

DOI:10.13350/j.cjpb.231210

• 论著 •

细粒棘球蚴 DNA 损伤修复蛋白 RAD51 的生物信息学分析及初步验证^{*}

巩月红^{1,2}, 林玉霞³, 王梦君⁴, 潘美驰³, 赵一聪³, 赵军^{1,2}, 王建华^{1,2,*}

(1. 新疆医科大学第一附属医院药理学部, 新疆 乌鲁木齐 830011; 2. 新疆医科大学省部共建中亚高发病因与防治国家重点实验室; 3. 新疆医科大学药学院; 4. 新疆医科大学医学部)

【摘要】 **目的** 利用生物信息学软件预测和分析细粒棘球蚴 DNA 损伤修复蛋白(EgRAD51)的分子特性,为 EgRAD51 的功能和药物靶点研究奠定基础。 **方法** 在 NCBI 数据库中检索下载 EgRAD51 蛋白的氨基酸序列,采用 ProtParam 预测 EgRAD51 蛋白的理化性质,SignalP-5 预测信号肽,ProtScale 预测其亲疏水性,TMHMM 分析跨膜区域,Euk-mPLoc2.0 预测亚细胞定位,SOMPA 和 Swissmodel 预测其二、三级结构,NetPhos3.1、NetOGlyc4.0 及 NetNGlyc1.0 预测磷酸化位点,ABC pred 和 SYFPEITHI 数据库预测其 T、B 细胞表位。采用 MEGA11 软件对不同物种来源的 RAD51 蛋白进行多序列比对并构建系统进化树。采用 qRT-PCR 检测不同药物干预后 EgRAD51 mRNA 表达水平。 **结果** 预测 EgRAD51 氨基酸数 378 个,分子式为 C₁₈₃₉H₂₉₅₀N₅₀₈O₅₄₄S₂₀,分子质量为 41.521 93 ku,pI 值为 7.95;该蛋白无信号肽和无跨膜区,定位于细胞核内;含有 32 个丝氨酸磷酸化位点,26 个苏氨酸磷酸化位点,9 个酪氨酸磷酸化位点,4 个 O-糖基化潜在位点,1 个 N-糖基化位点;含 PRK09302 超级家族保守结构域及 RAD51C 重组酶保守结构域;二级结构中 α -螺旋占 45.24%,延伸主链占 15.08%, β -转角占 3.97%,无规则卷曲占 35.71%,为低聚单体;含有 13 个 B 细胞抗原表位,具有能与 HLA-A * 02:01 的结合能力,且能被分子呈递。同源性及系统进化分析显示,EgRAD51 与多房棘球蚴 RAD51 的氨基酸序列一致,与亚洲带绦虫、中殖孔绦虫、血吸虫、欧洲锥虫等亲缘关系较近,与智人、果蝇、中华枝睾吸虫、带翅棘球蚴亲缘关系较远。qRT-PCR 检测不同药物干预组的 EgRAD51 mRNA 较空白对照组均上调。 **结论** EgRAD51 蛋白定位于细胞核内,含有丰富的 B 细胞抗原表位,具有能与 HLA-A * 02:01 的结合能力,是一个具有潜在研究价值的药物靶点。

【关键词】 细粒棘球蚴;EgRAD51;DNA 损伤修复;生物信息学

【中图分类号】 R383.33

【文献标识码】 A

【文章编号】 1673-5234(2023)12-1416-08

[Journal of Pathogen Biology. 2023 Dec;18(12):1416-1423.]

DNA damage repair protein RAD51 in *Echinococcus granulosus* bioinformatics analysis and preliminary validation

GONG Yuehong^{1,2}, LIN Yuxia³, WANG Mengjun⁴, PAN Meichi³, ZHAO Yicong³, ZHAO Jun^{1,2}, WANG Jianhua^{1,2} (1. Pharmacy, First Hospital Affiliated with Xinjiang Medical University, Urumqi 830011, China; 2. Xinjiang Medical University Jointly Established the State Key Laboratory of Pathogenesis, Prevention, and Treatment of Diseases Highly Prevalent in Central Asia; 3. College of Pharmacy, Xinjiang Medical University; Medical Department, Xinjiang Medical University) ***

【Abstract】 **Objective** To predict and analyze the molecular characteristics of the DNA damage repair protein (EgRAD51) in *Echinococcus granulosus* using bioinformatics software, in order to lay the foundation for further exploring the function of EgRAD51 and developing drug targets. **Methods** The amino acid sequence of EgRAD51 protein was retrieved and downloaded from the NCBI database. ProtParam was used to predict the physical and chemical properties of EgRAD51 protein, SignalP-5 was used to predict the Signal peptide, ProtScale was used to predict its hydrophobicity, TMHMM was used to analyze the transmembrane region, Euk-mPLoc2.0 was used to predict the subcellular location of the protein, SOMPA and Swissmodel were used to predict its secondary and tertiary structure, NetPhos3.1, NetOGlyc4.0 and NetNGlyc1.0 were used to predict the phosphorylation site, ABC pred and SYFPEITHI databases predict the T and B cell epitopes of EgRAD51 protein. Using MEGA11 software to perform multiple sequence alignment of RAD51 proteins

* **【基金项目】** 新疆维吾尔自治区科学技术厅自然科学基金重点项目(No. 2021D01D15);省部共建国家重点实验室开放课题(No. SKL-HIDCA-2022-9);新疆维吾尔自治区科学技术厅自然科学基金面上项目(No. 2020D01C240)。

