



(12) 发明专利

(10) 授权公告号 CN 101892239 B

(45) 授权公告日 2012.05.23

(21) 申请号 201010224645.5

C12Q 1/68(2006.01)

(22) 申请日 2010.07.15

G01N 33/53(2006.01)

(73) 专利权人 中国农业科学院上海兽医研究所
地址 200241 上海市闵行区紫月路 518 号

A61K 48/00(2006.01)

A61K 38/17(2006.01)

A61P 33/12(2006.01)

(72) 发明人 王孝波 苑纯秀 李红飞 冯新港
杨健美

审查员 张丽玮

(74) 专利代理机构 上海大邦律师事务所 31252
代理人 周东萍

(51) Int. Cl.

C12N 15/12(2006.01)

C07K 14/705(2006.01)

C07K 16/28(2006.01)

C12N 15/63(2006.01)

C12N 1/15(2006.01)

C12N 1/19(2006.01)

C12N 1/21(2006.01)

C12N 5/10(2006.01)

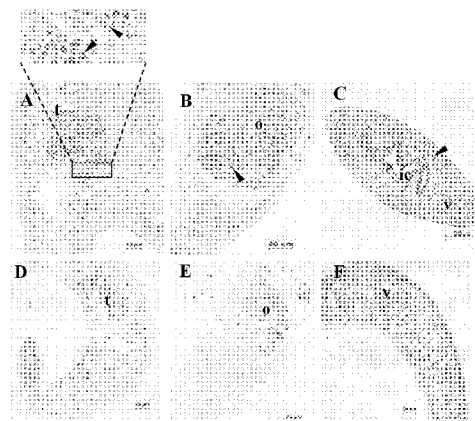
权利要求书 1 页 说明书 27 页 附图 4 页

(54) 发明名称

日本血吸虫 Frizzled9 基因、蛋白及用途

(57) 摘要

本发明公开了一种日本血吸虫基因,所述基因具有编码 Frizzled(Fz) 家族蛋白的核苷酸序列,该核苷酸序列包含:编码 SEQ ID NO. 1 氨基酸序列的 DNA 序列,或编码 SEQ ID NO. 1 所示氨基酸序列中缺失、添加、插入或取代一个或多个氨基酸所得到的具有 SEQ ID NO. 1 相同活性的氨基酸序列的 DNA 序列。本发明还公开了上述基因编码的蛋白质。本发明的日本血吸虫基因和蛋白,是日本血吸虫 Fz 受体家族的新成员,参与调控血吸虫的生长发育及生殖系统发育,对开发阻断血吸虫生长发育及生殖系统发育的新药具有很高的应用价值。



1. 一种日本血吸虫基因,其特征在于,所述基因为编码SEQ ID NO. 1所示氨基酸序列的Frizzled9 基因。
2. 根据权利要求1所述的日本血吸虫基因,其特征在于,所述基因的核苷酸序列为SEQ IDNO. 2 所示的 DNA 序列。
3. 一种日本血吸虫蛋白质,其特征在于,所述蛋白质为SEQ ID NO. 1所示氨基酸序列的Frizzled9 蛋白质。
4. 一种重组载体,其特征在于,包含编码 SEQ ID NO. 1 中 22-213 位氨基酸序列的 DNA 序列。
5. 一种宿主细胞,其特征在于,包含权利要求4所述的重组载体。
6. 一种抗血清,其特征在于,该抗血清是用权利要求4所述重组载体表达的重组蛋白免疫小鼠而制得。

日本血吸虫 Frizzled9 基因、蛋白及用途

技术领域

[0001] 本发明涉及生物工程技术领域,尤其涉及一种日本血吸虫 Frizzled9 基因、蛋白及用途。

背景技术

[0002] 日本血吸虫病是一种分布广泛、危害严重的人畜共患寄生虫病。至今仍有 76 个国家、近 2 亿人口受到该病的威胁。近 20 年来,我国对日本血吸虫病的防治主要依赖于药物(吡喹酮)化疗为主的综合措施,但长期单一用药不能解决重复感染问题,也容易引发血吸虫的耐药性,因此新的治疗药物亟待开发。然而,要使新药研究取得突破性进展,充分了解血吸虫的生长发育过程及其相关的调控机制是重要基础。

[0003] 日本血吸虫生活史复杂,在终末宿主体内虫体的生长发育过程都呈现显著的生物学和形态学变化,但是,以往研究者对血吸虫的研究主要集中于表观生物学及组织化学层面,对于虫体复杂的生长发育调控机理缺乏认识,因而在新药开发方面没有新的突破。生物体的生长发育由细胞间的信号网路系统调控,日本血吸虫也不例外。另外,全基因组测序分析发现,日本血吸虫存在 Wnt、Notch 等信号通路。因此对信号通路的研究有助于更好的理解血吸虫生长发育机制,也为新药的开发提供新思路。

[0004] 其中,Wnt 信号通路普遍存在于果蝇、小鼠、爪蟾等真核生物中,是调控胚胎发育以及成年生物体内环境平衡的重要信号途径。许多研究报道发现该通路的异常调控会导致胚胎发育发生表型突变,甚至肿瘤等疾病的产生。Wnt 信号通路的激活,是由 Wnt 信号分子(一种分泌性糖蛋白),与细胞膜表面的 Frizzled(Fz) 家族受体结合触发下游信号的传递,且涉及许多不同的信号元件。

[0005] 关于 Wnt 信号通路如何调控日本血吸虫生长发育的分子机制已经展开初步研究,目前已克隆到 SjWnt4 信号分子(陶丽红,姚利晓,傅志强等。日本血吸虫信号转导蛋白 SjWnt4 基因的克隆、表达及功能分析。生物工程学报 2007 年 5 月,23 卷 3 期:392-397) 和 SjWnt10a 信号分子(陶丽红,姚丽晓,苑纯秀等。日本血吸虫信号转导蛋白 SjWnt10a 基因的克隆及其在童虫和成虫中 mRNA 表达量的变化。中国兽医科学,2007,37(02):93-97)、SjFz5 和 SjFz7 受体分子及其他成员组分,还有待于对它们进行具体的生物学功能分析。Wnt 信号分子(配体)与 Fz 受体的相互作用是 Wnt 途径的起始步骤,阐明它们的相互作用机制,对了解血吸虫 WNT 途径的发育生物学功能具有重要意义。此外,由于一种 Wnt 可能和 Fz 受体家族不同成员作用,不同的 Wnt 也可能作用于同一种 Fz 受体,即受配体的结合主要取决于 Fz 受体及其所在的组织环境,因此有必要对日本血吸虫 Fz 家族成员的种类、表达特征、组织分布等进行研究。

发明内容

[0006] 本发明要解决由于对日本血吸虫的生长发育调控机制缺乏认识而导致在新药研发方面难以取得突破的技术问题,提供一种日本血吸虫 Frizzled9 基因(简称 SjFz9 基因)

和蛋白,该 SjFz9 基因是日本血吸虫 Fz 受体家族的新成员,参与调控血吸虫的生长发育及生殖系统发育,对开发控制血吸虫生长发育及生殖生理的新药具有较高的应用价值。

[0007] 此外,还需要提供一种日本血吸虫 Frizzled9 基因和蛋白质的用途。

[0008] 为了解决上述技术问题,本发明通过如下技术方案实现:

[0009] 在本发明的一个方面,提供了一种日本血吸虫基因,所述基因具有编码 Frizzled 家族蛋白的核苷酸序列,该核苷酸序列包含下述 DNA 序列之一:

[0010] (1) 编码 SEQ ID NO. 1 氨基酸序列的 DNA 序列;

[0011] (2) 编码 SEQ ID NO. 1 所示氨基酸序列中缺失、添加、插入或取代一个或多个氨基酸所得到的具有 SEQ ID NO. 1 相同活性的氨基酸序列的 DNA 序列。

[0012] 优选的,所述基因具有 SEQ ID NO. 2 所示的 DNA 序列。

[0013] 在本发明的另一方面,提供了一种日本血吸虫蛋白质,所述蛋白质选自:

[0014] (1) 具有 SEQ ID NO. 1 氨基酸序列的蛋白质;

[0015] (2) 具有 SEQ ID NO. 1 中部分氨基酸序列的活性片段或保守性变异蛋白质;

[0016] (3) SEQ ID NO. 1 所示氨基酸序列中缺失、添加、插入或取代一个或多个氨基酸所得到的具有 SEQ ID NO. 1 相同活性的氨基酸序列的蛋白质。

