非洲锥虫抗人血清相关基因的进化

赖德华^{*}, 王巧平, 李志, JULIUS Lukeš, 伦照荣^{*}

中山大学生命科学学院, 生物防治国家重点实验室, 寄生生物研究中心, 广州 510275

Biology Centre, Institute of Parasitology, Czech Academy of Sciences; Faculty of Biology, University of South Bohemia, 37005 České Budějovice (Budweis), Czech Republic

^{*} 联系人, E-mail: lsslzr@mail.sysu.edu.cn, ltacilai@hotmail.com

2008-11-22 收稿, 2009-01-15 接受

国家自然科学基金(批准号: 30570245, 30670275)、教育部创新团队基金(批准号: DPCCKSCU/IRT0447)、瑞典国际科学基金(批准号: B/4318-1)、捷克科学基金(批准号: Z60220518)和捷克教育基金(批准号: 2B06129)资助项目。

Lai D H, Wang Q P, Li Z, et al. Evolution of the serum resistance-associated *SRA* gene in African trypanosomes.

Chinese Science Bulletin, 2009, 54(7): 1275—1278, doi: 10.1007/s11434-009-0137-z

摘要 抗人血清相关蛋白(SRA)是非洲睡眠病的病原体布氏锥虫罗得西亚亚种抵抗人血清中的锥虫溶解因子溶解的关键蛋白。该蛋白同时也作为鉴定罗得西亚亚种的重要分子标记。本研究从锥虫亚属各物种的基因组中克隆出与 *SRA* 同源的基因片段 *SRAbc*, 并对其进行了测序和结构分析。测序结果显示, 各物种之间的 *SRAbc* 核酸序列同源性高达 82%, 表明 *SRAbc* 在锥虫亚属的种类中具有起源相同并相对保守的特性。由于在罗得西亚亚种内同时存在 *SRA* 和 *SRAbc*, 故认为 *SRAbc* 可以作为 *VSG* 重组的“供体 *VSG*”, 通过复制和插入机制整合表达位点。在自然选择的压力下, *SRAbc* 将可能再次进化成有功能的 *SRA* 基因。

关键词
抗人血清相关基因
SRA 基本拷贝
进化
非洲锥虫

血鞭毛虫中的锥虫亚属(*Trypanozoon*)种类多, 分布极其广泛, 可感染绝大多数哺乳动物, 引起人或家畜的锥虫病^[1,2]。由于正常人血清中存在可溶解锥虫的锥虫溶解因子(trypantolytic factors, TLFs), 家畜锥虫病的病原体, 如布氏锥虫布氏亚种(*Trypanosoma brucei brucei*)、马媾疫亚种(*T. b. equiperdum*)和伊氏亚种(*T. b. evansi*)都不感染人。然而, 布氏锥虫的罗得西亚亚种(*T. b. rhodesiense*)和冈比亚亚种(*T. b. gambiense*)却能抵抗TLFs, 引起人的急性或慢性非洲锥虫病。

目前, 罗得西亚亚种的抗人血清溶解因子已经被发现, 并被命名为抗人血清相关蛋白(serum resistance associated protein, SRA)。实验证明, 该蛋白能阻碍由TLFs介导的

杀锥虫作用^[3-5]。由于*SRA*基因仅存在于罗得西亚亚种中, 因此该基因被用作该亚种鉴定的特异分子标记^[6-9]。

*SRA*基因属于锥虫可变表面抗原(VSG)基因家族^[3,5,10], 编码区的核苷酸序列比普通VSG基因短, 全长仅为 1.2 kb。然而, 在应用*SRA*基因特异引物的PCR检测中, 从部分布氏锥虫亚种, 如布氏亚种和罗得西亚亚种中也能扩增出比目的片段大的另一片段^[11,12]。这个未知片段与*SRA*基因同源性达到 76%, 因此被命名为*SRA*基本拷贝(*SRAbc*)。但是与*SRA*基因相比, *SRAbc*内部存在移码和提前终止的现象^[11], 被认为是假基因。忽略这些错误编码信息后, 软件分析表明, *SRAbc*对应的蛋白氨基酸三维结构比*SRA*拥有更为复杂的典型VSG表面环结构^[12]。

