

DNA 拓扑异构酶 II 与 IA 的进化分析

任永明^{1,2}, 唐素妮^{1,2}, 黄京飞^{1,3,*}

(1. 中国科学院昆明动物研究所 细胞与分子进化重点实验室, 云南 昆明 650223;

2. 中国科学院研究生院, 北京 100049; 3. 中国科学院昆明动物研究所 云南省畜禽分子生物学重点实验室, 云南 昆明 650223)

摘要: DNA 拓扑异构酶与断开 DNA 链和改变其拓扑结构有关。先前的研究认为, 拓扑异构酶 Type II B 是由 Type II A 通过基因的复制、重组和缺失所造成。然而, 本研究结果则显示, Type IA 与 Type II B 的进化关系较近, 而 Type II A 与 Type II B 的关系较远。因此, Type II B 可能是由 Type IA 演化而来, 或者说, Type II B 是 Type IA 在古细菌中的特化形式; Type II B 可能是由 Type II A 通过形成 Type IA 后演化而成的, 而不是由 Type II A 直接通过基因的复制、重组和缺失所造成。在由 Type II A 演化成为 Type IA, 最后演化为 Type II B 的过程中, DNA 拓扑异构酶的催化机制也发生了变化: 从需要金属离子的协助, 演化到了不需要镁离子的存在协助其催化。

关键词: DNA 拓扑异构酶; CAP-like 结构域; Toprim 结构域; 分子进化, 催化机制

中图分类号: Q523; Q588 **文献标识码:** A **文章编号:** 0254 – 5853 (2006) 06 – 0595 – 06

Evolutionary Analyses of DNA Topoisomerases II and IA

REN Yong-ming^{1,2}, TANG Su-ni^{1,2}, HUANG Jing-fei^{1,3,*}

(1. Key Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650223, China;

2. Graduate School of the Chinese Academy of Sciences, Beijing 100049, China; 3. Yunnan Key Laboratory of Molecular Biology of

Domestic Animals, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650223, China)

Abstract: DNA strand scission and topology changes are catalyzed by DNA topoisomerases. Previous studies suggest that Type II B topoisomerase evolved from Type II A by gene duplication, recombination and absence. However, this study shows that the evolutionary relationship between Type IA and II B is closer than that between Type II A and II B. Thus, Type II B is possibly from Type IA, or Type II B is a special form of IA in archaeobacteria; Type II B is probably evolved from II A via IA rather than from II A directly. During the evolution of II A to II B via IA, the catalytic mechanisms of DNA topoisomerases have also been changed; that is, metal ions are necessary in the catalyses for both IA and II A, but no ion is needed for II B.

Key words: DNA topoisomerase; CAP-like domain; Toprim domain; Molecular evolution; Catalytic mechanism

DNA 拓扑异构酶 (DNA topoisomerase) 是生命活动中一类重要的酶, 它可催化单链或双链 DNA 断开和结合的偶联反应, 从而改变 DNA 的拓扑结构 (Wang, 1985, 1996; Champoux, 2001)。根据 DNA 链断裂反应过程中产生断口性质的不同, DNA 拓扑异构酶可分为两类: I 类 DNA 拓扑异构酶 (Type I) 引起 DNA 糖 – 磷酸骨架上的单链 DNA 的断裂; 而 II 类 DNA 拓扑异构酶 (Type II)

引起双链 DNA 的断裂。根据作用端的不同, Type I 又可分为两类: IA 类 DNA 拓扑异构酶 (Type IA) 作用于 5' – 端; 而 IB 类 DNA 拓扑异构酶 (Type IB) 作用于 3' – 端。Type IA 包括细菌 Topoisomerase I、Topoisomerase III、古细菌 Reverse gyrase、古细菌 Topo III、真核生物 Topo III (Gangloff et al, 1999; Dai et al, 2003); Type IB 包括真核生物 Topoisomerase I (Topo I)、病毒

* 收稿日期: 2006 – 06 – 30; 接受日期: 2006 – 10 – 17

基金项目: 国家自然科学基金资助项目 (30470939 和 30021004)

