

New classification of prokaryotic DNA methyltransferases

Samokhina M.^{1, 2*}, Alexeevski A.^{1, 2, 3}

¹ Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia

² Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia

³ Research Institute for System Analysis, RAS, Moscow, Russia

* mashila6799@gmail.com

Key words: DNA methyltransferases; classification; evolution

Motivation and Aim: Prokaryotic DNA methyltransferases (MTases) are essential enzymes in bacterial immune systems like restriction and modification systems and others. Also the enzymes play an important role in regulation gene expression, cell cycle and DNA repair. MTase classification by sequence similarity was built more than 30 years ago [1]. Our goal was to build new MTases classification that describe modern diversity of MTases and construct an evolutionary model.

Methods and Algorithms: To built classification we use MTase sequences from database REBASE. For splitting MTases into groups we used HMM profiles from Pfam and SUPFAM databases. Profile hits were used to classify MTases into classes.

Results: We divided all MTases into 10 classes based on profile region hits (Fig. 1). The classes are named in Latin letters from A to J. Each class has a specific topology. Topologies A–D obtained from data on resolved spatial structures, the remaining topologies are obtained from the predicted structures from AlphaFold database. We identified catalytic motifs in all MTases. Based on the similarity of catalytic motifs, MTases are divided into 3 evolutionary groups. Within each group, the sequences of catalytic motifs intersect, but between groups they do not intersect.

We have developed an automatic algorithm that detects the topology of MTase by region hits of cat-profiles in MTase sequences (<https://github.com/MVolobueva/MTase-classification.git>) and a web interface to run it (<https://mtase-pipeline-6g1yf9ugw8.streamlit.app/>) [2].

We have compared the developed classification with existing ones, namely into methylation types, restriction-modification system (R-M) types they belong and permutations of S-Adenosyl methionine (SAM)-binding motifs, catalytic motifs and TRD (target recognition domain) (Table 1). MTases from class C and K have m5C methylation type. MTases from the P-M I system are found only in class A, MTases from the P-M III system are found only in class B.

The comparison with the classification by motifs shows that classes A, L, K, F correspond to class γ , classes M and D correspond to class α . The K class was previously allocated to a separate class ζ^* . However, there is no experimental evidence that class ζ is different from class γ .

Conclusion: We have built new MTase classification based on sequence similarity of catalytic domain. All known MTases fell into 10 classes that comprise 3 evolutionary groups.

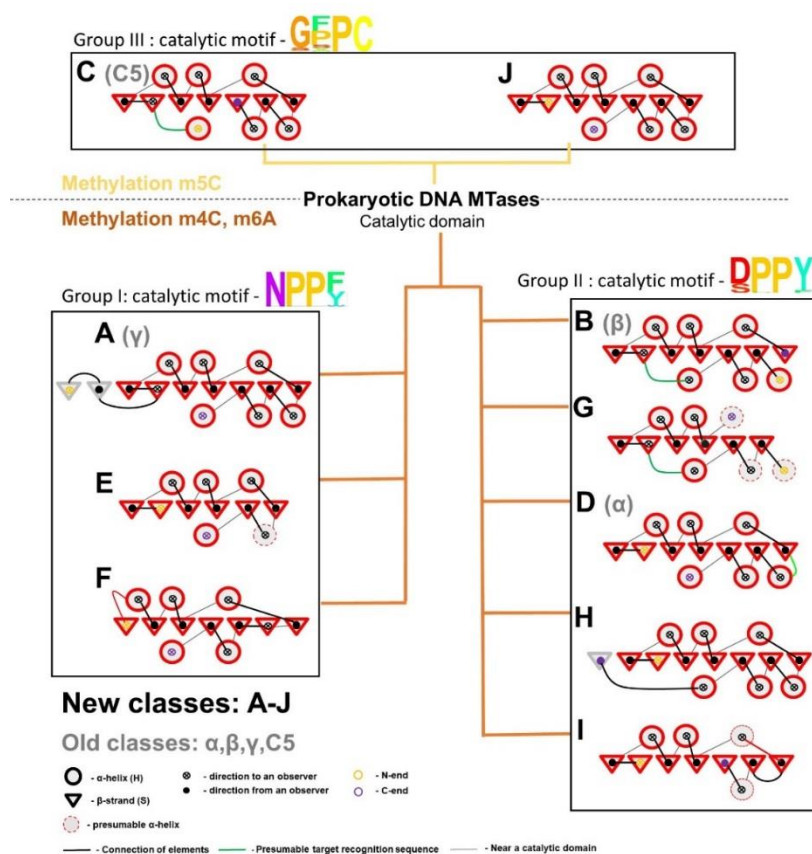


Fig. 1. New classification of MTases. Classes are named in Latin letters from A to J. For each class, a topology scheme is provided, which displays the sequence of connection of the secondary structure elements

Table 1. Comparison of the new classification with the previous classifications

Class	Methylation type	R-M type	Subtype
A	m4C, m6A	I, II	γ
B	m4C, m6A	II, III	β
D	m4C, m6A	II	α
L	m4C, m6A	II	γ
M	m4C, m6A	II	α
C	m5C	II	C5
K	m5C	II	ζ*, γ
E	m6A	II	
F	m6A	II	γ
G	m6A	II	
I	m4C, m6A	II	γ

* Buinitsky [1] classified zeta into a separate class.

References

1. Bujnicki J.M. Sequence permutations in the molecular evolution of DNA methyltransferases. *BMC Evol Biol.* 2002;2:3. doi 10.1186/1471-2148-2-3
2. Samokhina M., Rusinov I., Alexeevski A. Classification of prokaryotic DNA methyltransferases by topology and sequence similarity. *bioRxiv.* 2023. doi 10.1101/2023.12.13.571470