鉴于 *SRAbc* 片段很有可能是 *SRA* 基因的进化前体, 我们希望能从锥虫亚属的多个锥虫 DNA 样本中获得该片段。相关的材料包括来自南美洲、亚洲和非洲的 8 株布氏锥虫伊氏亚种, 来自非洲和亚洲的 4 株马媾疫亚种, 来自非洲的布氏亚种, 罗得西亚亚种和冈比亚亚种各为 2 株, 以及 1 株刚果锥虫(*T. congolense*)作为外组(图 1)。

扩增 *SRAbc* 片段所用的正向引物(5'-CACACCTCTAAGAATC-ACAATAG-3')和反向引物(5'-ATTCTTGTGCCTTGGTGC-3')都源于基因数据库中布氏锥虫的布氏亚种和罗得西亚亚种的序列(NW_001076895 和 DQ427094)。PCR 反应采用 25 μL 体系, 包括约 10 ng 总 DNA, 0.8 μmol/L 各引物, 200 μmol/L

图 1 基于 SRAbc 序列构建的分子进化树

基于邻接法构建的非洲锥虫株SRAbc序列进化树(NJ树),把SRA基因作为外群,分析软件采用MEGA3.1^[13], Bootstrap重复 10000 次,成对删除法处理缺失碱基,其他为默认参数. 伊氏亚种虫株标记为Tev; 马媾疫亚种虫株标记为Teq; 布氏亚种, 冈比亚亚种和罗得西亚亚种分别标记为Tbbbru, Tbrho和Tbgam. 对应SRA的活性相关的氨基酸残基用阴影标记

dNTP, 和 1 U 的Ex Taq聚合酶 (TaKaRa, 日本). 反应程序为: 94 变性 5 min; 94 变性 30 s, 55 退火 1 min, 72 延伸 1 min, 35 个循环; 最后 72 延伸 10 min. 扩增产物通过电泳确认后回收目标片段, 克隆入 pBS-T载体(Tiangen, 北京)进行测序. 序列的拼接和比对分析使用 MEGA 3.1 软件^[13].

在所有检测的样品, 除刚果锥虫外(非锥虫亚属成员), 均扩增出 1 条长约 1.5 kb 的 SRAbc 特异产物, 测序结果表明该片段的序列间同源性高达 82%. 出乎意料的是, 罗得西亚亚种虽然已经具有 SRA 基因, 但也同样能扩增出 SRAbc 片段. 这个结果确证了 SRAbc 基因在锥虫亚属 (Trypanozoon), 即布氏锥虫及其亚种中普遍存在, 并可作为鉴定该种的一个特异分子标记.

我们把获得的SRAbc序列与SRA基因序列进行比对, 并构建分子进

化树(图 1). 结果显示, SRAbc序列分成两大支, 6 株伊氏亚种集中地聚在一起, 表现出极近的亲缘关系, 而肯尼亚的伊氏亚种虫株KETRI 2479 处在相对较远的地位(图 1). 由于伊氏亚种可分为A型和B型, 两型虫株之间在线粒体基因组^[14]和核基因组^[15]都有明显的差异. 因此, SRAbc的序列分析结果进一步支持伊氏锥虫的分型, 且这些类型可能并非单一起源. 有趣的是, 源于摩洛哥的伊氏亚种虫株Zagora I1.7 则落入到另一支. 马媾疫亚种的情况也与之相似, 分处于两支上. 这些结果为伊氏亚种和马媾疫亚种的多起源假说提供了新的证据^[16]. 更有趣的是, 仅有的两株冈比亚亚种都处于第一支上, 表现出与伊氏亚种有较高的亲缘关系, 但与同样能感染人的罗得西亚亚种的亲缘关系则反而较远. 事实上, 有学者认为, 冈比亚亚种和罗得西亚亚种两者对人的感染力分别由

不同的进化路线所获得^[6]. 本研究的结果支持这种观点.