* 通讯作者 (Corresponding author), Tel: 0871 – 5195183, E-mail: huangjf@mail.kiz.ac.cn

Topoisomerase I、原核生物 Topoisomerase V (Topo V) (Krogh & Shuman, 2002; Forterre, 2006)。虽然, Type II 都是作用于 5' - 端, 但根据其基因结构和作用机制的不同, 又分为 Type II A 和 Type II B。Type II A 类中的 toprim (topoisomerase/primase) 和 CAP-like (catabolite-activator protein-like) 结构域分别分布于 A 和 B 两个亚基上, 并且反应中需要金属离子参与, 在断开 DNA 双链后产生 4 个碱基的错位, 而 Type II B 类中的 toprim 和 CAP-like 结构域都存在于 A 亚基上, 反应中不需要金属离子参与, 断口只有两个碱基错位 (Buhler et al, 2001)。Type II A 包括真核生物 Topoisomerase II (Topo II)、细菌 DNA gyrase、极少数古细菌 gyrase 以及细菌 Topoisomerase IV (Topo IV) (Wang, 1985, 1996; Champoux, 2001); 而 Type II B 仅包括古细菌 Topo VI (Buhler et al, 2001)。

Type I A、II A 和 II B 具有两个相似的结构域, 在断裂和结合 DNA 的过程中起作用, 一个是 CAP-like 结构域, 由于其折叠方式类似于 *E. coli* 的分解代谢物活化蛋白而得名, 研究表明, CAP-like 结构域不仅存在于 Type I A、II A 和 II B 中, 也存在于真核生物减数分裂重结合因子 (Spo11) 中 (Champoux, 2001; Nichols et al, 1999); 另一个是 toprim 结构域, 不仅存在于 Type I A、II A、II B 中, 而且在细菌的引物酶 (Primase) 和核酸酶中也有发现 (Corbett & Berger, 2004)。此外, 在 Type I A 中发现两个 CAP-like 结构域, 即具有催化功能的 CAP-like (d3) 结构域和不具催化功能的 CAP-like (d4) 结构域, 而 CAP-like (d4) 结构域的二级结构排列方式与 Type II A、II B 的类似 (Liu & Wang, 1999)。先前的研究表明, Type II B 是祖先细胞的 Type II A 的一个或者是两个亚单位在古细菌以及真核生物中通过复制、重组和缺失所造成; 古细菌失去了 Type II A, 保留了 Type II B, 而真核生物失去了 Type II B 中的 B 亚单位, 保留了 A 亚单位, 并且发展成为 Spo11 (Nichols et al, 1999; Corbett & Berger, 2003a)。由于 Type I B 的结构和功能与其他类酶的不同, 显然没有同源性, 鉴于 Type II B 基因在 toprim 和 CAP-like 这两个重要结构域的连接方式与 Type I A、II A 的顺序完全不同, 因此, 本研究以 Type I A、II A 和 II B 中都含有的 toprim 和 CAP-like 结构域为重点, 进一步探讨了 Type I 和 Type II 之间的分子进化关系。

1 材料与方法

本文所用的序列都取自 SWISS-PROT 蛋白质序列数据库 (Bairoch & Apweiler, 2000), 结构域的定义来自于 SCOP 数据库 (Murzin & Brenner, 1995), 所用到的基因组信息均来自于 GenBank 数据库 (Benson et al, 1997)。序列用 Clustal W 程序 (Thompson et al, 1994) 进行比对, 比对结果利用 PHYLIP 软件包构建 NJ 树进行系统演化分析 (Felsenstein, 1996)。各个结构域在各种生物基因中的分布, 运用 PSI-BLAST (E 值 ≤ 0.005) 对美国国立生物技术信息中心 (NCBI) 的微生物基因组数据库 (Altschul et al, 1997) 进行搜索后确定。运用 COMPARE on-line 程序 (Sali & Blundell, 1990) 对引物酶和三类拓扑异构酶中 toprim 结构域进行基于结构的序列比对。

2 结果

基因组搜索结果显示, Type I A 广泛分布于细菌、古细菌、真核生物以及极少数病毒中; Type II A 存在于细菌、古细菌和真核生物中; Type II B 仅存在于古细菌中。

