



Metagenomic mining reveals novel Cas12 subtypes and their evolutionary diversification

Meixia Yu^{1#}, Zilong Zhao^{1#}, Peiyang Li¹, Yunpeng Teng¹, Feng Xu¹,
Zhimeng Li¹, Zhijia Yang², Hainan Zhao², Huanbin Zhou^{3,4,5},
Lanqin Xia^{4,6}, Jinsheng Lai^{1,2,7,8,9}, Beibei Xin^{1,9*}

¹State Key Laboratory of Maize Bio-breeding, National Maize Improvement Center, Department of Plant Genetics and Breeding, China Agricultural University, Beijing 100193, China

²Frontiers Science Center for Molecular Design Breeding (Ministry of Education), China Agricultural University, Beijing 100193, China

³State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China

⁴Ministry of Agriculture and Rural Affairs Key Laboratory of Gene Editing Technologies (Hainan), National Nanfan Research Institute, Chinese Academy of Agricultural Sciences, Sanya 572024, China

⁵Scientific Observing and Experimental Station of Crop Pests in Guilin, Ministry of Agriculture and Rural Affairs, Guilin 541399, China

⁶State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences (ICS), Chinese Academy of Agricultural Sciences (CAAS), Beijing 100081, China

⁷Center for Crop Functional Genomics and Molecular Breeding, China Agricultural University, Beijing 100193, China

⁸Sanya Institute of China Agricultural University, Sanya 572025, China

⁹Ministry of Agriculture and Rural Affairs Key Laboratory of Gene Editing Technologies (Beijing), Beijing 100193, China

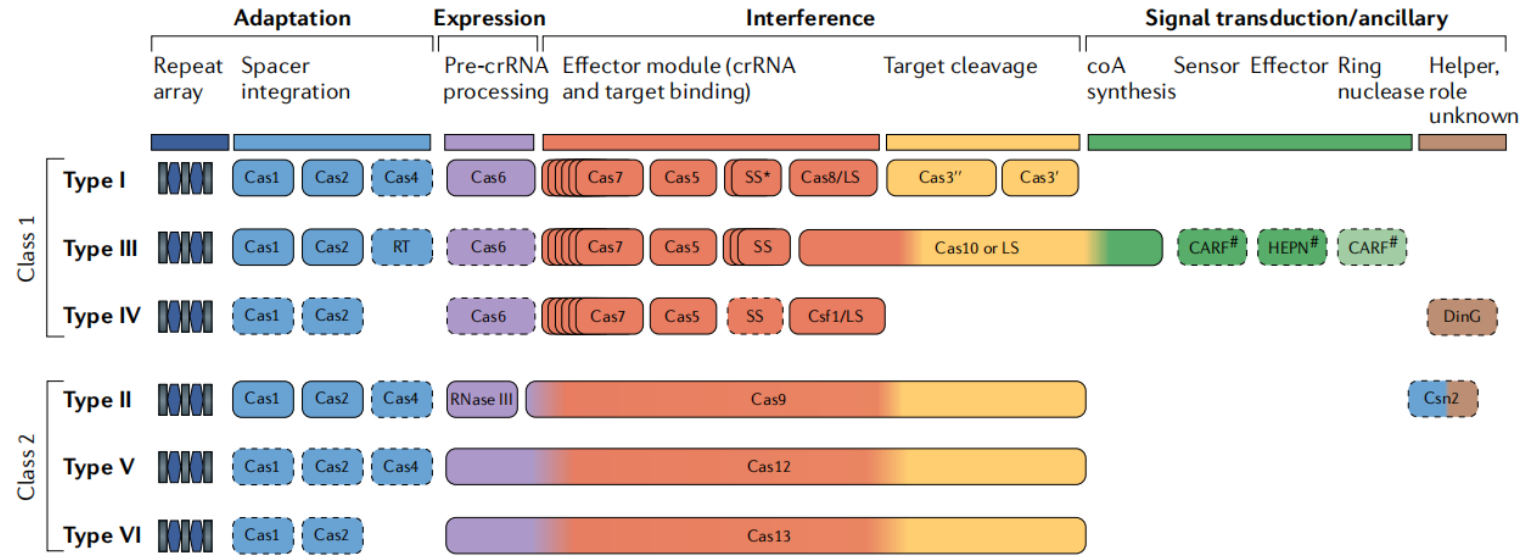
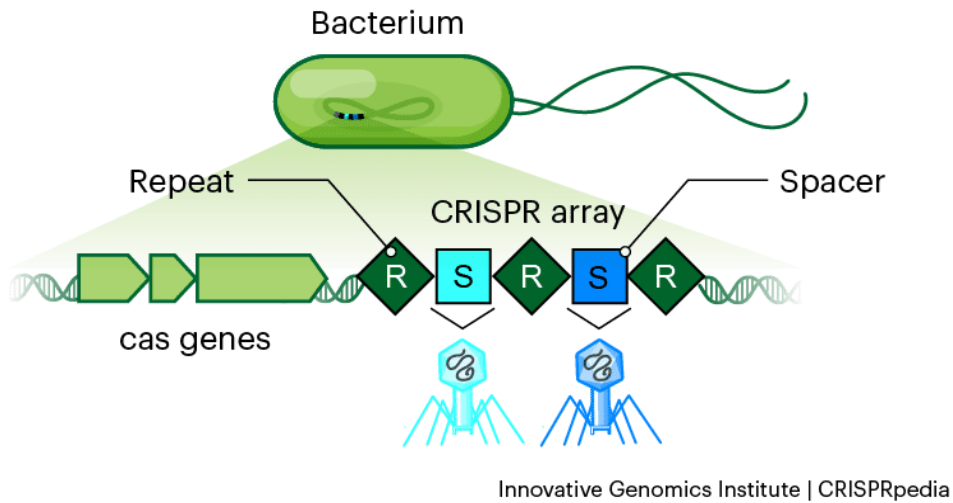