[0017] 优选的,所述蛋白质具有 SEQ ID NO. 1 所示氨基酸序列。

[0018] 在本发明中,由于体内或纯化期间蛋白质本身的修饰等或由于编码蛋白基因的多态性及变异,天然存在的蛋白质会出现氨基酸的缺失、添加、插入、取代或其他变异。事实上,存在着一些生理和生物活性上基本等同于无变异蛋白质的蛋白。这些结构不同于相应的蛋白质但与该蛋白质没有明显的功能差异的多肽或蛋白称为功能等同变异体。

[0019] 功能等同变异体同样适用于通过人工手段向一种蛋白质的氨基酸序列中导入这类变异而制成的多肽。尽管这样能获得更多不同形式的变异体,但所得的变异体作为功能等同变异体的前提是其生理活性基本等同于原始无变异蛋白质的活性。

[0020] 例如,人白介素 2 (IL-2) 的氨基酸序列中用丝氨酸替换特定的半胱氨酸残基所得到的多肽仍保留着 IL-2 的活性 [Science, 224, 1431 (1984)]。

[0021] 此外,在用基因工程方法生产蛋白质时,常把所需蛋白质表达为融合蛋白质,例如,把源于其他蛋白质的 N-末端肽链加到所需蛋白质的 N-末端以提高所需蛋白质的表达,或者把一个合适的肽链加到所需蛋白的 N-或 C-末端,表达该蛋白并使用一种对所加肽链具有亲和力的载体使所需蛋白的纯化变得更加容易。

[0022] 就决定基因上特定氨基酸的密码子(三联体碱基组合)而言,每种氨基酸存在 1 到 6 个密码子。因此尽管依赖于氨基酸序列,但仍有许多编码一种氨基酸的基因。在自然界中,基因并不稳定,常发生核酸变异。基因的变异可能不影响所编码的氨基酸序列(沉默变异),这种情况下,会产生编码相同氨基酸序列的不同基因。因此,即使分离出了编码特定氨基酸序列的基因,随着含有该基因生物体的传代,仍不可避免地会产生编码相同氨基酸序列的不同基因。

[0023] 用各种基因工程技术人工产生编码相同氨基酸序列的各种基因并不困难。例如,当编码所需蛋白质的天然基因的密码子在用于以基因工程技术产生蛋白质的宿主中可利用性较低、表达的蛋白量不够时,可以人工将密码子转变成在该宿主中可利用性高的另一密码子而不改变所编码的氨基酸序列,即可增强所需蛋白质的表达。因此,这类人工生产的

不同多核苷酸也包括在本发明的范围内,只要它能编码出本发明公开的氨基酸序列即可。

[0024] 另外,由至少一种改变(如蛋白质的氨基酸序列中有一个或多个氨基酸残基的缺失、添加、插入或取代)所得的多肽或蛋白一般具有功能上等同于所述蛋白质的活性,编码这类多肽或蛋白的基因也包括在本发明的范围内,不管它是从天然来源分离还是人工生产的。

[0025] 一般的,编码功能等同变异体的基因是同源的。因此,能与本发明的基因杂交且编码具有 Frizzled9 受体活性的蛋白质的核酸分子也包括在本发明的范围内。

[0026] 在本发明的另一方面,还提供了一种包含上述日本血吸虫基因 DNA 序列的全部或部分序列的重组载体。

[0027] 所述重组载体包括重组克隆载体或重组表达载体,重组表达载体包括重组原核表达载体、重组真核表达载体。

[0028] 优选的,所述重组载体包含编码 SEQ ID NO. 1 中 22-213 位氨基酸序列的 DNA 序列。

[0029] 在本发明的另一方面,还提供了一种宿主细胞,包含上述重组载体,或已用上述基因序列转化或转染。

[0030] 在本发明的另一方面,还提供了一种上述日本血吸虫基因的用途,用于制备预防或治疗血吸虫病的药物;或用于筛选预防或治疗血吸虫病的药物。

[0031] 所述药物通过阻断血吸虫生长发育或阻断血吸虫生殖系统发育来发挥作用。

[0032] 在本发明的另一方面,还提供了一种上述日本血吸虫蛋白质的用途,用于制备预防或治疗血吸虫病的药物;或用于筛选预防或治疗血吸虫病的药物。

[0033] 所述药物通过阻断血吸虫生长发育或阻断血吸虫生殖系统发育来发挥作用。

[0034] 本发明日本血吸虫 Frizzled9 基因,是日本血吸虫 Fz 受体家族的新成员。实时荧光定量 PCR 分析结果表明 SjFz9 基因可能介导了童虫阶段的发育及产卵初期、高峰期雌雄虫的发育调控;免疫组化实验结果显示 SjFz9 基因可能与日本血吸虫的运动相关,并且还参与调控血吸虫成虫生殖系统的发育,为开发阻断虫体发育的新药提供了新思路。

附图说明

[0035] 下面结合附图和具体实施方式对本发明作进一步详细的说明。

[0036] 图 1 是本发明实施例 1 简并 PCR 反应产物的电泳检测图;

[0037] 图 2 是本发明实施例 1 的 SjFz9 基因的 3' 和 5' RACE 产物电泳鉴定图;

[0038] 图 3 是本发明实施例 1 利用生物信息学工具在线预测 SjFz9 蛋白的结构图;

[0039] 图 4 是本发明实施例 1 不同物种的 Frizzled 9 分子亲缘关系和进化分析图;

[0040] 图 5 是本发明实施例 2 的 SjFz9 基因在日本血吸虫不同发育阶段的 mRNA 表达水平差异分析图;

[0041] 图 6 是本发明实施例 3 重组原核表达质粒 pET28a-SjFz9 的酶切鉴定图;

[0042] 图 7 是本发明实施例 3 重组蛋白经纯化后的 SDS-PAGE 电泳图;

[0043] 图 8 是本发明实施例 3 的 pET28a-SjFz9 重组表达蛋白的 Western-blot 鉴定图;

[0044] 图 9 是本发明实施例 3 的 SjFz9 多抗的 Western-blot 检测图;

[0045] 图 10 是本发明实施例 3 免疫组化分析 SjFz9 蛋白在日本血吸虫童虫、成虫体壁肌

肉层细胞的定位图；

[0046] 图 11 是本发明实施例 3 免疫组化分析 SjFz9 蛋白在日本血吸虫成虫生殖细胞的定位图。

具体实施方式

[0047] 下列实施例中,未注明具体条件的实验方法,通常按常规条件,如《分子克隆实验指南》(J. 萨姆布鲁克, D. W. 拉塞尔著,黄培堂,汪嘉玺,朱厚础,等译. 第 3 版,北京:科学出版社,2002) 中所述的方法进行。

[0048] 本发明利用简并引物 PCR 方法从日本血吸虫 13d 童虫 cDNA 中获得 Frizzled 家族新成员——Fz9 同源的 EST,然后通过 RACE 技术获得全长 cDNA,再利用生物信息学软件分析该基因及其编码蛋白的结构。结果表明,该基因具有 Fz 家族受体的基本结构特征:含有一个信号肽,一段保守的半胱氨酸富集区(CRD,也是 Wnt 信号分子的区域),和一段七次跨膜疏水区。利用荧光实时定量 PCR 方法分析了 SjFz9 基因在日本血吸虫发育的不同时期的 mRNA 表达图式。成功制备了能够特异性识别天然虫体中 SjFz9 蛋白的多克隆抗体,并以该制备的多抗作为一抗,利用免疫组化方法对 SjFz9 蛋白进行组织定位研究分析,实验结果显示 SjFz9 可能与日本血吸虫的运动相关,并且还参与调控血吸虫成虫生殖系统的发育。表明该 SjFz9 基因编码蛋白可能对控制血吸虫生长发育及生殖生理的新药开发具有较高的应用价值。

[0049] 实施例 1SjFz9 新基因的克隆

[0050] 1. 血吸虫虫体材料的收集及总 RNA 的提取

[0051] 日本血吸虫(安徽株)尾蚴感染雄性新西兰杂交兔(3kg/只),腹部贴片,2,000~10,000 条每只。在标准卫生条件下饲养,以肝门静脉灌注法收集感染兔体内不同期别的虫体,包括 7d、13d、18d、23d、26d、29d、32d、35d、42d 和 50d 虫体,并将 26d、32d、35d、42d 和 50d 虫体进行雌雄分开。将感染 42d 的兔肝组织剪碎、匀浆后,通过多重筛网过滤纯化方法收集兔肝组织中的虫卵。Trizol(invitrogen)抽提虫卵及各期别虫体总 RNA(具体操作参照 Trizol 说明书进行),DNase I(Takara)除去残留 DNA。1%非变性琼脂糖凝胶电泳检测 RNA 完整性,分光光度计(Eppendorf A G,Hambury)检测浓度,DEPC 水溶解的 RNA 于 -80℃ 冰箱保存。