细菌 Type II A 包含两个亚基, 分别由两个基因编码, 即 *gyrA* 和 *gyrB* (在 *gyrase* 中), 或 *parC* 和 *parE* (在 Topo IV 中), 而 toprim 和 CAP-like 结构域分别位于这两个亚基上。真核生物和极少数双链病毒 Type II A 的 toprim 结构域在前, CAP-like 结构域在后, 由一个基因编码。通常, Type I A 具有两个 CAP-like 结构域, 即 d3 和 d4 结构域 (图 1), 然而, 进一步的基因组搜索发现, 许多细菌的 Type I A 往往具有多个序列长度不同的拷贝, 但其中至少有一个拷贝同时具有 d3 和 d4 结构域。

Type II B 包含两个亚基, 由两个基因分别编码, 其 CAP-like 结构域和 toprim 结构域在同一个亚基上, CAP-like 结构域在前, toprim 结构域在后, 这与 Type I A、II A 的顺序完全不同 (图 1)。

在 Type I A、II A 和 II B 中都具有一个结构相似的同源 toprim 结构域, 并且 Type I A 的 d4 结构域与 Type II A、II B 的 CAP-like 结构域的二级结构类似, 其顺序均为 $\alpha_1\beta_1\alpha_2\alpha_3\beta_2\beta_3$; 而 Type I A d3 结构域的二级结构顺序则为 $\beta_1\alpha_2\alpha_3\beta_2\beta_3\alpha_1$ 。

CAP-like 结构域序列间最大相似性的分析显示 (表 1), Type I A 的 d4 结构域与 Type II B 序列间

的相似性最高, 其次为同属于 Type I A 的 d3 与 d4 的相似性, 而 Type II A 与 Type I A 的 d4 和 II B 的相似性基本相同, Type II A 与 I A 的 d3 的相似性最低。同样, toprim 结构域序列间最大相似性分析也表明 (表 2), Type I A 与 Type II B 的相似性最高。Type I A、II A 和 II B 的 CAP-like 结构域的系统演化分析表明, Type II A 与属于 Type I A 的 d3 关系较近, 而 Type II B 与属于 Type I A 的 d4 关系较近 (图 2)。Type I A、II A 和 II B 的 toprim 结构域的系统演化分析则显示, Type I A 与 Type II A 和 II B 的关系基本相同 (图 3)。

对 toprim 结构域进行基于结构的序列比对显示, 除了 4 个保守位点外, 它们的序列差异很大 (图 4), 但这 4 个保守位点均是 DNA 结合相关的活性位点 (Gangloff et al, 1999; Dai et al, 2003; Berger et al, 1998)。

利用 CAP-like 结构域进行基因组搜索, 结果表明, 并不是所有物种中的 Type I A 都同时具有 d3

和 d4 结构域, 有的只具有 d3 结构域, 而有的则只具有 d4 结构域。例如, 在 *Haemophilus influenzae* R2846 中, 其 Type I A (NR 数据库, 检索号为 ZP00155446) 没有发现 d4 结构域, 而在 *Haemophilus somnus* 129PT 的 Type I A (NR 数据库, 检索号为 ZP-00123147) 中则缺少 d3 结构域。

3 讨论

Type I A 和 Type II A 在细菌、古细菌、真核生物中广泛分布, 而 Type II B 则仅分布于古细菌中, 因此, Type I A 和 II A 可能是比 Type II B 较为古老的拓扑异构酶。

虽然先前对 Type II 的 B 亚基上的两个与断裂和结合 DNA 功能无关的拓扑异构酶结构域研究显示, Type II B 是由 Type II A 通过基因的复制、重组和缺失所造成的 (Nichols et al, 1999; Corbett & Berger, 2003b)。然而, 本研究通过对 Type I A 和 Type II 中共同存在的、具有结合或断裂 DNA 功能

表 1 Type I A、II A 和 II B 的 CAP-like 结构域序列间的最大相似性 (%)

Tab. 1 The maximum similarity (%) between CAP-like domain sequences in Type I A, II A and II B