Meixia Yu, Zilong Zhao, Peiyang Li, Yunpeng Teng, Feng Xu, Zhimeng Li, Zhijia Yang, et al. 2025.
Metagenomic mining reveals novel Cas12 subtypes and their evolutionary diversification.

iMetaOmics 2: e70057. <https://doi.org/10.1002/imo2.70057>



Introduction

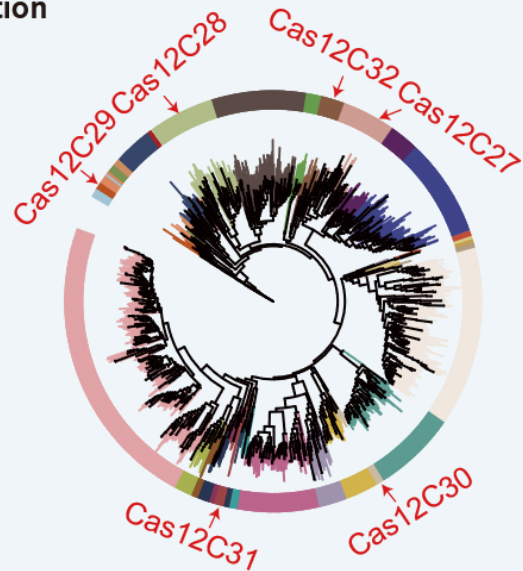


Subtype	Cas12a	Cas12b	Cas12c	Cas12d (casY)	Cas12e (casX)	Cas12f	Cas12g	Cas12h	Cas12i	Cas12j	Cas12k	Cas12l (casλ)	Cas12l (casπ)	Cas12m	Cas12n	
Length	~1300aa	~1100aa	1209-1330aa	~1200aa	~980aa	400-700aa	767aa	870-933aa	1033-1093aa	700-800aa	500-700aa	~700aa	~860aa	~600aa	400-700aa	
tracrRNA	×	√	scoutRNA	scoutRNA	√	√	√	×	×	×	√	×	√	×	√	
Cleavage Activity	dsDNA Pre-crRNA	dsDNA	dsDNA	dsDNA	dsDNA	ssDNA dsDNA	RNA ssDNA (Non-specific)	dsDNA	dsDNA	dsDNA	dsDNA Pre-crRNA	None	dsDNA	dsDNA	None	dsDNA
Adapt module	Cas1,Cas2, Cas4	Cas1-Cas4, Cas2	Cas1	Cas1	Cas1,Cas2, Cas4	Cas1*,Cas2*, Cas4*	None	None	None	None	Tn7-like	None	Cas1,Cas2, Cas4	None	None	
PAM	AT-rich	AT-rich	T-rich	5'-TA	5'-TA	ssDNA (None) dsDNA (T-rich)	No requirement	TTN	TTN	TTN	T-rich	GTN	TTR	CCN	TTN	AAN