典型的VSG蛋白的三级结构中含有 4 个功能/结构域, 包括N端的信号肽区、两个相互折叠的 α 螺旋区、表面环区以及C端区. 通过软件预测的SRA蛋白三级结构与VSG类似, 仅缺少明显的表面环结构^[11]. 通过把SRAbc基因序列中的移码和终止子去掉我们得到一条与SRA基因类似的蛋白序列(结果未显示). 这段序列的三级结构预测显示其与VSG结构一致^[17]. 所有分离株的预测结果中都包括保守的信号肽区(氨基酸残基 1~31)和 2 个 α 螺旋区(氨基酸残基 32~81 和 87~139)(图 2), 但表面环区则呈现高度可变. 此外, 构建的SRAbc和SRA (AF097331)模型中氨基酸残基 149~211 之间的二硫键较为保守(图 2). 同样, 在核酸序列分析中, 该片段可以简单地分成 3 个区: A区长 496 或 497 bp, C区为 436

图2 预测的 *SRAbc* 和 *SRA* 的二级和三级结构

(a) 预测的 *SRAbc* 和 *SRA* 的二级结构的多样性。图中显示保守的信号肽区(SP)和 2 个 α 螺旋区(Ha和Hb)、高度可变的表面环区(SL)、二硫键(灰线)以及序列保守的羧端(CT)。类型一: TevSTIB803, TevCAN86k, TbgamPAKWAITMAP2570, TbrhoAnTat25.1; 类型二: TevSTIB808, TevStockColumbia; TeqBoTat1.1, TeqSTIB841, TbbSTIB940; 类型三: TevSTIB815, TevSTIB810, TevKETRI2479, TeqSTIB818, TeqSTIB842, TbrhoAnTat12.1, TbgamAnTat11.17, TbbNW_001076895; 以及 *SRA* 基因: AF097331。(b) MODELLER (<http://salilab.org/modweb/>)^[18] 模拟的 *SRAbc* 和 *SRA* 的三级结构(模板为 1VSG 和 2VSG)。 α 螺旋区为灰色, β 折叠区为白色。 *SRAbc* 中用箭头标出的 3 个关键氨基酸残基(58, 61 和 62), 而在 *SRA* 中的相应氨基酸残基可能因为表面环的缺失, 引起结构的改变而暴露在外

到 440 bp, 这两个区与 *SRA* 基因比, 同源性分别为 70% (64%~78%) 和 76% (72%~79%)。而 B 区则在碱基序列和序列长度上都高度可变, 其中 *SRAbc* 片段中 B 区长 470~628 bp, 但在 *SRA* 基因里 B 区仅长 174~286 bp。本研究的结果支持 Vanhamme 等人^[12] 基于较少样品作出的推测。

SRAbc 的序列和 *VSG* 序列的两端保守性良好, 说明两者之间可能发生过重组现象。对 *SRAbc*, *SRA* 和 *VSG* 三者之间的进化关系有以下两种可能: () *SRAbc* 和 *SRA* 源于同一个 *VSG* 祖先, 并且由平行进化而来;

() *SRA* 源于 *SRAbc*, 而后者由 *VSG* 祖先进化而来。本文的数据支持第 2 种假说。布氏锥虫 *VSG* 基因的广泛分布和频繁变换, 源于 *VSG* 重组机制, 即处于激活表达状态的 *VSG* 被一个非表达的供体 *VSG* 所复制出来的拷贝替换^[19], 这种替换机制可同样解释 *SRAbc* 和 *SRA* 基因的现象。基于 *SRAbc* 和 *SRA* 共同存在于罗得西亚锥虫基因组中, 我们猜测原始的 *SRAbc* 基因可作为“供体 *VSG*”, 通过复制和重组插入到表达位点。支持这个猜想的又一证据是 *SRAbc* 的上游序列中有 1.7 kb 的片段与 *SRA* 共转录区

85% 同源^[12]。在自然选择下(如人血清), *SRAbc* 经过多次缺失变异形成 *SRA*^[11,20]。而原始的 *SRAbc* 则从选择压力下解脱出来, 积累移码等突变丧失编码普通 *VSG* 的能力。然而, *VSG* 基因马赛克式进化模式假说提出, 布氏锥虫可通过整合若干个 *VSG* 假基因在表达位点形成一个新的可表达的 *VSG* 基因^[21,22]。这意味着布氏锥虫仍可通过 *SRAbc* 基因部分片段, 与其他 *VSG* 基因重组形成新的 *SRA*。但目前对为何只有罗得西亚锥虫存在 *SRA* 基因仍缺乏令人信服的解释。