	I Ad3	I Ad4	II A	II B
I Ad3	100	20.0	13.5	13.7
I Ad4	20.0	100	17.3	20.7
II A	13.5	17.3	100	17.0
II B	13.7	20.7	17.0	100

表 2 Type I A、II A 和 II B 的 toprim 结构域序列间的最大相似性 (%)

Tab. 2 The maximum similarity (%) between toprim domain sequences in Type I A, II A and II B

	I A	II A	II B
I A	100	19.7	21.5
II A	19.7	100	21.0
II B	21.5	21.0	100

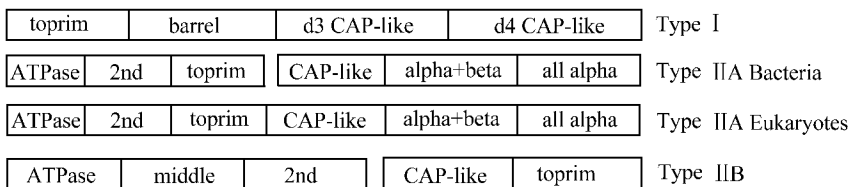


图 1 DNA 拓扑异构酶的结构域排列

Fig. 1 The arrangement of DNA topoisomerase domains

Toprim: toprim 结构域 (toprim domain); barrel: 桶状结构域 (barrel domain); d3 CAP-like: Type I A 类第 3 个结构域 (the third domain of Type I A); d4 CAP-like: Type I A 类第 4 个结构域 (the fourth domain of Type I A); ATPase: Type II A 类 ATP 酶结构域 (the ATPase domain of Type II A); 2nd: II A 类第 2 个结构域 (the second domain of Type II A); CAP-like: CAP-like 结构域 (CAP-like domain); alpha + beta: Type II A 类 $\alpha + \beta$ 结构域 (the alpha + beta domain of Type II A); all alpha: Type II A 类全 α 结构域 (all alpha domain of Type II A); middle: Type II B 类中间结构域 (the middle domain of Type II B)。

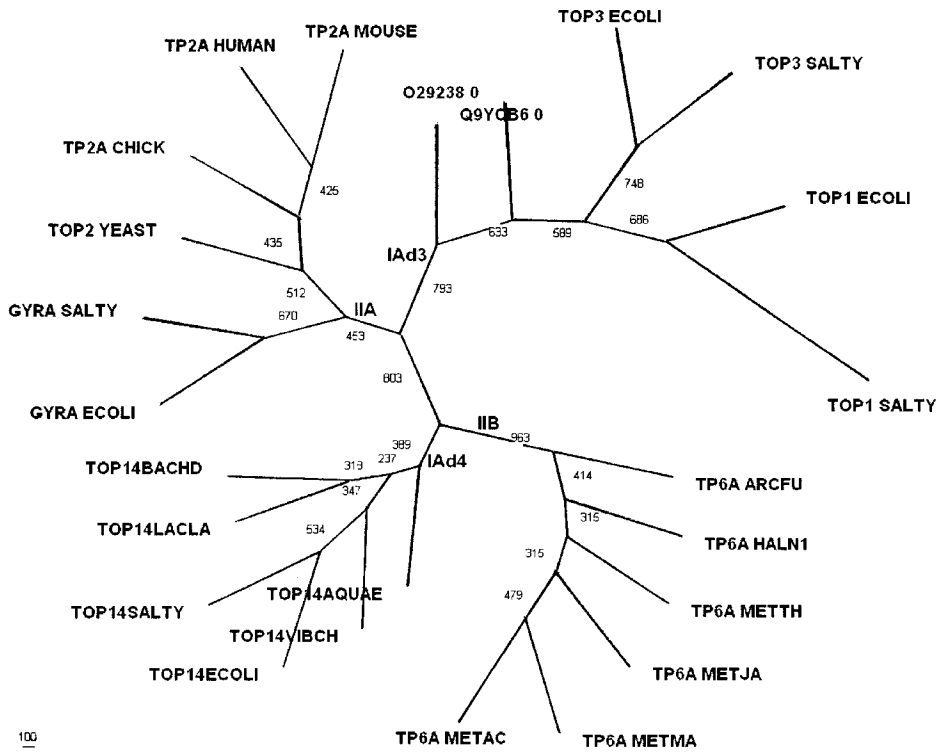


图 2 CAP-like 结构域的进化树 (1 000 bootstrap)