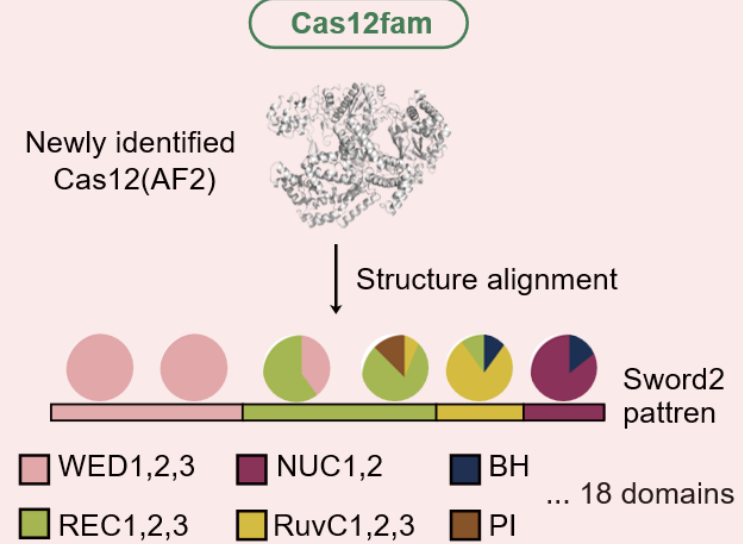
Highlights

Identification

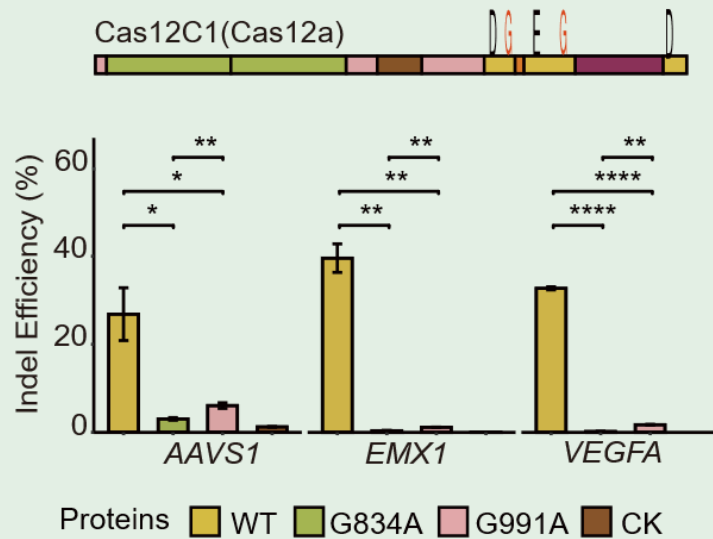


4112 proteins, including six new subtypes

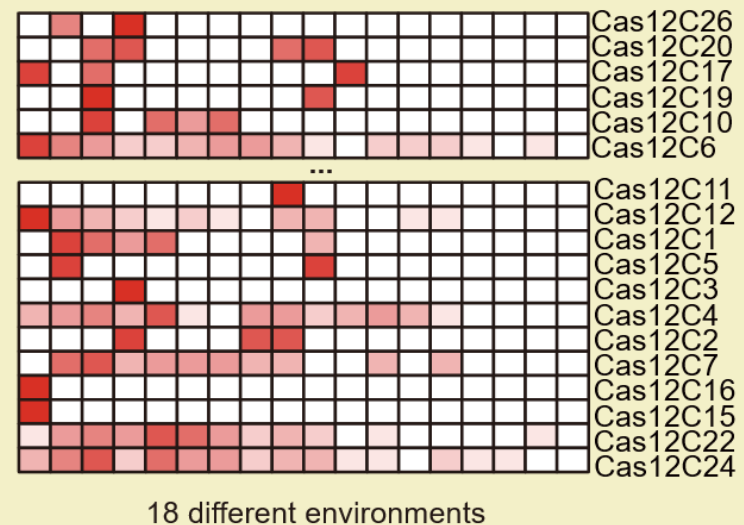
Annotation



Features

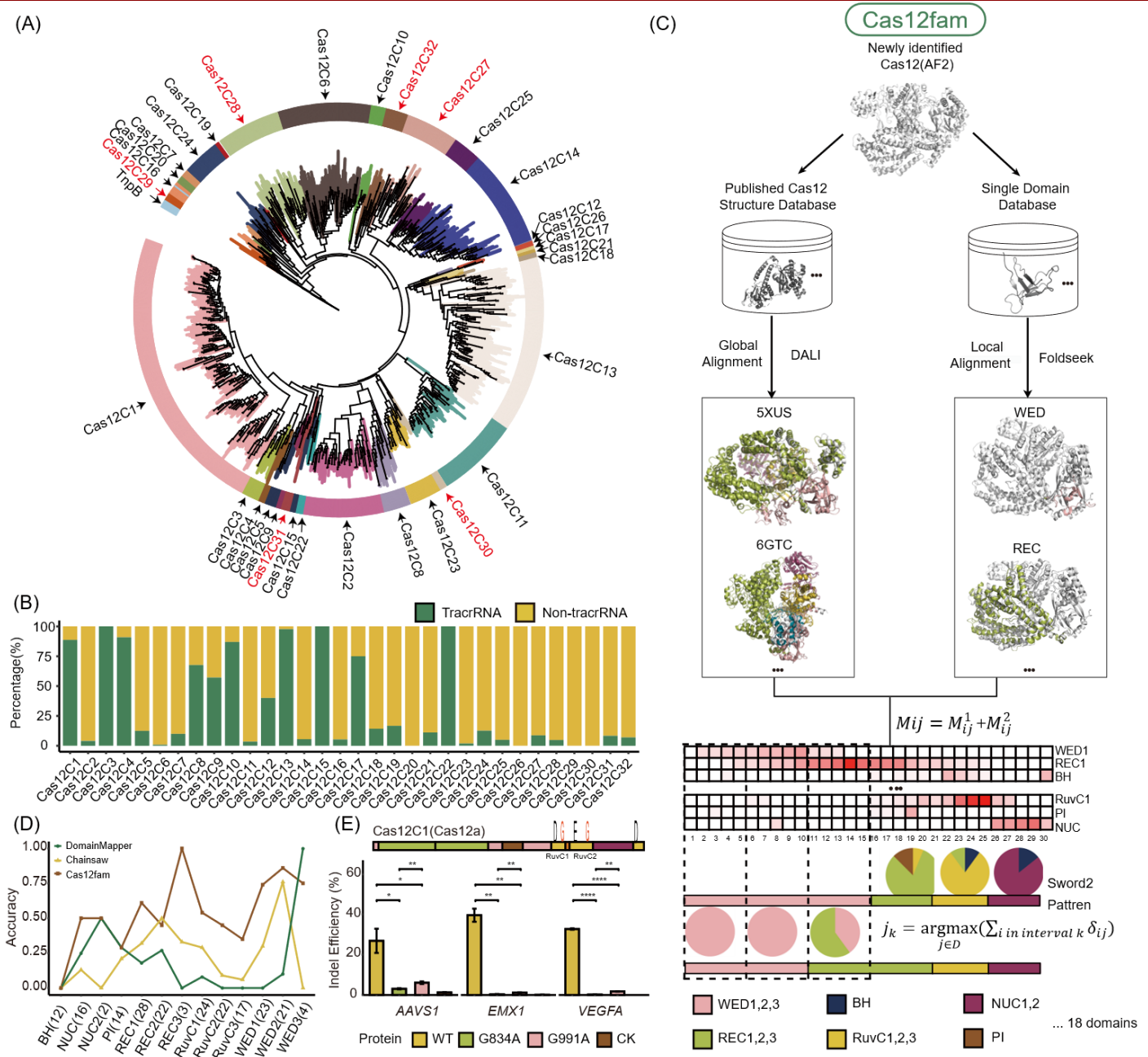


Adaptability





Cas12 Subtypes Nomenclature, Diversity and Annotation



- Identified **4112 Cas12 homologs** and proposed an expandable “**Cas12Cn**” nomenclature
- Revealed **high diversity** in Cas12 family (avg. sequence similarity 30.54%) and **tracrRNA dependence** linked to conserved motifs
- Developed an **AI-based domain annotation** tool, accurately identifying **18 domains** and **outperforming** existing tools
- Discovered **two conserved glycine residues** in the RuvC domain; mutations greatly affect **Cas12a cleavage activity**

Figure 1. Nomenclature of novel Cas12 subtypes and the annotation workflow of the Cas12fam software.