[0052] 2. 日本血吸虫 Fz9EST 新基因的获得

[0053] 参考文献(Daudet N, Ripoll C, Molès JP, et al. Expression of members of Wnt and Frizzled gene families in the postnatal rat cochlea. Mol Brain Res, 2002, 105 :98-107 ;Wang Y, Macke JP, Abel la BS, et al. A large family of putative transmembrane receptor homologous to the product of the Drosophila tissue polarity gene frizzled. J Biol Chem. 1996, 271 :4468-4476),根据已获得的两个日本血吸虫 Frizzled 受体:SjFz5(Genbank 登录号:EU370926)、SjFz7(Genbank 登录号:EU370927),与其他物种的 Fz 家族成员比对结果,选取 YPERP、WWVIL 及 VGIT 氨基酸保守位点,根据氨基酸的简并性设计简并引物(见表 1),引物由英俊生物技术公司合成。

[0054] 表 1 钓取日本血吸虫 Fz 家族新基因的 PCR 反应简并引物

[0055]

Primers	Sequences (5' - 3')
Fzp1 / 2	F: NNNGAATTCTAYCCNGARMGNCCNAT (SEQ ID NO. 3) R: NNNAAGCTTNARNAYNACCCACCA (SEQ ID NO. 4)
Fzp1 / 3	F: NNNGAATTCTAYCCNGARMGNCCNAT (SEQ ID NO. 5) R: NNNAAGCTTNGTDTANCCNTC (SEQ ID NO. 6)

[0056] 表 1 中, F :上游引物, R :下游引物 ;根据核苷酸标准命名法确定简并引物序列 (N = A/C/G/T ;R = A/G ;M = A/C ;Y = C/T ;D = A/G/T)。

[0057] 以保存的血吸虫 13 天童虫 cDNA 为模板进行 touch-up PCR。PCR 反应条件 : 94 °C 5min ;(94 °C 1min, 50 °C 1.5min, 72 °C 2min) × 5cycles ;(94 °C 1min, 53 °C 1.5min, 72 °C 2min) × 35cycles ;72 °C 再延伸 10min。Touch-up PCR 开始以低于理论 T_m 值的温度退火, 确保引物能与模板结合 ;然后逐渐升高退火温度, 随着循环的增加特异性扩增产物得到积累, 最终得到较特异的 PCR 产物。1.2% 的琼脂糖凝胶电泳检测扩增产物, 目的片段切胶回收后与 pGEM-T Easy 载体连接, 转化大肠杆菌 JM109 感受态细胞, PCR 鉴定为阳性的克隆由上海桑尼生物科技有限公司测序。

[0058] 结果 :以简并引物进行 touch-up PCR (上升 PCR) 可见多个条带 (见图 1)。在图 1 中, 1 :以 Fzp1/2 为引物对进行 PCR 反应的产物 ;2 :以 Fzp1/3 引物对进行 PCR 反应的产物。由设计引物的保守氨基酸序列间长度推测产物长度在 200 ~ 400bp 之间, 选择图 1 中箭头所指约 250bp 大小的条带进行克隆。挑选 150 个阳性克隆测序, 测序结果经 blastx 在线分析, 发现其中一个克隆的序列与其它物种 Fz 家族成员相似, 且不同于已获得的 SjFz5 和 SjFz7 基因, 是血吸虫 Fz 家族的新成员。

[0059] 3. 日本血吸虫 SjFz9 新基因全长克隆及生物信息学分析

[0060] 利用 RACE (rapid amplification of cDNA ends, cDNA 末端快速扩增技术) 方法获得 SjFz9 基因全长序列。以获得的 SjFz9EST 序列为模板设计 3' 和 5' RACE 引物 (见表 2)。利用 SMARTMRACE cDNA Amplification kit (clontech) 将 13d 童虫 RNA 反转录成 3' race ready cDNA 和 5' race ready cDNA (参照试剂盒说明书完成)。1.0% 琼脂糖凝胶电泳检测 PCR 产物, 扩增片段切胶回收后与 pMD18T 载体 (TaKaRa) 连接, 转化 DH5 α 感受态细胞, 阳性克隆测序结果用 gene tool 软件进行电子拼接。利用常规生物信息学软件对 SjFz9 基因及其编码蛋白结构进行预测分析。

[0061] 表 2SjFz9RACE PCR 反应用引物

[0062]

Primers	Sequences (5' - 3')
3' GSP1/NGSP1	CTATCCGGAGCGACCCATTATTTAC (SEQ ID NO. 7) ACACTCTCTGATTGGGCGTGATTTG (SEQ ID NO. 8)
5' GSP1/NGSP1	TGACCCACCATAAATAACTCGCTTG (SEQ ID NO. 9) CACCATGTACCTTCATATCCTGAAG (SEQ ID NO. 10)
3' GSP2/ NGSP2	TAACGGAATACGGTACAAGTTGGTC (SEQ ID NO. 11) TATCACTTCAGGAATGTGGGTATGG (SEQ ID NO. 12)
5' GSP2/NGSP2	TAAGTTCGATTGGCCATGGAAAGCT (SEQ ID NO. 13) ACCAACATTGATGAGGGAACGATAG (SEQ ID NO. 14)

[0063] 结果 :

[0064] 以该克隆序列为模板, 利用引物 3' GSP1/NGSP1 和 5' GSP1/NGSP1 (表 2) 进行 3'

和 5' RACE PCR (RACE PCR 产物电泳鉴定图见图 2)。图 2 中, 1 : 3' RACE 产物 ; 2 : 5' RACE 产物。其中箭头所指片段为切胶回收作克隆用。利用 gene tool 软件将两端延伸后的序列与模板序列进行拼接, 拼接后的序列 (2572bp) 在 NCBI 网站的 ORF Finder 及 blastp 在线分析, 预测的 ORF 编码的 Fz 蛋白结构不完整。在首次拼接序列基础上再设计 3' GSP2/NGSP2 和 5' GSP2/NGSP2 (表 2), 进行第二次 RACE 延伸。同样方法拼接得到的序列为 3670bp, 再次预测出的 ORF 编码蛋白具有 Fz 家族蛋白的完整结构, 至此获得了血吸虫 Frizzled 家族新成员的全长 cDNA 序列。氨基酸序列相似性分析显示, 该新获得 Fz 家族成员与其他物种的 Fz9 相似性最高, 其中与斑马鱼、小鼠和人的同源性分别为 33%、32% 和 31%, 因此将该新获得的基因命名为日本血吸虫 Frizzled9 基因, 简称 SjFz9 基因。

[0065] SjFz9 基因的 cDNA 全长 3670bp (SEQ ID NO. 2), 146 ~ 2917bp 为编码区序列, 编码 923 个氨基酸。理论相对分子量为 103.97Ku, 理论等电点 8.65。

[0066] SjFz9 基因的核苷酸和氨基酸序列如下, 起始、终止密码子均用方框标出 ; 曲线标出的序列为信号肽 ; 连续 10 个半胱氨酸为保守的 CRD 结构, 推测为 Wnt 结合位点 ; 下划直线处为七个跨膜区段。