Fig. 2 The phylogenetic tree of CAP-like domains (1 000 bootstrap)

II A: II A 类 (Type II A); II B: II B 类 (Type II B); I Ad3: I A 类第 3 个结构域 (the third domain of Type I A); I Ad4: I A 类第 4 个结构域 (the fourth domain of Type I A)。

的 CAP-like 和 toprim 结构域的比较分析表明, Type I A 与 II A 中 toprim 和 CAP-like 结构域排列的先后顺序是一致的, 而与 Type II B 的明显不同 (图 1): 因而, 从 toprim 和 CAP-like 结构域排列顺序来看, Type I A 和 Type II A 与 Type II B 都应该相距较远。然而, 序列和系统演化分析则表明, Type II B 可能是由 Type I A 演化而来, 或者说, Type II B 是 I A 在古细菌中的特化形式 (图 2, 3): 因此, Type II B 可能是由 Type II A 通过形成 Type I A 后演化而成的, 而不是由 Type II A 直接通过基因的复制、重组和缺失所造成的。此外, 在催化机制上, Type I A 和 Type II A 较之 Type II B 更为相似, Type I A 和 Type II A 在行使催化功能时都需要金属离子 (如镁离子) 的存在, 而 Type II B 的在行使催化功能时, 却不需要金属离子的存在, 通常, 同源蛋白质催化机制的相似性, 是其进化关系的反应: 因此, 催化机制的相似性也提示, Type II B 可能是由 Type II A 通过形成 Type I A 后演化而成的。

通常, Type I A 具有两个 CAP-like 结构域, 即 d3 和 d4 结构域, 然而基因组搜索发现, 并不是所

有物种中的 Type I A 都同时具有 d3 和 d4 结构域, 有的只具有 d3 结构域; 有的则只具有 d4 结构域。研究表明, d3 结构域具有催化功能, 而 d4 结构域不具催化功能 (Berger et al, 1998)。然而, d4 结构域是否具有其他的功能呢? 进一步的基因组搜索显示, 许多细菌往往具有多个序列长度不同的 Type I A 拷贝, 但其中至少有一个拷贝同时具有 d3 和 d4 结构域。因此, 在具有多拷贝 Type I A 的物种中, 缺少 d3 或 d4 结构域的 Type I A 应该都不具有功能, 真正行使功能的可能是同时具有 d3 和 d4 结构域的那个 Type I A。虽然, d4 结构域不具催化功能 (Berger et al, 1998), 但它的存在可能对于 Type I A 完成催化功能是必需的, 它可能对催化功能的完成起到间接或者辅助的作用。

综上所述, Type II B 很可能是由 Type I A 演化而来, 并且在由 Type II A 演化为 Type I A 最后演化为 Type II B 的过程中, DNA 拓扑异构酶的催化机制也发生了变化: 从需要金属离子的协助, 演化到了不需要镁离子的存在协助其催化, 同时, 其 DNA 断口的错位碱基数目也发生了变化。