参考文献

- 1 Hoare C A. The Trypanosomes of Mammals. Oxford: Blackwell Scientific Publications, 1972
- 2 Maudlin I, Holmes P H, Miles M A. The Trypanosomiases. Oxfordshire: CABI publishing, 2004
- 3 Xong H V, Vanhamme L, Chamekh M, et al. A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma brucei rhodesiense*. Cell, 1998, 95(6): 839—846[doi]
- 4 Vanhamme L, Paturiaux-Hanocq F, Poelvoorde P, et al. Apolipoprotein L-I is the trypanosome lytic factor of human serum. Nature, 2003, 422(6927): 83—87[doi]
- 5 Pays E, Vanhollebeke B, Vanhamme L, et al. The trypanolytic factor of human serum. Nat Rev Microbiol, 2006, 4(6): 477—486[doi]
- 6 De Greef C, Imberechts H, Matthyssens G, et al. A gene expressed only in serum-resistant variants of *Trypanosoma brucei rhodesiense*. Mol Biochem Parasitol, 1989, 36(2): 169—176[doi]
- 7 De Greef C, Chimfwembe E, Kihang'a Wabacha J, et al. Only the serum-resistant bloodstream forms of *Trypanosoma brucei rhodesiense* express the serum resistance associated (SRA) protein. Ann Soc Belg Med Trop, 1992, 72(1): 13—21
- 8 Gibson W, Backhouse T, Griffiths A. The human serum resistance associated gene is ubiquitous and conserved in *Trypanosoma brucei rhodesiense* throughout East Africa. Infect Genet Evol, 2002, 1(3): 207—214[doi]
- 9 Radwanska M, Chamekh M, Vanhamme L, et al. The serum resistance associated gene as a diagnostic tool for the detection of *Trypanosoma brucei rhodesiense*. Am J Trop Med Hyg, 2002, 67(6): 684—690
- 10 De Greef C, Hamers R. The serum resistance associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. Mol Biochem Parasitol, 1994, 68(2): 277—284[doi]
- 11 Campillo N, Carrington M. The origin of the serum resistance associated (SRA) gene and a model of the structure of the SRA polypeptide from *Trypanosoma brucei rhodesiense*. Mol Biochem Parasitol, 2003, 127(1): 79—84[doi]
- 12 Vanhamme L, Renauld H, Lecordier L, et al. The reference *Trypanosoma brucei* strain TREU927/4 contains *T. brucei rhodesiense*-specific SRA sequences, but displays a distinct phenotype of relative resistance to human serum. Mol Biochem Parasitol, 2004, 135(1): 39—47[doi]
- 13 Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinfo, 2004, 5(2): 150—163[doi]
- 14 Borst P, Fase-Fowler F, Gibson W C. Kinetoplast DNA of *Trypanosoma evansi*. Mol Biochem Parasitol, 1987, 23(1): 31—38[doi]
- 15 Li F J, Gasser R B, Zheng J Y, et al. Application of multiple DNA fingerprinting techniques to study the genetic relationships among three members of the subgenus *Trypanozoon* (Protozoa: Trypanosomatidae). Mol Cell Probes, 2005, 19(6): 400—407[doi]
- 16 Lai D H, Hashimi H, Lun Z R, et al. Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *T. equiperdum* and *T. evansi* are *petite* mutants of *T. brucei*. Proc Natl Acad Sci USA, 2008, 105(6): 1999—2004[doi]
- 17 Carrington M, Miller N, Blum M, et al. Variant specific glycoprotein of *Trypanosoma brucei* consists of two domains each having an independently conserved pattern of cysteine residues. J Mol Biol, 1991, 221(3): 823—835
- 18 Lambert C, Léonard N, De Bolle X, et al. ESyPred3D: Prediction of proteins 3D structures. Bioinformatics, 2002, 18(9): 1250—1256[doi]
- 19 Donelson J E. Antigenic variation and the Africa trypanosome genome. Acta Trop, 2003, 85(3): 391—404[doi]
- 20 Gibson W, Ferris V. Conservation of the genomic location of the human serum resistance associated gene in *Trypanosoma brucei rhodesiense*. Mol Biochem Parasitol, 2003, 130(2): 159—162[doi]
- 21 Roth C, Bringaud F, Layden R E, et al. Active late-appearing variable surface antigen genes in *Trypanosoma equiperdum* are constructed entirely from pseudogenes. Proc Natl Acad Sci USA, 1989, 86(23): 9375—9379[doi]
- 22 Barry J D, Marcello L, Morrison L J, et al. What the genome sequence is revealing about trypanosome antigenic variation. Biochem Soc Transact, 2005, 33(5): 986—989[doi]