** **【通讯作者】** 王建华, E-mail: jhw716@163.com

【作者简介】 巩月红(1976-),女,甘肃天水人,副主任药师。主要研究方向:抗虫病机制研究。E-mail: gongyh0602@126.com

from different species and construct a phylogenetic tree. Using qRT PCR to detect the expression level of EgRAD51 mRNA after different drug interventions. **Results** The results predicted that EgRAD51 had 378 amino acid numbers, a molecular formula of $C_{1839}H_{2950}N_{508}O_{544}S_{20}$, a molecular weight of 41, 521.93 ku, and a pI value of 7.95; The protein has no signal peptide or transmembrane region and is localized within the nucleus; Contains 32 serine phosphorylation sites, 26 threonine phosphorylation sites, 9 tyrosine phosphorylation sites, 4 O-glycosylation potential sites, and 1 N-glycosylation site; Contains the conserved domains of PRK09302 superfamily and RAD51C recombinase; In the secondary structure α -Helixes account for 45.24%, and extended main chains account for 15.08%, β -Corners account for 3.97%, irregular curls account for 35.71%, and are oligomers; It contains 13 B cell antigen epitopes and has the ability to bind to HLA-A * 02-01, and can be presented by molecules. Homology and phylogenetic analysis showed that the amino acid sequences of EgRAD51 were consistent with those of *Echinococcus multilocularis* RAD51, and were closely related to *Taenia asiatica*, *Taenia media*, *Schistosoma*, *Trypanosoma europaeans*, and far from *Homo sapiens*, *Drosophila*, *Clonorchis sinensis*, and *Echinococcus winged*. The expression of EgRAD51 mRNA in different drug intervention groups was upregulated by qRT PCR compared to the blank control group. **Conclusion** The EgRAD51 protein is located in the nucleus and contains rich B cell antigen epitopes. It has the ability to bind to HLA-A * 02-01 and is a potential drug target with research value.

【Key words】 *Echinococcus granulosus*; EgRAD51; DNA damage repair; bioinformatics

囊型棘球蚴病(cystic echinococcosis, CE)是细粒棘球绦虫的幼虫细粒棘球蚴(*Echinococcus granulosus*, Eg)感染所致的人畜共患病^[1]。该病在全世界流行,病死率为2%~4%^[2]。我国是全球棘球蚴病流行最严重的国家之一,估计有38万患者,近6600万人有感染风险,是导致我国西部地区因病返贫的主要病种^[3]。CE临床治疗首选方案为手术摘除,但因多重因素限制,并非所有患者均符合手术适应证^[4]。药物治疗成为治疗CE有力的辅助手段,并成为该病研究的主要方向。但目前WHO推荐用于治疗棘球蚴病的药物较少,除阿苯达唑(Albendazole, ABZ)外尚无其他新药上市^[5],因此亟需开发CE的替代化疗方案以实现CE的药物高效治疗。

在CE的化疗中明确药物作用机制及寻找药物靶点是提高治疗效率的关键。研究发现DNA损伤修复对于基因组稳定性以及在疾病的预防治疗中起重要作用,针对损伤的DNA修复系统对于疾病细胞的存活也至关重要^[6]。DNA损伤过程中DNA双链断裂(DSB)是最具细胞毒性的DNA损伤之一,其错误修复会导致突变和易位。RAD51的重组酶活性是3个主要DSB修复途径的核心,RAD51及其基因家族在双链断裂修复、复制应激和减数分裂中发挥多种作用,具有发现和侵入同源DNA序列的功能,成为DNA保真度的关键调节因子,能够及时地进行DNA修复,对于疾病治疗后的复发起重要作用^[7]。研究表明,EgDNA可能是生物碱类药物主要攻击部位,药物作用后可导致同源重组修复通路相关基因EgBRCA1、EgMRE11等DNA末端修饰因子表达上调,募集EgRad51发挥修复功能,EgRad51几乎参与同源重组修复的全部阶段,可能随着疾病的发生发展,部分Eg会启动DNA修复功能而成功存活并使CE复发。因

此推测EgRad51可能是CE治疗过程中产生耐药的关键因素,也是Eg治疗潜在的关键靶点^[8-12]。深入探究EgRad51蛋白功能在生物碱类药物治疗囊型包虫病的作用机制研究中至关重要,可能有助于进一步理解生物碱类药物治疗囊型包虫病的病理机制,并为开发新的治疗方法提供新思路。

本研究利用生物信息学方法预测EgRad51蛋白的理化特性、结构特性和生物学功能,并采用实时荧光定量PCR检测去氢骆驼蓬碱及其衍生物对EgRad51基因相对表达量的影响,为深入研究EgRad51在生物碱类药物治疗CE中的作用奠定理论基础。

材料与方法

1 材料

1.1 细粒棘球蚴细粒棘球蚴 采自新疆乌鲁木齐华凌屠宰场自然感染Eg的绵羊肝脏。

1.2 主要试剂和仪器 总RNA提取试剂盒与反转录试剂盒均购于日本TakaRa生物技术公司;RPMI1640培养基购自美国Hyclone公司;Trizol购自美国Thermo Fisher公司。核酸定量仪和恒温培养箱购自美国Thermo公司;QuantStudioTM6qRT-PCR仪购自美国Life公司。

2 方法

2.1 EgRAD51氨基酸序列的获取 登录NCBI,在Protein数据库(<http://www.ncbi.nlm.nih.gov/protein>)中检索获取EgRAD51的氨基酸序列,登录号为:CDS15925。

2.2 EgRAD51蛋白的生物信息学分析

2.2.1 理化性质预测 运用在线软件ProtParam(<http://web.expasy.org/protparam/>)预测EgRAD51的分子质量、理论等电点、氨基酸组成;运

用在线预测软件 ProtScale(<http://web.expasy.org/ProtScale/>)中的 Hphob./Kyte&Doolittle 算法预测蛋白质的亲/疏水性。

2.2.2 信号肽、跨膜结构域预测及亚细胞定位 运用 SignalP5.0 (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>)预测 EgRAD51 是否含有信号肽序列;使用 TMHMM 在线服务器(<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>)分析 EgRAD51 是否含有跨膜区域;运用 Euk-mPLoc2.0 (<http://www.cbsio.sjtu.edu.cn/bioinf/euk-multi-2/>)预测该蛋白的亚细胞定位。

2.2.3 空间结构预测 访问在线预测软件 SOPMA (https://npsa-pbil.ibcp.fr/cgi-bin/secpred_sopma.pl)进行蛋白质二级结构预测;使用 SWISS-MODEL 在线服务器对 Eg RAD51 的三维结构进行预测。

2.2.4 磷酸位点、O-糖基化位点及 N-糖基化位点预测 运用 NetPhos3.1 (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>)在线软件预测 Eg RAD51 蛋白的磷酸化位点,运用在线软件 NetOGlyc4.0 (<https://services.healthtech.dtu.dk/service.php?NetOGlyc-4.0>)和 NetNGlyc1.0 (<https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0>)预测 Eg RAD51 蛋白的 O-糖基化位点及 N-糖基化位点。