[0067] 1 GGTCAAATAACAATTTAACATTATGAAAGATATATTTTCAAGTGTATATTTATGAAT
GTACTTACTTGTA

[0068] 71 AAATGAAGTCATCTTAACAGAAGCTGAGCTGTTTAGTTGCTTTGTTGTGTTAATTTACATAAA
TTGTCACTCAGAG

[0069] 146 ATG AAAATTGTAAGAATGAGTTTATTTAAAAAATTTAATTTGGTCAATTTCATAAATATTATAAC
AGTCATGTTA

[0070] I M K I V R M S L F K K F N L V N F I N I I T V M L

[0071] 221 ACTCAATGTCTCTCAGATACTACCTGTGCTGTCTATAAAATTCGAAGCTCAAATTTTCACTT
TTATGATCAAATT

[0072] 26 T Q C L S D T T C A V Y K I R S S N F H F Y D Q I

[0073] 296 GATAAAGATGTTAATTTCAAGTGGAAAGTATTTGCCTCGCTTACGAACACTAAGAGATGTTAT
ACCTCTCATGTTT

[0074] 51 D K D V N F S G K Y L P R L R T L R D V I P
L M F

[0075] 371 CCAATAAATTCAAAA GAAATTGAAGTACCCGATCATTTGAAAAATCTAATTGAAGTTAATC
AAAATGGAAACCTA

[0076] 76 P I N S K E I E V P D H L K N L I E V N Q N
G N L

[0077] 446 CCGAATCAAGTTTATCCAGATGAGAGTCCAGGTGAACAACAAGGAGAAATTGAAAATGGTAT
AGATGATGTAAAT

[0078] 101P N Q V Y P D E S P G E Q Q G E I E N G I D D
V N

[0079] 521 CACTGTATTCCAATTCATGAACCAATTTGTCATGGTCTTCAATATACTCATACTTGGTTACC
TAATTCAGTGGGT

[0080] 126 H C I P I H E P I C H G L Q Y T H T W L P N

S V G

[0081] 596 CTAACGACACAAGAAGAGGCTGCAAAACGCATGAATGACTATCGTTCCCTCATCAATGTTGG
TTGTTCACTAT

[0082] 151 L T T Q E E A A K R M N D Y R S L I N V G C
S H Y

[0083] 671 TTAAAATTTTTCTG TGTACAGTTATTTTCCAATGTGTACCCCAACTGAATCCACCAG
CTGCCTTACAGCCA

[0084] 176 L K F F L C T V Y F P M C T P T L N P P A A
L Q P

[0085] 746 TGTCAAATTTATGTCGTTATGTACAATCAAAGTGTGAACCAATTATGAAAAGTTTTAGCTT
TCCATGGCCAATC

[0086] 201 C Q N L C R Y V Q S K C E P I M K S F S F P
W P I

[0087] 821 GAACTTAATTGTAAAGCATTACCCTCAGATTCTGGAATGTGTATAACAACCACAAAATTATGA
TTTAGATAAACAA

[0088] 226 E L N C K A L P S D S G M C I Q P Q N Y D L
D K Q

[0089] 896 GGACAAATTATAAGCCCAGGGATCCCATCTGAAGCCCATGGAGCTATAGCTAATTTATTGCC
AAATTTTAGTGAG

[0090] 251 G Q I I S P G I P S E A H G A I A N L L P N
F S E

[0091] 971 TTTTTGGCCCAACAA AGTGCTCAAACCAACTTTTGGCGAGTACACAGAACAATCAAACAG
ATAATACTAACGAG

[0092] 276 F L A Q Q S A Q N Q L L A S T Q N N Q T D N
T N E

[0093] 1046 AAAATACTTTCAGTGAATGATAACCAATCCGTTTGACAAGACATCATCTAATCCACAGGAC
AAAATACTAACATCT

[0094] 301 K I L S V N D N Q S G L T R H H L I H R T K
L Q S

[0095] 1121 TCAAAGTGCTATGTTGTCGAAATTCTTTTATATTCTGAACAGGCACATAATAATATTACATG
CGCTCGTCGTTGT

[0096] 326 S K C Y V V E I L L Y S E Q A H N N I T C A
R R C

[0097] 1196 AATGCTCATTTATTTTACAAACCAGTTGAAAAACGTTTCGCCAATATTTGGATGCTAGTTTG
GTCCATAATATGT

[0098] 351 N A H L F Y K P V E K R F A N I W M L V W S
I I C

[0099] 1271 TTGGGATCGTGTGCATTGACTATTATCACCTTACCTTAAACCGCACACGTTTTGCTTATCC
AGAACGACCGATT

[0100] 376 L G S C A L T I I T F T L N R T R F A Y P E
R P I

[0101] 1346 ATTTACATATCAGTTTGTGCTTCGTTTACTCTGCTGGATTTATTTTACACTCTCTGATTGG
GCGTGATTTGGTG

[0102] 401 I Y I S V C C F V Y S A G F I L H S L I G R
D L V

[0103] 1421 GCATGTAGGAACAAAGTTGAAACAATTTCCGCATCATTATTGACAAACGGGGTATCACAAGT
TGGTACAAAGAAA

[0104] 426 A C R N K V E T I S P S L L T N G V S Q V G
T K K

[0105] 1496 CTTTTTACAGGTTTAACTCAATTAAGTTAACGCTAGTTTCTTGATAACTTCAGGATATGA
AGGTACATGGTGT

[0106] 451 L F T G L T Q L K V N A S F L I T S G Y E G
T W C

[0107] 1571 ACTATTATATTTATT ATCCTGTTTTATTTTCAGTCAAGCGAGTTATTTATGGTGGGTTATGT
TAGCTACCTCATGG

[0108] 476 T I I F I I L F Y F S Q A S Y L W W V M L A
T S W

[0109] 1646 TTTTTATCAGCTTCCTGTAAGTGGGGATGTGAAGGCATAGAGGCTGTCAGTAGTATCTTCCA
TATGCTCGCCTGG

[0110] 501 F L S A S C K W G C E G I E A V S S I F H M
L A W

[0111] 1721 GCATTACCAGCATTA AAAACTATTATCATATTAATCATGCATCGTATTGACGCAGATGAGTT
AACTGGATTATGC

[0112] 526 A L P A L K T I I I L I M H R I D A D E L T
G L C

[0113] 1796 AACGTTGGTTATCAAAATTC AACGAGTTTACTCACATTGATATTACTACCACAAATCATCTA
CCTATTGTTGGGC

[0114] 551 N V G Y Q N S T S L L T L I L L P Q I I Y L
L L G

[0115] 1871 ATCATGTTTTTATTG GTTGGTTTTCACTCCTTGATTTCCCTTCGATCAA ACTTAAAACAAC
ACGCTAAAAGCCTC

[0116] 576 I M F L L V G F H S L I S L R S N L K Q H A
K S L

[0117] 1946 TTGCCTGCTACCAAAACAACAACAGCACCAGCAA ACTGTTATAATTCTCAGTGCAAAAAGAAA
TGCTTCAGCAACA

[0118] 601 L P A T K T T T A P A N C Y N S Q C K R N A
S A T

[0119] 2021 GTAGCAATACCAACTATGACTACTACAAACAATGACCACATCTAGTCCAGTCATCGCCCTCAA

CAATGGAAATATA

[0120] 626 V A I P T M T L Q T M T T S S P V I A L N N
G N I

[0121] 2096 CGACGTTTAGATAAACTAATGATAAAAATAGGCATATTTTCAATGCTGTATATTATTCCATG
TGCATGTGTTATT

[0122] 651 R R L D K L M I K I G I F S M L Y I I P C A
C V I

[0123] 2171 GGTGTGAATATTTACAATTATATTAATTATCCGGAATGGATGAAAACATTGGATAGACTATC
TAAACAAACAAAC

[0124] 676 G V N I Y N Y I N Y P E W M K T L D R L S K
Q T N

[0125] 2246 TGTTTAACGGAATACGGTACAAGTTGGTCACTGGTCACTCAGTGTCTAGAAGATAATCACT