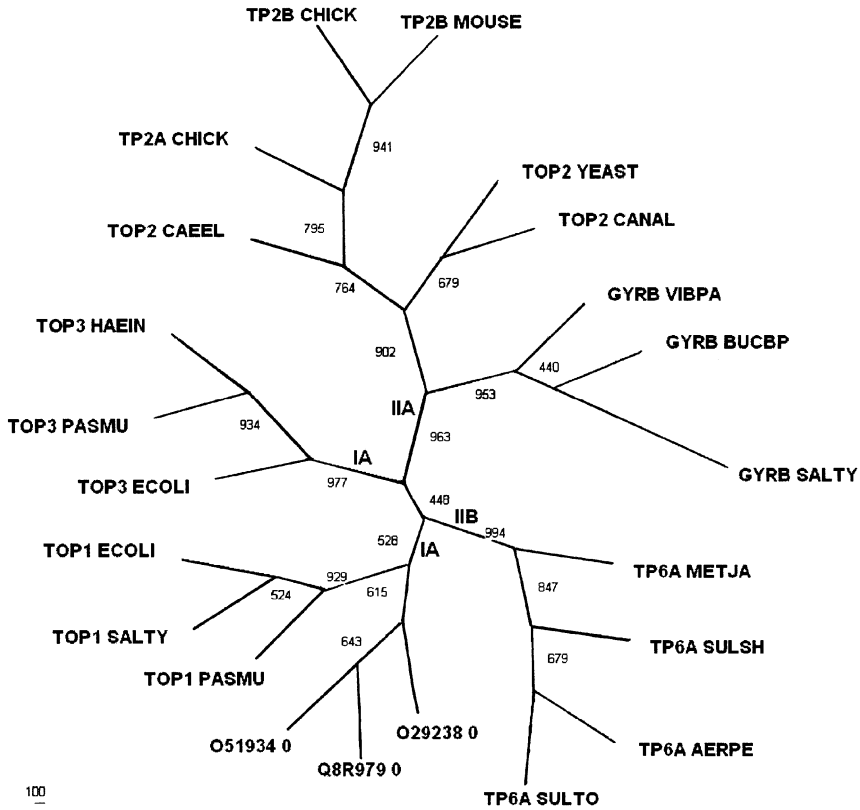


图 3 toprim 结构域的进化树 (1 000 bootstrap)

Fig. 3 The phylogenetic tree of toprim domains (1 000 bootstrap)
 I A: I A 类 (Type I A); II A: II A 类 (Type II A); II B: II B 类 (Type II B)。

1i7da (IA)	(1)	MRLFI A EKPSLARA I A D V L P -----GQVVTW C I G H L V T K Q L N V I K R F L H E
1bjt0 (IIA)	(443)	CTLV L T E G D S A L S L A V A G L A V V G R D Y Y G C Y P L R G K M --A E I Q A I K K I M G -
1d3ya (IIB)	(192)	F I L A I E T S G M F A R L N A E -----N C I L V S L K G V P A R A T R R F I K R L H E E
		bbb b a
1i7da (IA)	(95)	A S - E I V H A G D P D R E G Q L L V D E V L D Y L Q
1bjt0 (IIA)	(519)	--G H L M I M T D O D - H I K G L I I N F L E S S
1d3ya (IIB)	(240)	H D L P V L V F T D G D P Y G Y L N I Y R T L K V
		bbb b a

图 4 不同类型的拓扑异构酶 toprim 结构域基于结构的序列比对

Fig. 4 The structure-based sequence alignment of toprim domains in different Topoisomerases

1i7d: 大肠杆菌 Topo III (I A 类) (*Escherichia coli* Topo III (Type I A)); 1d3y: 古细菌 Topo VI (II B 类) (*Methanococcus jannaschii* Topo VI (Type II B)); 1bjt: 酵母菌 Topo II (II A 类) (*Saccharomyces cerevisiae* Topo II (Type II A)); 阴影 (shadow): 保守位点 (the conservative amino acid residues); a: α 螺旋 (α -helix); b: β 折叠 (β -sheet)。

参考文献:

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs [J]. *Nucleic Acids Res*, 25

(17): 3389 - 3402.

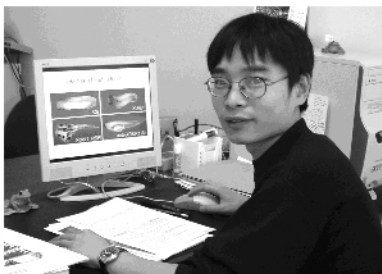
Bairoch A, Apweiler R. 2000. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000 [J]. *Nucleic Acids*