2.2.5 EgRAD51 抗原表位预测 使用 ABC pred、SVMTriP 和 IEDB 数据库(<https://www.iedb.org/>)预测 EgRAD51 的优势 B 细胞表位;使用 IEDB 和 SYFPEITHI 数据库预测蛋白的优势 T 细胞表位。

2.2.6 蛋白的保守结构域预测 运用在线网站 CDD (ncbi.nlm.nih.gov/cdd)预测 Eg RAD51 蛋白的保守结构域。

2.3 EgRAD51 蛋白的同源性分析及系统发育树构建 通过 NCBI 网站 BLAST 程序 (Basiclocal Alignmeng Search Tool, <http://www.NCBI.nlm.nih.gov/blast>)分析蛋白序列同源性,运用 MEGA11 软件 (<http://www.Megasoftware.net/>)对 EgRAD51 和其他物种的 RAD51 序列构建系统发育树。

2.4 qRT-PCR 检测去氢骆驼蓬碱及其衍生物干预对 Eg RAD51 基因表达量的影响 将 Eg 加入 24 孔培养板中,密度为 3 000 头/孔。试验设空白对照组、阿霉素组(DOX)组、100 μ mol/L 去氢骆驼蓬碱(Hermine, HM)组、100 μ mol/L DH-330 组、DOX 联合 DNA 损伤修复抑制剂(RI-1)组、HM+RI-1(1:1)组、DH-330+RI-1(1:1)组。药物干预 24 h 后收集各组 Eg,采用 Trizol 法提取总 RNA,并测定其浓度及纯度。采用

PrimeScript™ reagent Kit(NO. RR037A)试剂盒进行 cDNA 的合成,采用 qRT-PCR 检测 Eg RAD51 mRNA。引物序列见表 1,由上海生工有限公司合成。试验设阴性质控组,每个待测样品均设 3 个重复, $2^{-\Delta\Delta C_t}$ 法数据分析,目的基因药物干预组/空白对照组 ≥ 2 倍为表达增高, ≤ 0.5 倍为表达降低。

表 1 引物信息
Table 1 Primer information

基因 Gene	引物序列 Primer sequences(5'→3')
EgRAD51	F:GGCACTGATGGTGGCAATCTCC R:CAGAAGATCCGCTCCACTTATCGC
内参 β -actin	F:GCGATGTATGTAGCTATCCAGGCAGTGCTCTCGCT R:CAATCCAGACAGATATTTGCGTTCCGGAGGA

结 果

1 EgRAD51 蛋白的氨基酸序列

登录 NCBI 查询 EgRAD51 基因 cDNA 全长 1 604 bp,编码 378 个氨基酸,其序列为:MLSD PDFREL VMLPLPTS VLRELSRAGYSNVGEISGLT IEQIESVCKIPKPRAASIATHIQEHASLRKDSVFTT NFPNKNLFYPRSA LQMLHPSLSNLVNPSEFRVA NVLTMSKSFDDLMGGGFPCGRITELCGQPGVGK TQFCLQTCLTVQIPQWCGGLAGEAIFLDTEGSFI PKRLHQMAVELVKHCQTMLAPYDSYADAEEK VACLEALSRIPTESLLSKVHYIRCTGYLQLLAA VQRLGEFCKYHPNIRLIIVDSIALPFRYEFEDIPQ RNRLLA AVAQSLSTASAQNAAVIMTNQITTKF VTSETSATTRSTLVPALGESWGHICSVRLLSK AEGSDAEQRRTARLLKHPGRPPGTAAYQVTA GGIRDCF。

2 Eg RAD51 蛋白的理化性质

利用生物信息学软件 ProtParam 进行分析,Eg RAD51 蛋白的分子式为 $C_{1839}H_{2950}N_{508}O_{544}S_{20}$,分子质量为 41.521 93 ku,理论 pI 值为 7.95,带负电荷的残基总数(Asp + Glu)为 35,带正电荷的残基总数(Arg + Lys)为 37,原子总数为 5 861;其不稳定系数(II)为 44.42(>40 不稳定),属于不稳定蛋白;脂肪族指数为 93.70,总亲水性指数平均值(GRAVY)为-0.009,为亲水蛋白。该蛋白的半衰期为 30 h (mammalian reticulocytes),体内 >20 h (yeast),体内 >10 h (Escherichia coli);消光系数为 25 160。

3 Eg RAD51 蛋白信号肽、跨膜结构域及亚细胞定位预测

使用 SignalP、TMHMM 在线服务器预测 Eg RAD51 的信号肽,Sec/SPI 为 0.0022(图 1)。表明该蛋白无信号肽裂解位点,不含有信号肽序列,可能是非