Ecology and Evolution of Cas12 Subtypes

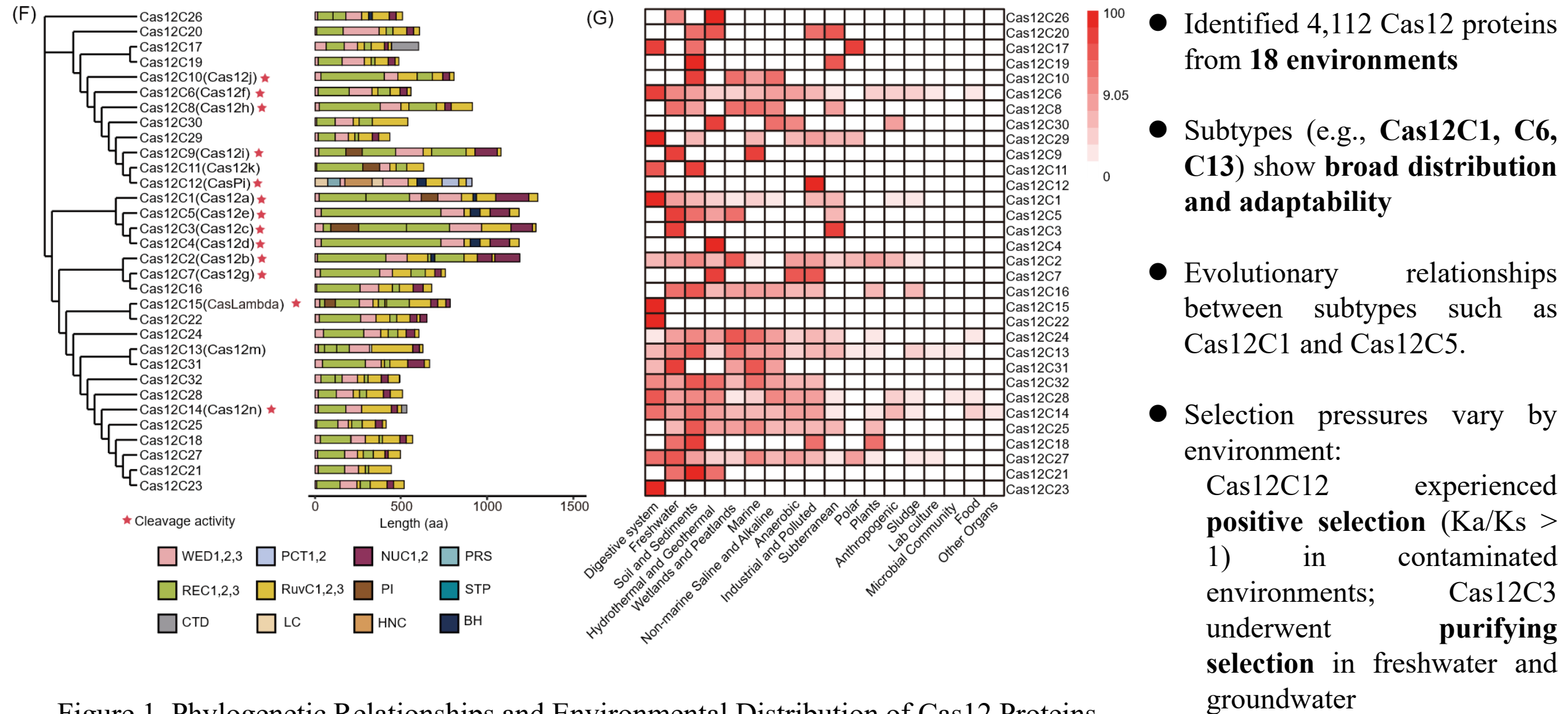


Figure 1. Phylogenetic Relationships and Environmental Distribution of Cas12 Proteins



Summary

- ❑ In this study, **4,112 Cas12 sequences** were identified from 8 TB of microbial genomic data, expanding known subtype diversity. An extensible naming scheme (**Cas12Cn**) was proposed to resolve classification inconsistencies.
- ❑ We developed **Cas12fam**, an AI-based tool that accurately annotates **18 functional domains**, offering a unified standard for future Cas12 research.
- ❑ Cas12 was found enriched in **digestive, anaerobic, and contaminated environments**, with untapped potential in other niches. Environment-specific Ka/Ks ratios confirmed **ecology as a major driver of Cas12 diversity**.
- ❑ Integration of **domain, evolution, and ecology** provides a framework for efficient screening of highly active and specific novel Cas12 subtypes.

Meixia Yu, Zilong Zhao, Peiyang Li, Yunpeng Teng, Feng Xu, Zhimeng Li, Zhijia Yang, et al. 2025.

Metagenomic mining reveals novel Cas12 subtypes and their evolutionary diversification.

iMetaOmics 2: e70057. <https://doi.org/10.1002/imo2.70057>

iMeta: To be top journals in biology and medicine

