* AAV-delivered suppressor tRNA overcomes a nonsense mutation in mice. *Nature* 2022;604:343–8. https://doi.org/10.1038/s41586-022-04533-3
26. Shi N, Yang Q, Zhang H *et al.* Restoration of dystrophin expression in mice by suppressing a nonsense mutation through the incorporation of unnatural amino acids. *Nat Biomed Eng* 2022;6:195–206. https://doi.org/10.1038/s41551-021-00774-1
27. Kim Y, Cho S, Kim J-C *et al.* tRNA engineering strategies for genetic code expansion. *Front Genet* 2024;15:1373250. https://doi.org/10.3389/fgene.2024.1373250
28. Tamaki S, Tomita M, Suzuki H *et al.* Systematic analysis of the binding surfaces between tRNAs and their respective aminoacyl tRNA synthetase based on structural and evolutionary data. *Front Genet* 2018;8:227. https://doi.org/10.3389/fgene.2017.00227
29. Sprinzl M, Horn C, Brown M *et al.* Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* 1998;26:148–53. https://doi.org/10.1093/nar/26.1.148
30. Katoh K, Misawa K, Kuma K *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;30:3059–66. https://doi.org/10.1093/nar/gkf436

31. Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol* 2019;1962:1–14.
32. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
33. Varadi M, Anyango S, Deshpande M *et al.* AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 2022;50:D439–44. https://doi.org/10.1093/nar/gkab1061
34. Cappannini A, Ray A, Purta E *et al.* MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Res* 2024;52:D239–44. https://doi.org/10.1093/nar/gkad1083
35. Li H, Wang G, Ye C *et al.* Quantitative RNA pseudouridine maps reveal multilayered translation control through plant rRNA, tRNA and mRNA pseudouridylation. *Nat Plants* 2025;11:234–47. https://doi.org/10.1038/s41477-024-01894-7
36. Lu J-L, Dai Y, Ji K *et al.* Taurine hypomodification underlies mitochondrial tRNATrp-related genetic diseases. *Nucleic Acids Res* 2024;52:13351–67. https://doi.org/10.1093/nar/gkae854
37. Yusupov MM, Yusupova GZ, Baucom A *et al.* Crystal structure of the ribosome at 5.5 A resolution. *Science* 2001;292:883–96. https://doi.org/10.1126/science.1060089
38. Lewis P, Perez E, Piktus A *et al.* Retrieval-augmented generation for knowledge-intensive NLP tasks. *Adv Neural Inf Process Syst* 2020;33:9459–74.
39. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. In: Inui K, Jiang J, Ng V, Wan X, (eds.), *(eds.)*, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong: Association for Computational Linguistics, 2019, pp. 3982–92.
40. Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421. https://doi.org/10.1186/1471-2105-10-421
41. Kerpedjiev P, Hammer S, Hofacker IL. Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. *Bioinformatics* 2015;31:3377–9. https://doi.org/10.1093/bioinformatics/btv372
42. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res* 2015;43:W576–9. https://doi.org/10.1093/nar/gkv402
43. Bostock M, Ogievetsky V, Heer J; $D^3$ data-driven documents. *IEEE Trans Vis Comput Graph* 2011;17:2301–9. https://doi.org/10.1109/TVCG.2011.185
44. Miao S, Li H, Song X *et al.* tRNA m1A modification regulates cholesterol biosynthesis to promote antitumor immunity of $CD8^+$ T cells. *J Exp Med* 2025;222:e20240559. https://doi.org/10.1084/jem.20240559
45. Xiong Q-P, Li J, Li H *et al.* Human TRMT1 catalyzes $m^2G$ or $m^2{}_2G$ formation on tRNAs in a substrate-dependent manner. *Sci China Life Sci* 2023;66:2295–309. https://doi.org/10.1007/s11427-022-2295-0