分泌蛋白。也不存在跨膜区域(图 2)。Euk-mPLoc2.0 在线软件预测 Eg RAD51 蛋白可能定位于细胞核内。

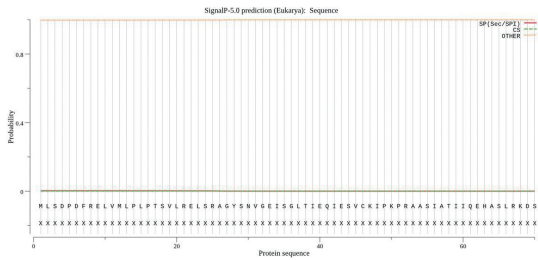


图 1 Eg RAD51 的信号肽分析
Fig. 1 Signal Peptide analysis of Eg RAD51 Protein

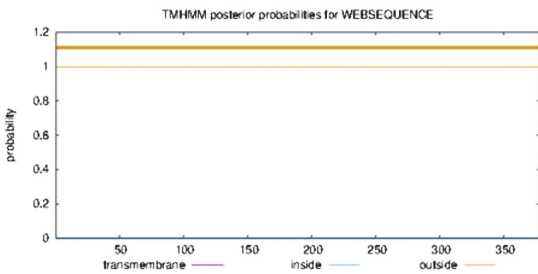
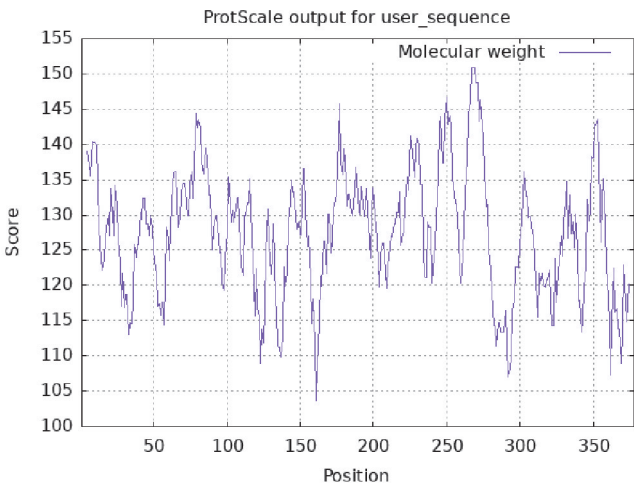


图 2 Eg RAD51 的跨膜结构域
Fig. 2 Transmembrane domain of Eg RAD51 Protein

4 Eg RAD51 蛋白的亲/疏水性

利用在线预测软件 ProtScale 中的 Hphob./Kyte&Doolittle 算法预测 Eg RAD51 蛋白的亲/疏水性,结果如图 3。该蛋白在第 346 个氨基酸残基处亲水性最强,亲水性最大值为-2.489。在第 259 位氨基酸残基处的疏水性最强,最大值为 2.533。该蛋白具有多个亲/疏水性区域分布。



注:横轴代表 Eg RAD51 的氨基酸残基位点,纵轴代表亲/疏水性得分(负值表示亲水性,正值表示疏水性)。

图 3 Eg RAD51 蛋白的亲/疏水性分析

Notes: The horizon axis represents the position of amino acids, and The long-term axis representation of hierarchy (positive values report) Hydrophobicity and negative values represent hydrophilicity.

Fig. 3 Hydrophobic map of Eg RAD51protei

5 Eg RAD51 蛋白的保守结构域

Eg RAD51 蛋白具有含核苷三磷酸水解酶的 P-loop(P-loop NTPase)结构域超家族(即:PRK09302 超级家族)的保守结构域(位于 106-146 氨基酸残基之间)和 RAD51C 重组酶保守结构域(位于 127-358 氨基酸残基之间)(图 4)。

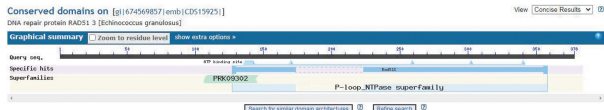
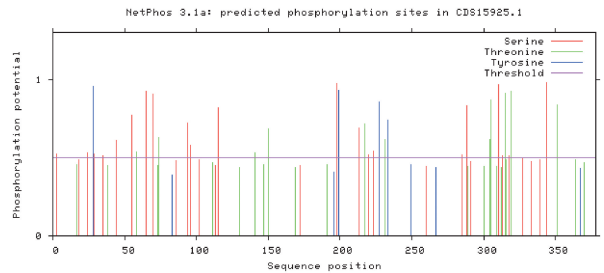


图 4 Eg RAD51 蛋白的保守结构域
Fig. 4 Conservative Domain of Eg RAD51 Protein

6 Eg RAD51 蛋白的磷酸化位点、O-糖基化位点和 N-糖基化位点

NetPhos3.1 在线网站分析 Eg RAD51 蛋白质含有 67 个磷酸化位点,其中有 32 个丝氨酸(Ser)磷酸化位点,26 个苏氨酸(Thr)磷酸化位点,9 个酪氨酸(Tyr)磷酸化位点(图 5)。NetOGlyc4.0 在线网站分析该蛋白存在 4 个 O-糖基化潜在位点,位于第 18、65、309 和 315 位氨基酸处,分值分别为 0.5、0.58、0.65、0.73。NetNGlyc1.0 在线软件分析该蛋白含有 1 个 N-糖基化位点,位于第 100 位氨基酸残基处(图 6)。



注:红色竖线代表丝氨酸磷酸化位点;绿色竖线代表苏氨酸磷酸化位点;蓝色竖线代表酪氨酸磷酸化位点。

图 5 Eg RAD51 蛋白的磷酸化位点预测

Notes: Red indicates the red vertical bar represents the serine phosphorylation site; Green vertical bars represent threonine phosphorylation sites; The blue vertical bars represent tyrosine phosphorylation sites.

Fig. 5 Prediction of phosphorylation sites of Eg RAD51 protein

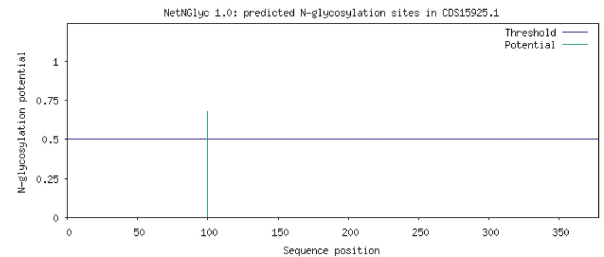
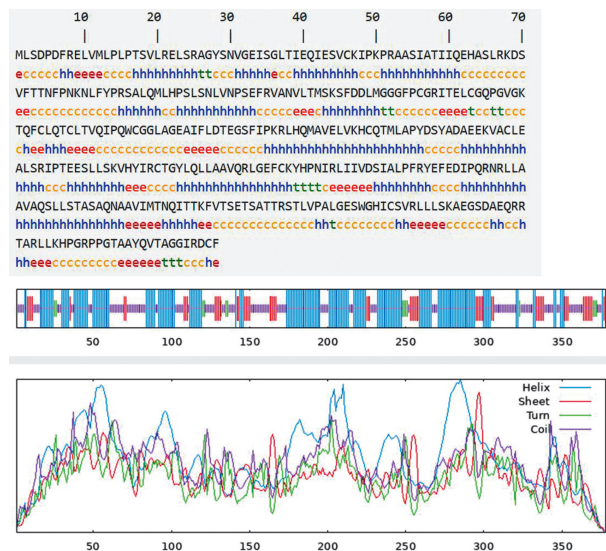