WILEY



“**iMeta**” launched in 2022 by iMeta Science Society, **impact factor (IF) 33.2**, ranking **top 65/22249 in world and 2/161 in the microbiology**. It aims to publish innovative and high-quality papers with broad and diverse audiences. **Its scope is similar to Cell, Nature Biotechnology/Methods/Microbiology/Medicine/Food**. Its unique features include video abstract, bilingual publication, and social media with 600,000 followers. Indexed by **SCIE/ESI, PubMed, Google Scholar** etc.

“**iMetaOmics**” launched in 2024, with a **target IF>10**, and its scope is similar to **Nature Communications, Cell Reports, Microbiome, ISME J, Nucleic Acids Research, Briefings in Bioinformatics**, etc.

“**iMetaMed**” launched in 2025, with a **target IF>15**, similar to **Med, Cell Reports Medicine, eBioMedicine, eClinicalMedicine** etc.



Society: <http://www.imeta.science>

Publisher: <https://wileyonlinelibrary.com/journal/imeta>

iMeta: <https://wiley.atyponrex.com/journal/IMT2>

Submission: iMetaOmics: <https://wiley.atyponrex.com/journal/IMO2>

iMetaMed: <https://wiley.atyponrex.com/journal/IMM3>



[iMetaScience](#)



[iMetaScience](#)



office@imeta.science
imetaomics@imeta.science



[Promotion Video](#)

Update
2025/7/6