TCCCATCAGCTGAA

[0126] 701 C L T E Y G T S W S L V T Q C L E D N H F P
S A E

[0127] 2321 GCTAACATGTTACAAATATTCATGTCTCTAGTCGTCGGTATCACTTCAGGAATGTGGGTAT
GGTGTAATAGAAAA

[0128] 726 A N M L Q I F M S L V V G I T S G M W V W C
N R K

[0129] 2396 TCTTCAGAAACATGGTTAAAATGCATAAAGAGGAGAAAGAAAAATTTATCTCAGCATAGTT
TGAATTACCGTTAT

[0130] 751 S S E T W L K C I K R R K K N L S Q H S L N
Y R Y

[0131] 2471 CATATTTGTGGGGATGTTGTCGCTGGTGATACTGTTGGAGAACTAATAATGTGTCCCAAG
TTGATCAAAAAGGCA

[0132] 776 H I C G D V V A G D T V G E T N N V S Q V D
Q K A

[0133] 2546 CAACTAATGAATACGACTTGTGTAATTCAAGAGCGGCATGTGGAAAAACAACCAGTATC
CAGGAAACACCATG

[0134] 801 Q L M N T T C C N S R A A C G K N N Q Y P
G N T M

[0135] 2621 AATTCGATAATACAATGAACAATATTAACCAGCTGGTGGCTGATTCCACTCAGTATCATC
AACAACAACACCAA

[0136] 826 N F D N T M N N I N Q L V A D S T Q Y H Q Q
Q H Q

[0137] 2696 AATGAAATCGCACGAAACAACGTAAGAAATTTAATGCTTTCTGGATCATTTAACATGTCAT
CTGGAACAGCTTCA

[0138] 851 N E I A R N N V R N L M L S G S F N M S S G
T A S

[0139] 2771 TGTACAAGTGCTCCACTATTATCACATGGTTTAGGAACAATAAATATGTCATTATTAATTCATTTGAATCCAAT

[0140] 876 C T S A P L L S H G L G T I N M S L L N S F E S N

[0141] 2846 TATTGCACCCGTTTCATCACAACGACAAACATTTAAAACCTAGACCATCTAATACGGGAAGAC CAGACTGT TTT

[0142] 901 Y C T R S S Q R Q T F K T R P S N T G R P D C *

[0143] 2921 GTTTGCTTCCTTGGTAAATAATGTAAATTTTTCTTGAATACACTTACCCATGATTACAAAT TTATTTGTAACAGC

[0144] 2996 TGATATTTACTCTACATTGCATAAGCCATTGCACAAACAAAACCTTGAGCGAAGATCAAAT ATGAAAAATAATCA

[0145] 3071 TAAATCTTGTTTCATCTCATTTTAATTTTGTATCAATTGATCGTTCATATTACATTCTGTAA TCATTTTCTAAATG

[0146] 3146 ATAGTTCCTTCGTTCTGCTTGTTCATTACGATTTATCAATTTGAACAAGTGATTACGCC TATCCGCACAGAAT

[0147] 3221 GCTCATTCTGTATATAAATTATTTTGTGACTTGTATTCCGGTTATCTTTAACTTGAAAA TAATTTCAATTTATT

[0148] 3296 CTTTATTTTTTAACTACAAATCAAGAAGCTATTGCTTGTGCTTCACCTATTATGCTTTAA TTTGATATTTTTTA

[0149] 3371 CGATTAATTCAATTCAGACTTTGGTCAAACGTTTTCTTTAACTTCACACTTTCAGAAAAA GAAAAACAAACCAT

[0150] 3446 ACTAGAACAGTGGTACGGGTTAGTAATGAGGAATAAATGAACAATAAATAACTCAATTTGA TCACAAAGCTCATT

[0151] 3521 AATCCATTTTCTTGAGTCACCTTGTATATGAGTAACTGTAGACTTGTATTTCCCTTCAGTT TTCATTTGTTAATT

[0152] 3596 TATTCACCTTTGCCAATAATTTAAAGACACTAATTATAGCCTTGAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAA

[0153] 利用生物信息学软件分析 Sjfz9 蛋白结构,表明该蛋白具有 Frizzled 家族蛋白结构特征(见图 3 和 SEQ ID NO. 1):N-端 1~31 个氨基酸残基为信号肽;126~244 个氨基酸残基段为半胱氨酸富集区(cystein rich domain, CRD),含有 10 个保守的半胱氨酸;364~748 氨基酸区段为七个跨膜的 α 螺旋疏水区;C-端尾部位于胞内,含有磷酸化位点。

[0154] 选择来自人(human, NP_003499. 1)、类人猿(Pan troglodytes, XP_527779. 2)、小鼠(Musculus, NP_034376. 1)、斑马鱼(Danio, NP_571586. 1)、爪蟾(Xenopus, AAD44332. 1)、斑马(Taeniopygia, XP_002187952. 1)等几个物种的 Fz9 蛋白以及曼氏血吸虫 Frizzled 受体(CAZ36662)进行氨基酸序列的多重比对。结果显示, Sjfz9 蛋白与曼氏血吸虫 Frizzled 受体相似性最高,达 79%,其次与爪蟾 Fz9 的相似性为 33%。另外,参与比对序列中除类人猿在 CRD 区缺少多个半胱氨酸,可能该序列在 5' 端不完整,其余物种的序列在 CRD 区内均含有 10 个保守的半胱氨酸,且跨膜区序列相当保守。从图 4 所示的 Fz9 的进化树上看,

SjFz9 与曼氏血吸虫 Frizzled 受体的亲缘关系最接近,其次,与爪蟾 Fz9 在进化关系上比较接近。

[0155] 实施例 2 荧光定量 PCR 方法分析 SjFz9 基因的 mRNA 表达图式

[0156] 1. 实验方法

[0157] 根据 real time PCR 引物设计原则,设计 SjFz9 基因引物:F(上游引物),5' TATCACTTCAGGAATGTGGGTATGG 3' (SEQ ID NO. 15);R(下游引物),5' TTAGTTTCTCCAACAGTATCACCAG 3' (SEQ ID NO. 16)。选择 18s rRNA 作内参,引物为:F,5' TTAGTTGGTGGAGCGATTTGTCTGG 3' (SEQ ID NO. 17);R,5' CTGTTATTGCTCAATTCGTGCGAC 3' (SEQ ID NO. 18)。按照 PrimeScript™RT reagent Kit (Takara) 说明书,将虫卵及各期别虫体总 RNA 反转录成 cDNA,SYBR Green I (SYBR Premix Ex Taq, Takara) 作染料,在 Rotor gene-3000A (corbett) 仪器上进行 Real time PCR。 $\Delta\Delta$ GT 法计算出各期别虫体中的目的基因相对于内参基因的初始模板量。

[0158] 2. 结果

[0159] 以 18s rRNA 为内参,通过 qPCR 方法对 SjFz9 在血吸虫不同发育阶段的转录水平差异情况进行分析,重复 3 次实验,分析结果如图 5。7、13、18d 的虫体较小,无法进行雌雄分开,因此图 5A 显示的是在雌雄虫混合样品中对 SjFz9 mRNA 表达水平的检测结果,可以看出在血吸虫童虫 (7、13、18d) 阶段,SjFz9 转录水平较高。23d 及其后的成虫虫体样品雌雄分开,图 5B 分析结果显示,23d 雌雄虫中 mRNA 均有较高表达。但 23d 后的成虫体内 mRNA 表达水平变化较大;SjFz9 mRNA 在 26、32、35、50d 雌虫虫体中的表达水平较低,而在 29d、42d 又分别呈现很高的转录水平;SjFz9 mRNA 在 26、29、32、35d 雄虫虫体的表达量较低,而在 42d、50d 虫体又呈现很高表达水平(在 42d 虫体的表达量几乎是 35d 的 6 倍, $P < 0.05$)。

[0160] 实施例 3 SjFz9 蛋白的组织定位分析

[0161] 1. SjFz9 蛋白在原核系统的表达、纯化与鉴定

[0162] 选取 SjFz9 蛋白 CRD 结构功能域作为免疫组化定位的目标抗原,设计引物如下:F,5' GTGGATCCACAGTCATGTAACTCAATGTCTCT 3' (SEQ ID NO. 19);R,5' GTCTCGAGTTCACACTTTGATTGTACATAACGA 3' (SEQ ID NO. 20),上下游引物中分别引入 BamHI 和 Xho I 限制性内切酶位点(引物序列中划线处标出)。以 13d 童虫 cDNA 为模板进行 PCR 扩增。将纯化的目的片段 (567bp) 克隆至 pMD18T 载体,转化大肠杆菌 DH5 α 感受态细胞。提取阳性克隆的质粒,经 BamH I 和 Xho I 双酶切,切胶回收,连接到经同样双酶切的 pET28a 载体中,构建成重组原核表达质粒 pET28a-SjFz9,将该 pET28a-SjFz9 重组质粒转化入 DH5 α 感受态细胞后,阳性克隆进行测序和双酶切鉴定。图 6 所示为重组原核表达质粒 pET28a-SjFz9 的酶切鉴定图,在图 6 中,“1”代表重组表达质粒 pET28a-SjFz9 的 BamH I 和 Xho I 双酶切产物。经鉴定正确的重组表达质粒转化表达用工程菌 BL21 (DE3) 感受态细胞。