图 6 Eg RAD51 蛋白的 N-糖基化位点预测
Fig. 6 Prediction of N-glycosylation sites of Eg RAD51 protein

7 Eg RAD51 蛋白的空间结构

运用 SOMPA 及 SWISS-MODEL 分析 Eg RAD51 蛋白二、三级结构,其中二级结构中 α 螺旋占 45.24%,延伸链占 15.08%, β 转弯占 3.97%,无规卷曲占 35.71%(图 7)。经盘绕、折叠和卷曲等一系列复杂的过程形成的三级结构如图 8。



注:h(蓝色)表示 α 螺旋;e(红色)表示延伸链;t(绿色)表示 β 转角;c(黄色)表示无规卷曲

图 7 Eg RAD51 蛋白的二级结构预测

Notes:h (blue) indicates α helix. e (red) denotes extended strand; t (green) indicates β Angle;c (yellow) indicates random coiling.

Fig. 7 Prediction of secondary structure of Eg RAD51

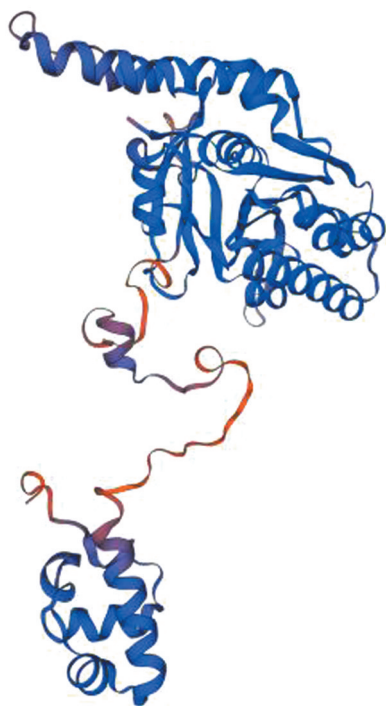


图 8 Eg RAD51 蛋白的三级结构预测

Fig. 8 Prediction of tertiary structure of Eg RAD51

8 Eg RAD51 蛋白的 T、B 细胞表位

当选定阈值为 0.8 时,ABCpred 软件预测 Eg RAD51 蛋白共有 13 个 B 细胞表位(表 2)。SVMTrIP 得到的 B 抗原表位为:31-35,103-122,169-188,218-237,276-295,339-358(表 3)。IEDB 软件预测 Eg RAD51 蛋白共有 15 个 B 细胞表位(表 4,图 9)。取 3 个软件预测结果交集综合分析,共得到 4 个 B 细胞抗原表位:32-115,171-181,187-222,343-347。经 EDBI 在线软件分析,HLA-DRB1 * 0701 限制性细胞 T 细胞表位和 HLA-A * 1101 限制性辅助性 T 细胞表位分值大于 0.3 的分别有 8 个和 17 个(表 5,表 6)。

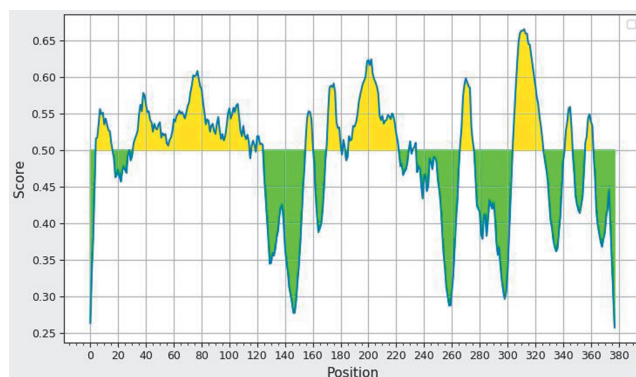


图 9 IEDB 预测 Eg RAD51 蛋白 B 细胞表位

Fig. 9 IEDB predicted Eg RAD51 protein B cell epitopes

表 2 ABCpred 预测 Eg RAD51 蛋白 B 细胞表位

Table 2 ABCpred predicts Eg RAD51 protein B cell epitopes

序号 No.	氨基酸序列 Amino acid sequence	起始位置 Start position	分值 Score
1	KRLHQMAVELVKHCQT	176	0.90
2	SALQMLHPSLSNLVNP	86	0.89
3	EQRRARLLKHPGRPP	347	0.87
4	SGLTIEQIESVCKIPK	35	0.86
5	KHCQTMLAPYDSYADA	187	0.86
6	KHPGRPPGTAAYQVTA	341	0.85
7	PCGRITELCGQPGVGK	356	0.84
8	GESWGHICSVRLLLSK	125	0.84
9	ASAQNAAVIMTNQITT	290	0.83
10	GESWGHICSVRLLLSK	325	0.82
11	LSNLVNPSEFRVANVL	24	0.82
12	QWCGGLAGEAIFLDTE	155	0.81
13	ATIIQEHASLRKDSVF	57	0.80

表 3 SVMTrIP 预测 Eg RAD51 蛋白 B 细胞表位

Tab 3 SVMTrIP predicts Eg RAD51 protein B cell epitopes

序号 No.	位置 Location	氨基酸序列 Amino acid sequence	得分 Score
1	218-237	EESLLSKVHYIRCTGYLQLL	1.00
2	276-295	NRLLAAVAQSLSTASAQNA	0.71
3	31-35	VEISGLTIEQIESVCKIPK	0.65
4	169-188	TEGSFIPKRLHQMAVELVKH	0.55
5	103-122	EFRVANVLTMSKSFDDLMMGG	0.42
6	339-358	SKAEGSDAEQRRARLLKHP	0.33

表 4 IEDB 预测 Eg RAD51 蛋白 B 细胞表位
Tab 4 IEDB predicts Eg RAD51 protein B cell epitopes