46. Li H, Dong H, Xu B *et al.* A dual role of human tRNA methyltransferase hTrmt13 in regulating translation and transcription. *EMBO J* 2022;41:e108544. https://doi.org/10.15252/embj.2021108544
47. Nagata N, Kurosaka H, Higashi K *et al.* Characteristic craniofacial defects associated with a novel USP9X truncation mutation. *Hum Genome Var* 2024:11:21. https://doi.org/10.1038/s41439-024-00277-w
48. Niida Y, Kobayashi A, Togi S *et al.* Recessive dystrophic epidermolysis bullosa caused by a novel COL7A1 variant with isodisomy. *Hum Genome Var* 2023;10:29. https://doi.org/10.1038/s41439-023-00257-6
49. Schmitz R, Young RM, Ceribelli M *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* 2012;490:116–20. https://doi.org/10.1038/nature11378
50. Ojha J, Ayres J, Secreto C *et al.* Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. *Blood* 2015:125:492–8. https://doi.org/10.1182/blood-2014-06-580563
51. Shain AH, Garrido M, Botton T *et al.* Exome sequencing of desmoplastic melanoma identifies recurrent NFKBIE promoter mutations and diverse activating mutations in the MAPK pathway. *Nat Genet* 2015;47:1194–9. https://doi.org/10.1038/ng.3382
52. Miao D, Margolis CA, Vokes NI *et al.* Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet* 2018;50:1271–81. https://doi.org/10.1038/s41588-018-0200-2
53. Panea RI, Love CL, Shingleton JR *et al.* The whole-genome landscape of Burkitt lymphoma subtypes. *Blood* 2019;134:1598–607. https://doi.org/10.1182/blood.2019001880
54. Čapková Pavlíková Z, Miletínová P, Roithová A *et al.* Ribosomal A-site interactions with near-cognate tRNAs drive stop codon readthrough. *Nat Struct Mol Biol* 2025;32:662–74.
55. Galan A, Kraeva N, Záhonová K *et al.* Converting blastocrithidia nonstop, a trypanosomatid with non-canonical genetic code, into a genetically-tractable model. *Mol Microbiol* 2025;123:586–92. https://doi.org/10.1111/mmi.15365
56. Luo N, Huang Q, Dong L *et al.* Near-cognate tRNAs increase the efficiency and precision of pseudouridine-mediated readthrough of premature termination codons. *Nat Biotechnol* 2025;43:114–23. https://doi.org/10.1038/s41587-024-02165-8
57. Yuan C, Li Z, Luo X *et al.* Mammalian trans-editing factor ProX is able to deacylate tRNAThr mischarged with alanine. *Int J Biol Macromol* 2023;253:127121. https://doi.org/10.1016/j.ijbiomac.2023.127121
58. Li Z-H, Zhou X-L. Eukaryotic AlaX provides multiple checkpoints for quality and quantity of aminoacyl-tRNAs in translation. *Nucleic Acids Res* 2024;52:7825–42. https://doi.org/10.1093/nar/gkae486
59. Wang J-H, Chen W-X, Mei S-Q *et al.* tsRFun: a comprehensive platform for decoding human tsRNA expression, functions and prognostic value by high-throughput small RNA-seq and CLIP-seq data. *Nucleic Acids Res* 2022;50:D421–31. https://doi.org/10.1093/nar/gkab1023
60. Lei H-T, Wang Z-H, Li B *et al.* tModBase: deciphering the landscape of tRNA modifications and their dynamic changes from epitranscriptome data. *Nucleic Acids Res* 2023;51:D315–27. https://doi.org/10.1093/nar/gkac1087
61. Li J-H, Liu S, Zhou H *et al.* starBase v2.0: decoding miRNA–ceRNA, miRNA–ncRNA and protein–RNA interaction networks from large-scale CLIP-seq data. *Nucleic Acids Res* 2014;42:D92–7. https://doi.org/10.1093/nar/gkt1248


**Received:** August 14, 2025. **Revised:** September 16, 2025. **Accepted:** September 29, 2025