[0163] 分别挑取重组质粒转化表达菌的克隆及空载体转化克隆于试管 LB 培养基,37°C 过夜培养。次日,以 1 : 100 转接新鲜 LB 培养基,待菌液 OD₆₀₀ 为 0.8 ~ 1.0 时加入 IPTG 进行诱导表达,IPTG 的终浓度为 0.8mmol/L。8h 后收集菌体诱导物,用 1×PBS 溶解,冰浴超声后,离心分离上清,用 8M 尿素溶解沉淀蛋白,分别取样进行 12% SDS-PAGE 凝胶电泳。参照 Ni 亲和柱使用说明书进行蛋白纯化。纯化的蛋白再利用 SDS-PAGE 凝胶电泳进行检测,结果如图 7 所示。与 His 标签融合的重组蛋白理论分子量为 30.7kD,图 7 中所示纯化的重

组蛋白大小与理论预测一致。

[0164] 用 Western-blot 方法鉴定原核表达的重组蛋白（参照分子克隆操作指南进行）。用含 0.05% 吐温-20 的 TBS 稀释一抗、二抗，以 1 : 100 比例稀释血吸虫卵抗原免疫兔血清作一抗（本实验室制备），以 1 : 2000 比例稀释 HRP 标记的羊抗兔 IgG 作二抗。（沉淀型）TMB（天根生物）为底物进行显色。Western-Blot 结果如图 8 所示，表明 pET28a-SjFz9 原核表达的蛋白能被血吸虫卵蛋白免疫血清识别，具有良好的抗原性。

[0165] 2. SjFz9 多抗的制备及其特异性鉴定

[0166] 取 500 μ l 纯化的重组蛋白 (1mg/ml) 与等体积完全弗氏佐剂混合乳化进行首免，背部脊柱两侧两点皮下注射免疫 6 只雄性 Balb/c 小鼠，免疫剂量为 50 μ g/ 只；以后每隔两周进行加强免疫，将纯化蛋白与不完全弗氏佐剂混合，免疫剂量为 75 μ g/ 只。在第 4 次加强免疫后一周进行摘眼球采血，分离血清。-70 $^{\circ}$ C 冰箱冻存储备用。PBS 与佐剂混合物免疫小鼠作对照。

[0167] 通过 Western-blot 方法鉴定用重组蛋白制备的抗血清。用蛋白抽提试剂盒（Pierce, 美国）抽提血吸虫 18 天虫体蛋白，在蛋白抽提过程中加入蛋白酶抑制剂 PMSF 和 EDTA。抽提的虫体蛋白经 SDS-PAGE 凝胶电泳后电转移到硝酸纤维（NC）膜上（转印方法同上）。以 1 : 500 比例稀释制备的多抗作一抗；以 1 : 2000 的比例稀释 IRDyeTM800 标记的羊抗小鼠 IgG (Rockland) 作二抗，避光条件下孵育，通过 Odyssey 红外荧光扫描成像系统分析结果。

[0168] 结果显示该抗血清识别的单一蛋白条带约 104kD（见图 9），与理论预测的 SjFz9 蛋白大小一致，说明制备的多抗血清能够特异性地识别天然虫体中的 SjFz9 蛋白，而检测 PBS 免疫血清的对照结果为阴性，表明该多抗可用于 SjFz9 蛋白的免疫组化定位分析。在图 9 中，a 为阳性血清检测，b 为 PBS 免疫血清检测对照。

[0169] 3. SjFz9 蛋白的虫体定位分析

[0170] 选择性收集 13d、18d 童虫及 23d、32d、42d 雌、雄虫新鲜虫体，制备石蜡切片。通过小鼠 S-P 免疫组化检测试剂盒对 SjFz9 进行组织定位分析。

[0171] 具体步骤如下：

[0172] 经二甲苯脱蜡，2 \times 15min。由高到低的梯度浓度酒精进行水化：100% \rightarrow 95% \rightarrow 90% \rightarrow 85% \rightarrow 75% \rightarrow 50%，每个梯度放置 5min，然后放入蒸馏水中 20 ~ 30min，还可在水中放置过夜。3% 过氧化氢中放置 20min，去除内源性过氧化物酶。PBS 洗涤 3 次，3min/ 次。滴加经 37 $^{\circ}$ C 预热的 (0.25%) 胰酶于切片上约 3min 进行抗原修复。PBS 洗涤 3 次，3min/ 次。以正常山羊血清（工作浓度）封闭滴于切片上，湿盒内 37 $^{\circ}$ C 孵育 1h。甩掉封闭血清，滴加正常的小鼠血清及 SjFz9 的小鼠多抗血清的稀释液 (1 : 200) 各 50 ~ 80 μ L，37 $^{\circ}$ C 孵育 1h。PBS 洗涤 4 次，5min/ 次。滴加生物素化山羊抗鼠 IgG。50 ~ 80 μ L，37 $^{\circ}$ C 孵育 30min。PBS 洗涤同上。加 HRP 标记的链霉卵白素工作液 50 ~ 80 μ L，37 $^{\circ}$ C 孵育 20 ~ 30min，PBS 洗涤 4 次，加 DAB (DAKO) 显色液，同时光镜下观察显色情况。一旦着色即用蒸馏水洗涤。苏木素复染 2min，自来水冲洗 10min。脱水，透明，封片，镜检。

[0173] 结果：

[0174] 如图 10 所示，图 10 为免疫组化分析 SjFz9 蛋白在日本血吸虫童虫、成虫体壁肌肉层细胞的定位图，其中，A, B, C, D, E 分别为 13d、18d、23d、32d、42d 虫体定位结果，F 为阴性

对照,箭头所指为阳性染色;at 代表吸盘,ap 代表前部,sm 代表体壁肌细胞,p 代表实质细胞,gc 代表抱雌沟。图 10 所示的童虫虫体定位显示,SjFz9 蛋白在虫体的分布比较广,主要在虫体的前部、口腹吸盘、体壁肌细胞层(见图 10A,B),在实质细胞中也有少量表达,以正常阴性血清检测的对照没有任何染色(图 10F)。而在成虫中的定位显示,SjFz9 蛋白在 23d、32d 和 42d 虫体的吸盘、体壁肌肉层细胞上仍然有大量分布(图 10C,D,E)。另外,更有意思的是,SjFz9 蛋白在雌虫的卵细胞、卵黄腺细胞和雄虫的精子细胞内或细胞膜上都有大量表达(见图 11A,B,C),相应期别的虫体对照组没有表达信号(图 11D,E,F)。图 11 是免疫组化分析 SjFz9 蛋白在日本血吸虫成虫生殖细胞的定位图,是以 SjFz9 蛋白的多抗血清孵育虫体切片,以正常小鼠血清做阴性对照,以 DAB 为显色剂作的免疫组化实验结果图。图 11 中,A,D 为雄虫的睾丸定位;B,E 为雌虫的卵巢定位;C,F 为雌虫的卵黄腺定位;A-C 显示阳性结果,D-F 为阴性对照;箭头所指为 SjFz9 蛋白的阳性染色;o 代表卵巢;t 代表睾丸;v 代表卵黄腺细胞;ic 代表肠腔。

[0175] 以上所述实施例仅表达了本发明的实施方式,其描述较为具体和详细,但并不能因此而理解为对本发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干变形和改进,这些都属于本发明的保护范围。因此,本发明的保护范围应以所附权利要求为准。

[0176] 序列表

[0177] <110> 中国农业科学院上海兽医研究所

[0178] <120> 日本血吸虫 Frizzled9 基因、蛋白及用途

[0179] <160>20

[0180] <170>PatentIn version 3.3

[0181] <210>1

[0182] <211>923

[0183] <212>PRT

[0184] <213>Schistosoma japonicum

[0185] <220>

[0186] <221>SIGNAL

[0187] <222>(1)..(31)

[0188] <223> 信号肽

[0189] <220>

[0190] <221>DOMAIN

[0191] <222>(126)..(244)

[0192] <223> 半胱氨酸富集区

[0193] <220>

[0194] <221>TRANSMEM

[0195] <222>(264)..(748)