序号 NO.	起始位点 Start	结束位点 End	肽段 Peptide	长度 Length
1	5	16	PDFRELVMLPLP	12
2	29	29	S	1
3	32	115	GEISGLTIEOIESVCKIPKPRAASIAT IIQEHASLIKDSVFTTNFPKNLYP RSALQMLHPSLSNLVPSEFRVAN VLTMSSKS	84
4	117	119	DDL	3
5	121	125	GGGFP	5
6	156	161	WCGGLA	6
7	171	181	GSFIPKPLHQM	11
8	183	184	VE	2
9	187	222	KHCQTMLAPYDSYADAEEKVACL EASRIPTHEELS	36
10	230	231	CT	2
11	233	235	YLQ	3
12	267	276	YEFEDIPQRN	10
13	305	326	TKFVTSETSATTRSTLVPALGE	22
14	343	347	GSDAE	5
15	357	362	HPGRPP	6

表 5 IEDB 预测 Eg RAD51 蛋白 T 细胞表位(MHC 一类分子)
Tab 5 IEDB predicted Eg RAD51 protein T cell epitopes
(MHC class I molecules)

序号 NO.	起始位点 start	结束位点 end	长度(bp) length	核心肽段 core_peptide	肽段 peptide	分值 score
1	257	267	11	IVDSIALPFRY	IVDSIALRY	0.824038
2	73	83	11	TTNFPNKNLFY	TTNNKNLFY	0.755995
3	216	227	12	PTEESLLSKVHY	PTEESLLSY	0.705547
4	259	267	9	DSIALPFRY	DSIALPFRY	0.5502
5	72	83	12	FTTNFPNKNLFY	FTTNKNLFY	0.495711
6	190	199	10	QTMLAPYDSY	QTMLAPDSY	0.450054
7	299	307	9	MTNQITTKF	MTNQITTKF	0.32486
8	219	227	9	ESLLSKVHY	ESLLSKVHY	0.313143

表 6 IEDB 预测 Eg RAD51 蛋白 T 细胞表位(MHC 二类分子)
Tab 6 IEDB predicted Eg RAD51 protein T cell epitopes
(MHC class II molecules)

序号 NO.	起始位点 start	结束位点 end	长度(bp) length	核心肽段 core_peptide	肽段 peptide	分值 score
1	275	289	15	LLAAVAQSL	RNRLAAVAQSLST	0.7732
2	174	188	15	LHQMAVELV	IPKRLHQMAVELVKH	0.7224
3	274	288	15	LLAAVAQSL	QRNRLAAVAQSLLS	0.7129
4	101	115	15	FRVANVLT	PSEFRVANVLTMSKS	0.531
5	175	189	15	LHQMAVELV	PKRLHQMAVELVKHC	0.5066
6	85	99	15	LQMLHPSLS	RSALQMLHPSLSNLV	0.4492
7	100	114	15	FRVANVLT	NPSEFRVANVLTMSK	0.4297
8	273	287	15	LLAAVAQSL	PQRNRLAAVAQSLL	0.4254
9	4	18	15	FRELVMLPL	DPDFRELVMLPLPTS	0.4129
10	173	187	15	LHQMAVELV	FIPKRLHQMAVELVK	0.413
11	84	98	15	LQMLHPSLS	PRSALQMLHPSLSNL	0.4073
12	252	266	15	LIIVDSIAL	NIRLIIVDSIALPFR	0.393
13	163	177	15	FLDTEGSFI	EAIFLDTEGSFIPKR	0.3821
14	276	290	15	LLAAVAQSL	NRLAAVAQSLLSTA	0.3368
15	3	17	15	FRELVMLPL	SDPDFRELVMLPLPT	0.3201
16	83	97	15	LQMLHPSLS	YPSALQMLHPSLSN	0.3133
17	333	347	15	LLLSKAEGS	SVRLLLSKAEGSDAE	0.3001

9 Eg RAD51 氨基酸序列比对及系统进化分析

与其他物种的 RAD51 的氨基酸序列进行同源性比较,显示 EgRAD51 蛋白与绦虫类的多房棘球绦虫(*Echinococcus multilocularis*, Em)的氨基酸序列一致性为 100%,与亚洲带绦虫(*Taenia asiatica*)、中殖孔绦虫(*Mesocostoides corti*)、血吸虫(*Schistosoma haematobium*) 欧洲锥虫(*Spirometra erinaceieuropae*)亲缘关系较近,与哺乳动物纲中的智人(*Homo sapiens*)、非脊椎类的果蝇(*Drosophila melanogaster*, Dm) 中华枝睾吸虫(*Clonorchissiniensis*, Cs)及带翅棘球蚴亲缘关系较远(图 10)。

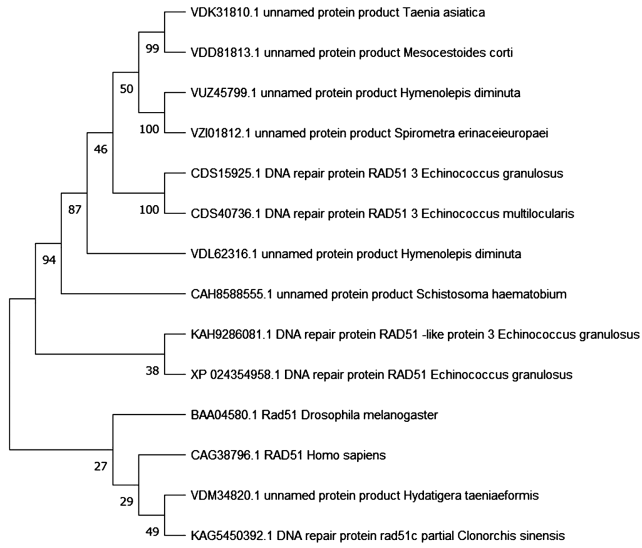


图 10 EgRAD51 蛋白进化树
Fig. 10 Phylogenetic tree of EgRAD51 prototype experience

10 药物干预对 EgRAD51 基因表达的影响

采用 qRT-PCR 检测药物干预后细粒棘球蚴 EgRAD51 基因表达量的变化,结果如图 11。与空白对照组相比,DOX、DOX + RI-1、HM、DH-330 组 EgRAD51 mRNA 表达量均显著上调($F=580.661$, $P<0.05$ 或 $P<0.01$);与单独给药组相比,联合用药组 EgRAD51mRNA 表达量均下调($F=193.035$, 均 $P<0.05$)。提示 HM 及其衍生物 DH-330 可对细粒棘球蚴造成 DNA 损伤,DNA 损伤修复抑制剂 RI-1 可抑制细粒棘球蚴的修复相关基因表达。RAD51 基因是同源重组修复通路上的重要基因,而 Eg 应激上调与 DNA 修复相关的基因,可能通过激活 EgRAD51 等修复基因开始修复过程。RI-1 可抑制 Eg 的修复过程,因此与 DNA 修复相关的基因被下调。