- Res*, **28** (1): 45–48.
- Benson DA, Boguski MS, Lipman DJ, Ostell J. 1997. GenBank [J]. *Nucleic Acids Res*, **25** (1): 1–6.
- Berger JM, Fass D, Wang JC. 1998. Structural similarities between topoisomerases that cleave one or both DNA strands [J]. *Proc Natl Acad Sci USA*, **95** (14): 7876–7881.
- Buhler C, Lebbink JH, Boes C, Ladenstain R, Forterre P. 2001. DNA topoisomerase VI generates ATP-dependent double-strand breaks with two-nucleotide overhangs [J]. *J Biol Chem*, **276** (40): 37215–37222.
- Champoux JJ. 2001. DNA topoisomerases: Structure, function, and mechanism [J]. *Annu Rev Biochem*, **70**: 369–413.
- Corbett KD, Berger JM. 2003a. Emerging roles for plant topoisomerase VI minireview [J]. *Chem Biol*, **10** (2): 107–111.
- Corbett KD, Berger JM. 2003b. Structure of topoisomerase VI-B subunit: Implications for type II topoisomerase mechanism and evolution [J]. *EMBO J*, **22** (1): 151–163.
- Corbett KD, Berger JM. 2004. Structure, molecular mechanisms, and evolutionary relationships in DNA Topoisomerases [J]. *Annu Rev Biophys Biomol Struct*, **33**: 95–118.
- Dai PG, Wang Y, Ye RS, Chen L, Huang L. 2003. DNA topoisomerase III from the hyperthermophilic archaeon *Sulfolobus solfataricus* with specific DNA cleavage activity [J]. *J Bacteriol*, **185** (18): 5500–5507.
- Felsenstein J. 1996. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods [J]. *Methods Enzymol*, **266**: 418–427.
- Forterre P. 2006. DNA topoisomerase V: A new fold of mysterious origin [J]. *Trends Biotechnol*, **24** (6): 245–247.
- Gangloff S, de Massy B, Arthur L, Rothstein R, Fabre F. 1999. The essential role of yeast topoisomerase III in meiosis depends on recombination [J]. *EMBO J*, **18** (6): 1701–1711.
- Hartung F, Puchta H. 2000. Molecular characterisation of two paralogous SPO11 homologues in *Arabidopsis thaliana* [J]. *Nucleic Acids Res*, **28** (7): 1548–1554.
- Krogh BO, Shuman S. 2002. A poxvirus-like Type I B topoisomerase family in bacteria [J]. *Proc Natl Acad Sci USA*, **99** (4): 1853–1858.
- Liu QY, Wang JC. 1999. Similarity in the catalysis of DNA breakage and rejoining by type I A and II A DNA topoisomerases [J]. *Proc Natl Acad Sci USA*, **96** (3): 881–886.
- Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures [J]. *J Mol Biol*, **247** (4): 536–540.
- Nichols MD, De Angelis K, Keck J L, Berger JM. 1999. Structure and function of an archaeal topoisomerase VI subunit with homology to the meiotic recombination factor Spo11 [J]. *EMBO J*, **18** (21): 6177–6188.
- Podobnik M, McInerney P, O'Donnell M, Kuriyan J. 2000. A TOPRIM domain in the crystal structure of the catalytic core of *Escherichia coli* primase confirms a structural link to DNA topoisomerases [J]. *J Mol Biol*, **300** (2): 353–362.
- Sali A, Blundell TL. 1990. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming [J]. *J Mol Biol*, **212**: 403–428.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Res*, **22** (22): 4673–4680.
- Wang JC. 1985. DNA topoisomerases [J]. *Annu Rev Biochem*, **54**: 665–697.
- Wang JC. 1996. DNA topoisomerases [J]. *Annu Rev Biochem*, **65**: 635–692.

本刊编委毛炳宇研究员简介

毛炳宇,男。1993年毕业于山东大学生物系,获学士学位;1998年获山东大学发育生物学专业博士学位。1998至2003年在德国癌症研究中心从事博士后研究。2003年12月至今任中国科学院昆明动物研究所研究员。

从事发育生物学研究,主要以非洲爪蟾为实验动物模型,研究脊椎动物神经诱导与图式形成的分子机制。目前任中德马普青年科学家小组组长,2004和2005年分别获得国家自然科学基金委杰出青年科学基金和重点项目的支持。2004年入选云南省中青年学术技术带头人后备人才。



毛炳宇研究员