[0196] <223> 跨膜区

[0197] <400>1

[0198] Met Lys Ile Val Arg Met Ser Leu Phe Lys Lys Phe Asn Leu Val Asn

[0199]	1	5	10	15
[0200]	Phe Ile Asn Ile Ile Thr Val Met Leu Thr Gln Cys Leu Ser Asp Thr			
[0201]		20	25	30
[0202]	Thr Cys Ala Val Tyr Lys Ile Arg Ser Ser Asn Phe His Phe Tyr Asp			
[0203]		35	40	45
[0204]	Gln Ile Asp Lys Asp Val Asn Phe Ser Gly Lys Tyr Leu Pro Arg Leu			
[0205]		50	55	60
[0206]	Arg Thr Leu Arg Asp Val Ile Pro Leu Met Phe Pro Ile Asn Ser Lys			
[0207]		65	70	75
[0208]	Glu Ile Glu Val Pro Asp His Leu Lys Asn Leu Ile Glu Val Asn Gln			
[0209]		85	90	95
[0210]	Asn Gly Asn Leu Pro Asn Gln Val Tyr Pro Asp Glu Ser Pro Gly Glu			
[0211]		100	105	110
[0212]	Gln Gln Gly Glu Ile Glu Asn Gly Ile Asp Asp Val Asn His Cys Ile			
[0213]		115	120	125
[0214]	Pro Ile His Glu Pro Ile Cys His Gly Leu Gln Tyr Thr His Thr Trp			
[0215]		130	135	140
[0216]	Leu Pro Asn Ser Val Gly Leu Thr Thr Gln Glu Glu Ala Ala Lys Arg			
[0217]		145	150	155
[0218]	Met Asp Asp Tyr Arg Ser Leu Ile Asn Val Gly Cys Ser His Tyr Leu			
[0219]		165	170	175
[0220]	Lys Phe Phe Leu Cys Thr Val Tyr Phe Pro Met Cys Thr Pro Thr Leu			
[0221]		180	185	190
[0222]	Asn Pro Pro Ala Ala Leu Gln Pro Cys Gln Asn Leu Cys Arg Tyr Val			
[0223]		195	200	205
[0224]	Gln Ser Lys Cys Glu Pro Ile Met Lys Ser Phe Ser Phe Pro Trp Pro			
[0225]		210	215	220
[0226]	Ile Glu Leu Asn Cys Lys Ala Leu Pro Ser Asp Ser Gly Met Cys Ile			
[0227]		225	230	235
[0228]	Gln Pro Gln Asn Tyr Asp Leu Asp Lys Gln Gly Gln Ile Ile Ser Pro			
[0229]		245	250	255
[0230]	Gly Ile Pro Ser Glu Ala His Gly Ala Ile Ala Asn Leu Leu Pro Asn			
[0231]		260	265	270
[0232]	Phe Ser Glu Phe Leu Ala Gln Gln Ser Ala Gln Asn Gln Leu Leu Ala			
[0233]		275	280	285
[0234]	Ser Thr Gln Asn Asn Gln Thr Asp Asn Thr Asn Glu Lys Ile Leu Ser			
[0235]		290	295	300
[0236]	Val Asn Asp Asn Gln Ser Gly Leu Thr Arg His His Leu Ile His Arg			
[0237]		305	310	315
				320

[0238]	Thr Lys Leu Gln Ser Ser Lys Cys Tyr Val Val Glu Ile Leu Leu Tyr
[0239]	325 330 335
[0240]	Ser Glu Gln Ala His Asn Asn Ile Thr Cys Ala Arg Arg Cys Asn Ala
[0241]	340 345 350
[0242]	His Leu Phe Tyr Lys Pro Val Glu Lys Arg Phe Ala Asn Ile Trp Met
[0243]	355 360 365
[0244]	Leu Val Trp Ser Ile Ile Cys Leu Gly Ser Cys Ala Leu Thr Ile Ile
[0245]	370 375 380
[0246]	Thr Phe Thr Leu Asn Arg Thr Arg Phe Ala Tyr Pro Glu Arg Pro Ile
[0247]	385 390 395 400
[0248]	Ile Tyr Ile Ser Val Cys Cys Phe Val Tyr Ser Ala Gly Phe Ile Leu
[0249]	405 410 415
[0250]	His Ser Leu Ile Gly Arg Asp Leu Val Ala Cys Arg Asn Lys Val Glu
[0251]	420 425 430
[0252]	Thr Ile Ser Pro Ser Leu Leu Thr Asn Gly Val Ser Gln Val Gly Thr
[0253]	435 440 445
[0254]	Lys Lys Leu Phe Thr Gly Leu Thr Gln Leu Lys Val Asn Ala Ser Phe
[0255]	450 455 460
[0256]	Leu Ile Thr Ser Gly Tyr Glu Gly Thr Trp Cys Thr Ile Ile Phe Ile
[0257]	465 470 475 480
[0258]	Ile Leu Phe Tyr Phe Ser Gln Ala Ser Tyr Leu Trp Trp Val Met Leu
[0259]	485 490 495
[0260]	Ala Thr Ser Trp Phe Leu Ser Ala Ser Cys Lys Trp Gly Cys Glu Gly
[0261]	500 505 510
[0262]	Ile Glu Ala Val Ser Ser Ile Phe His Met Leu Ala Trp Ala Leu Pro
[0263]	515 520 525
[0264]	Ala Leu Lys Thr Ile Ile Ile Leu Ile Met His Arg Ile Asp Ala Asp
[0265]	530 535 540
[0266]	Glu Leu Thr Gly Leu Cys Asn Val Gly Tyr Gln Asn Ser Thr Ser Leu
[0267]	545 550 555 560
[0268]	Leu Thr Leu Ile Leu Leu Pro Gln Ile Ile Tyr Leu Leu Leu Gly Ile
[0269]	565 570 575
[0270]	Met Phe Leu Leu Val Gly Phe His Ser Leu Ile Ser Leu Arg Ser Asn
[0271]	580 585 590
[0272]	Leu Lys Gln His Ala Lys Ser Leu Leu Pro Ala Thr Lys Thr Thr Thr
[0273]	595 600 605
[0274]	Ala Pro Ala Asn Cys Tyr Asn Ser Gln Cys Lys Arg Asn Ala Ser Ala
[0275]	610 615 620
[0276]	Thr Val Ala Ile Pro Thr Met Thr Leu Gln Thr Met Thr Thr Ser Ser

[0277]	625	630	635	640
[0278]	Pro Val Ile Ala Leu Asn Asn Gly Asn Ile Arg Arg Leu Asp Lys Leu			
[0279]		645	650	655
[0280]	Met Ile Lys Ile Gly Ile Phe Ser Met Leu Tyr Ile Ile Pro Cys Ala			
[0281]		660	665	670
[0282]	Cys Val Ile Gly Val Asn Ile Tyr Asn Tyr Ile Asn Tyr Pro Glu Trp			
[0283]		675	680	685
[0284]	Met Lys Thr Leu Asp Arg Leu Ser Lys Gln Thr Asn Cys Leu Thr Glu			
[0285]		690	695	700
[0286]	Tyr Gly Thr Ser Trp Ser Leu Val Thr Gln Cys Leu Glu Asp Asn His			
[0287]	705	710	715	720
[0288]	Phe Pro Ser Ala Glu Ala Asn Met Leu Gln Ile Phe Met Ser Leu Val			
[0289]		725	730	735
[0290]	Val Gly Ile Thr Ser Gly Met Trp Val Trp Cys Asn Arg Lys Ser Ser			
[0291]		740	745	750
[0292]	Glu Thr Trp Leu Lys Cys Ile Lys Arg Arg Lys Lys Asn Leu Ser Gln			
[0293]		755	760	765
[0294]	His Ser Leu Asn Tyr Arg Tyr His Ile Cys Gly Asp Val Val Ala Gly			
[0295]		770	775	780
[0296]	Asp Thr Val Gly Glu Thr Asn Asn Val Ser Gln Val Asp Gln Lys Ala			
[0297]	785	790	795	800
[0298]	Gln Leu Met Asn Thr Thr Cys Cys Asn Ser Arg Ala Ala Cys Gly Lys			
[0299]		805	810	815
[0300]	Asn Asn Gln Tyr Pro Gly Asn Thr Met Asn Phe Asp Asn Thr Met Asn			
[0301]		820	825	830
[0302]	Asn Ile Asn Gln Leu Val Ala Asp Ser Thr Gln Tyr His Gln Gln Gln			
[0303]		835	840	845
[0304]	His Gln Asn Glu Ile Ala Arg Asn Asn Val Arg Asn Leu Met Leu Ser			
[0305]		850	855	860
[0306]	Gly Ser Phe Asn Met Ser Ser Gly Thr Ala Ser Cys Thr Ser Ala Pro			
[0307]	865	870	875	880
[0308]	Leu Leu Ser His Gly Leu Gly Thr Ile Asn Met Ser Leu Leu Asn Ser			
[0309]		885	890	895
[0310]	Phe Glu Ser Asn Tyr Cys Thr Arg Ser Ser Gln Arg Gln Thr Phe Lys			
[0311]		900	905	910
[0312]	Thr Arg Pro Ser Asn Thr Gly Arg Pro Asp Cys			
[0313]		915	920	
[0314]	<210>2			
[0315]	<211>3670			