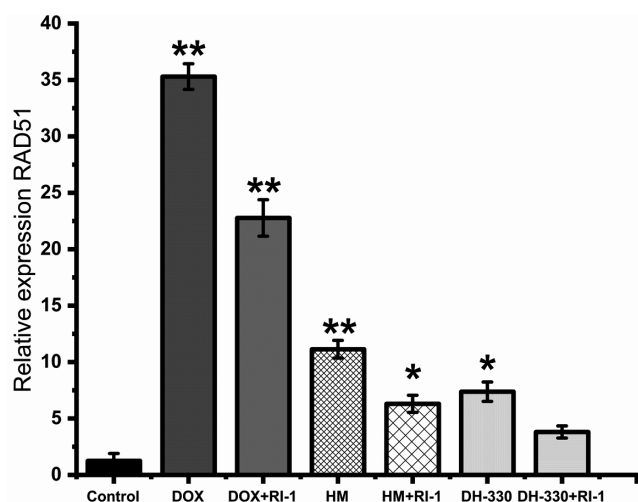


图 11 药物干预后的 EgRAD51 基因表达水平
Fig. 11 The expression level of EgRAD51 gene after drug intervention in parasites

讨论

包虫病的难治性和治疗后因耐药导致的包虫病复发是临床治疗中的难点。药物治疗作为包虫病治疗的重要手段,对于其耐药作用机制以及药物治疗的作用靶点的开发是亟待解决的重要问题。研究发现骆驼蓬碱类化合物具有抗包虫病作用,其机制可能通过导致细粒棘球蚴虫体基因组不稳定、诱发 DNA 损伤以及改变同源重组相关因子 EgRAD51、EgBrca1 和 EgMre11 等的 mRNA 水平^[13]抑制细粒棘球蚴的活性,并发现 EgRAD51 可能参与 DNA 损伤后的修复,如双链断裂 DNA 的同源重组修复^[14]。Rad51 介导的 HR 功能部分或全部恢复可导致肿瘤耐药^[15]。然而, EgRAD51 与骆驼蓬碱类化合物抗包虫治疗之间的关系尚未完全阐明。

随着分子生物学的发展,蛋白的生物信息学分析已经成为科研中的重要环节,主要运用数理和信息科学的观点,理论和方法进行序列比对、基因识别、蛋白质结构预测、蛋白质反应的预测以及建立进化模型等用于研究生命现象,现已广泛应用于药物治疗疾病新靶点的发现,为药物的研究和开发提供参考^[16]。本研究通过采用生物信息学方法预测 EgRAD51 蛋白的理化性质及结构特征分析 EgRAD51 在骆驼蓬碱类化合物抑制细粒棘球蚴中的分子靶点特性。结果表明,该蛋白具有含核苷三磷酸水解酶的 P-loop (P-loop NTPase) 结构域超家族(即:PRK09302 超级家族)的保守结构域,其二级结构中 α -螺旋占 45.24%,延伸主链占 15.08%, β -转角占 3.97%,无规则卷曲占 35.71%。该蛋白为低聚单体,可结合单链和双链 DNA,表现出 ATP 依赖性,通过 DNA 分子间的同源配对和链交换进行修复,再结合前期研究发现的

EgRAD51 蛋白可直接或间接地与 Eg53、EgBrca1 和 EgMre11 有相互作用,进一步说明 EgRAD51 可能增强细粒棘球蚴对骆驼蓬碱类化合物的耐受,适当下调 DNA 的修复功能,可能对包虫病治疗具有促进作用。

预测 EgRAD51 含有 13 个 B 细胞抗原表位,取 ABC pred 和 IEDB 预测交集综合分析共得到 4 个 B 细胞抗原表位(32-115,171-181,187-222,343-347),表明该蛋白能刺激机体的体液免疫应答。同源性及系统进化分析显示,EgRAD51 蛋白与多房棘球绦虫的 RAD51 氨基酸序列一致,与亚洲带绦虫、中殖孔绦虫、血吸虫、欧洲锥虫等亲缘关系较近,与智人、果蝇、中华枝睾吸虫、带翅棘球蚴亲缘关系较远,提示可能在 EgRAD51 靶向药物治疗中很大程度上能避免宿主的干扰。

陈苗生等^[17]报道,DNA 修复功能水平增高能增强肿瘤细胞对化疗药物的抵抗力,促进肿瘤细胞的转移。本研究中药物干预 Eg 后 EgRAD51 基因检测结果提示 HM 及其衍生物 DH-330 可对 Eg 造成 DNA 损伤,而 Eg 应激上调与 DNA 修复相关的基因,RAD51 的主要功能是促进和执行同源重组和修复,Rad51 作为 HRR 过程中的中心重组酶通过同源重组途径修复 DSBs,在维持基因组完整性和细胞生命功能方面发挥重要作用^[18]。因此推测 HM 及其衍生物可能通过激活 RAD51 等修复基因开始修复过程。EgRAD51 基因是同源重组修复通路上的重要基因,药物干预 Eg 后 EgRAD51 基因表达量分析发现,DNA 损伤修复抑制剂可使 RAD51 基因低表达,而单独使用骆驼蓬碱类化合物可使 EgRAD51 基因表达上调,表明 EgRAD51 与 DNA 损伤修复有关,可为抗包虫病药物的研发提供参考。

【参考文献】

- [1] Wen H, Vuitton L, Tuxun T, et al. Echinococcosis: Advances in the 21st Century[J]. Clin Microbiol Rev, 2019, 32(2): e00075-18.
- [2] 崔芳芳,赵杰,宇传华,等. 1990-2016 年中国囊型包虫病疾病负担变化趋势研究[J]. 现代预防医学, 2019, 46(2): 193-197, 220.
- [3] 黄嫣,伍卫平,韩帅,等. 2018-2019 年全国棘球蚴病监测分析[J]. 中国病原生物学杂志, 2021, 16(9): 1025-1029.
- [4] 郑灿军,杨柳,张光霞,等.《包虫病防治技术方案(2019 年版)》解读[J]. 热带病与寄生虫学, 2020, 18(4): 193-196, 201.
- [5] 李松,韩秀敏,郭亚民. 非苯并咪唑类药物用于棘球蚴病治疗研究进展[J]. 中国血吸虫病防治杂志, 2021, 33(2): 213-217.
- [6] Bonilla B, Hengel SR, Grundy MK, et al. RAD51 Gene family structure and function[J]. Annu Rev Genet, 2020, 54: 25-46.
- [7] Xu JF, Zhao LY, Xu YY, et al. Cryo-EM structures of human RAD51 recombinase filaments during catalysis of DNA-strand exchange[J]. Nat Struct Mol Biol, 2017, 24(1): 40-46.
- [8] Di Giorgio C, Delmas F, Ollivier E, et al. In vitro activity of the

- beta-carboline alkaloids harmine, harmine, and harmaline toward parasites of the species *Leishmania infantum* [J]. Exp Parasitol, 2004, 106(3-4): 67-74.
- [9] Lu S, Wen LM, Gong YH, et al. *In vitro* effects of harmine against *Echinococcus granulosus* protoscoleces by stimulating DNA damage [J]. Exp Parasitol, 2021. 226-227: 108121.
- [10] Gong YH, Tian CY, Lu S, et al. Harmine combined with Rad54 knockdown inhibits the viability of *Echinococcus granulosus* by enhancing DNA damage [J]. DNA Cell Biol, 2021, 40(1): 1-9.
- [11] Gong YH, Lv SZ, Wang JH, et al. Effect of harmine and its derivatives against *Echinococcus granulosus* and comparison of DNA damage targets [J]. J Biomed Nanotechnol, 2020. 16(6): 827-841.
- [12] Li YF, Wen LM, Zhao J, et al. *In vitro* and *In vivo* efficacy of DNA damage repair inhibitor veliparib in combination with artesunate against *Echinococcus granulosus* [J]. Dis Markers, 2020, 2020: 8259820.
- [13] Bhat KP, Cortez D. RPA and RAD51: fork reversal, fork protection, and genome stability [J]. Nat Struct Mol Biol, 2018, 25(6): 25: 446-453.
- [14] Feng Y, Wang D, Xiong L, et al. Predictive value of RAD51 on the survival and drug responsiveness of ovarian cancer [J]. Cancer Cell Int, 2021, 21(1): 249.
- [15] Qiao Y, Yu X, Zhou B, et al. KIAA1529 regulates RAD51 expression to confer PARP inhibitors resistance in ovarian cancer [J]. Transl Oncol, 2022, 24: 101497.
- [16] 肖静, 曲晨, 赵嘉庆. 细粒棘球绦虫抗原 Fis1 的生物信息学分析 [J]. 中国病原生物学杂志, 2022, 17(2): 191-194.
- [17] 陈苗生, 林翠红, 黄凌艳, 邱晓明. RAD51AP1 在结直肠癌组织中的表达及意义 [J]. 中国现代医生, 2022, 60(21): 23-27.
- [18] Isabel EW, Fumiko E. RAD51: Beyond the break [J]. Seminars in Cell and Developmental Biology, 2020 (prepublish).
- 【收稿日期】 2023-07-25 【修回日期】 2023-10-08

(上接 1415 页)

- [10] 陆飞宇, 李剑侠, 黄先锋, 等. miR-132 靶向 FoxO3a 抑制细胞自噬在脑出血模型大鼠中的神经保护作用 [J]. 脑与神经疾病杂志, 2021, 29(10): 629-634.
- [11] 刘韵, 肖丽, 曹明艳, 等. 蛋白激酶 B 抑制剂对糖尿病大鼠视网膜血管新生和炎症反应抑制作用的研究 [J]. 中国糖尿病杂志, 2020, 28(2): 127-132.
- [12] van de Beek D, Brouwer MC, Koedel U, et al. Community-acquired bacterial meningitis [J]. Lancet, 2021, 398 (10306): 1171-1183.
- [13] 高蒙, 黄娟. 白藜芦醇通过调节 miR-223-3p/NLRP3 通路发挥对幼年大鼠脑膜炎模型皮质神经元的保护作用 [J]. 中国病理生理杂志, 2020, 36(7): 1161-1169.
- [14] Depoortere S, Toeback J, Lunsken S, et al. Delayed cerebral thrombosis complicating bacterial meningitis [J]. Acta Clin Belg, 2022, 77(2): 462-469.
- [15] Walgrave H, Balusu S, Snoeck S, et al. Restoring miR-132 expression rescues adult hippocampal neurogenesis and memory deficits in Alzheimer's disease [J]. Cell Stem Cell, 2021, 28(10): 1805-1821.
- [16] 张洋, 陈美月, 崔莹, 等. 丹参酮 II A 对缺氧缺血性脑病新生大鼠海马组织中 miR-132 表达的调节作用及其机制 [J]. 吉林大学学报(医学版), 2022, 48(4): 905-914.
- [17] McLoughlin NM, Mueller C, Grossmann TN. The Therapeutic Potential of PTEN Modulation: targeting strategies from gene to protein [J]. Cell Chem Biol, 2018, 25(1): 19-29.
- [18] Takeuchi A, Koga K, Satake E, et al. Endometriosis triggers excessive activation of primordial follicles via PI3K-PTEN-Akt-Foxo3 pathway [J]. J Clin Endocrinol Metab, 2019, 104(11): 5547-5554.
- [19] Zhao Y, Zhao R, Wu J, et al. Melatonin protects against A β -induced neurotoxicity in primary neurons via miR-132/PTEN/AKT/FOXO3a pathway [J]. Biofactors, 2018, 44(6): 609-618.
- [20] Zhao M, Gao J, Cui C, et al. Inhibition of PTEN ameliorates secondary hippocampal injury and cognitive deficits after intracerebral hemorrhage: involvement of AKT/FoxO3a/ATG-mediated autophagy [J]. Oxid Med Cell Longev, 2021, 2021(1): 547-553.
- 【收稿日期】 2023-08-12 【修回日期】 2023-10-25