[0316] <212>DNA
 [0317] <213>Schistosoma japonicum
 [0318] <220>
 [0319] <221>CDS
 [0320] <222>(146).. (2917)
 [0321] <400>2
 [0322] ggtcaaataa caatttaaca ttatgaaaga tatatnttca agtgtatatt tatgaatgta
 60
 [0323] cttacttgta aaatgaagtc atcttaacag aactgagctg tttagttgct ttgttggtt
 120
 [0324] aatttacata aattgtcact cagag atg aaa att gta aga atg agt tta ttt
 172
 [0325] Met Lys Ile Val Arg Met Ser Leu Phe
 [0326] 1 5
 [0327] aaa aaa ttt aat ttg gtc aat ttc ata aat att ata aca gtc atg tta 220
 [0328] Lys Lys Phe Asn Leu Val Asn Phe Ile Asn Ile Ile Thr Val Met Leu
 [0329] 10 15 20 25
 [0330] act caa tgt ctc tca gat act acc tgt gct gtc tat aaa att cga agc 268
 [0331] Thr Gln Cys Leu Ser Asp Thr Thr Cys Ala Val Tyr Lys Ile Arg Ser
 [0332] 30 35 40
 [0333] tca aat ttt cac ttt tat gat caa att gat aaa gat gtt aat ttc agt 316
 [0334] Ser Asn Phe His Phe Tyr Asp Gln Ile Asp Lys Asp Val Asn Phe Ser
 [0335] 45 50 55
 [0336] gga aag tat ttg cct cgc tta cga aca cta aga gat gtt ata cct ctc 364
 [0337] Gly Lys Tyr Leu Pro Arg Leu Arg Thr Leu Arg Asp Val Ile Pro Leu
 [0338] 60 65 70
 [0339] atg ttt cca ata aat tca aaa gaa att gaa gta ccc gat cat ttg aaa 412
 [0340] Met Phe Pro Ile Asn Ser Lys Glu Ile Glu Val Pro Asp His Leu Lys
 [0341] 75 80 85
 [0342] aat cta att gaa gtt aat caa aat gga aac cta ccg aat caa gtt tat 460
 [0343] Asn Leu Ile Glu Val Asn Gln Asn Gly Asn Leu Pro Asn Gln Val Tyr
 [0344] 90 95 100 105
 [0345] cca gat gag agt cca ggt gaa caa caa gga gaa att gaa aat ggt ata 508
 [0346] Pro Asp Glu Ser Pro Gly Glu Gln Gln Gly Glu Ile Glu Asn Gly Ile
 [0347] 110 115 120
 [0348] gat gat gta aat cac tgt att cca att cat gaa cca att tgt cat ggt 556
 [0349] Asp Asp Val Asn His Cys Ile Pro Ile His Glu Pro Ile Cys His Gly
 [0350] 125 130 135
 [0351] ctt caa tat act cat act tgg tta cct aat tca gtg ggt cta acg aca 604

[0352]	Leu Gln Tyr Thr His Thr Trp Leu Pro Asn Ser Val Gly Leu Thr Thr	
[0353]	140	145 150
[0354]	caa gaa gag gct gca aaa cgc atg aat gac tat cgt tcc ctc atc aat	652
[0355]	Gln Glu Glu Ala Ala Lys Arg Met Asn Asp Tyr Arg Ser Leu Ile Asn	
[0356]	155	160 165
[0357]	gtt ggt tgt tca cac tat tta aaa ttt ttt ctg tgt aca gtt tat ttt	700
[0358]	Val Gly Cys Ser His Tyr Leu Lys Phe Phe Leu Cys Thr Val Tyr Phe	
[0359]	170	175 180 185
[0360]	cca atg tgt acc cca aca ctg aat cca cca gct gcc tta cag cca tgt	748
[0361]	Pro Met Cys Thr Pro Thr Leu Asn Pro Pro Ala Ala Leu Gln Pro Cys	
[0362]	190	195 200
[0363]	caa aat tta tgt cgt tat gta caa tca aag tgt gaa cca att atg aaa	796
[0364]	Gln Asn Leu Cys Arg Tyr Val Gln Ser Lys Cys Glu Pro Ile Met Lys	
[0365]	205	210 215
[0366]	agt ttt agc ttt cca tgg cca atc gaa ctt aat tgt aaa gca tta ccc	844
[0367]	Ser Phe Ser Phe Pro Trp Pro Ile Glu Leu Asn Cys Lys Ala Leu Pro	
[0368]	220	225 230
[0369]	tca gat tct gga atg tgt ata caa cca caa aat tat gat tta gat aaa	892
[0370]	Ser Asp Ser Gly Met Cys Ile Gln Pro Gln Asn Tyr Asp Leu Asp Lys	
[0371]	235	240 245
[0372]	caa gga caa att ata agc cca ggg atc cca tct gaa gcc cat gga gct	940
[0373]	Gln Gly Gln Ile Ile Ser Pro Gly Ile Pro Ser Glu Ala His Gly Ala	
[0374]	250	255 260 265
[0375]	ata gct aat tta ttg cca aat ttt agt gag ttt ttg gcc caa caa agt	988
[0376]	Ile Ala Asn Leu Leu Pro Asn Phe Ser Glu Phe Leu Ala Gln Gln Ser	
[0377]	270	275 280
[0378]	gct caa aac caa ctt ttg gcg agt aca cag aac aat caa aca gat aat	1036
[0379]	Ala Gln Asn Gln Leu Leu Ala Ser Thr Gln Asn Asn Gln Thr Asp Asn	
[0380]	285	290 295
[0381]	act aac gag aaa ata ctt tea gtg aat gat aac caa tcc ggt ttg aca	1084
[0382]	Thr Asn Glu Lys Ile Leu Ser Val Asn Asp Asn Gln Ser Gly Leu Thr	
[0383]	300	305 310
[0384]	aga cat cat cta atc cac agg aca aaa cta caa tct tca aag tgc tat	1132
[0385]	Arg His His Leu Ile His Arg Thr Lys Leu Gln Ser Ser Lys Cys Tyr	
[0386]	315	320 325
[0387]	gtt gtc gaa att ctt tta tat tet gaa cag gca cat aat aat att aca	1180
[0388]	Val Val Glu Ile Leu Leu Tyr Ser Glu Gln Ala His Asn Asn Ile Thr	
[0389]	330	335 340 345
[0390]	tgc gct cgt cgt tgt aat get cat tta ttt tac aaa cca gtt gaa aaa	1228

[0391]	Cys Ala Arg Arg Cys Asn Ala His Leu Phe Tyr Lys Pro Val Glu Lys		
[0392]		350	355 360
[0393]	cgt ttc gcc aat att tgg atg cta gtt tgg tcc ata ata tgt ttg gga		1276
[0394]	Arg Phe Ala Asn Ile Trp Met Leu Val Trp Ser Ile Ile Cys Leu Gly		
[0395]		365	370 375
[0396]	tcg tgt gca ttg act att atc acc ttt acc tta aac cgc aca cgt ttt		1324
[0397]	Ser Cys Ala Leu Thr Ile Ile Thr Phe Thr Leu Asn Arg Thr Arg Phe		
[0398]		380	385 390
[0399]	gct tat cca gaa cga ccg att att tac ata tca gtt tgt tgc ttc gtt		1372
[0400]	Ala Tyr Pro Glu Arg Pro Ile Ile Tyr Ile Ser Val Cys Cys Phe Val		
[0401]		395	400 405
[0402]	tac tct gct gga ttt att tta cac tct ctg att ggg cgt gat ttg gtg		1420
[0403]	Tyr Ser Ala Gly Phe Ile Leu His Ser Leu Ile Gly Arg Asp Leu Val		
[0404]		410	415 420 425
[0405]	gca tgt agg aac aaa gtt gaa aca att tcg cca tca tta ttg aca aac		1468
[0406]	Ala Cys Arg Asn Lys Val Glu Thr Ile Ser Pro Ser Leu Leu Thr Asn		
[0407]		430	435 440
[0408]	ggg gta tca caa gtt ggt aca aag aaa ctt ttt aca ggt tta act caa		1516
[0409]	Gly Val Ser Gln Val Gly Thr Lys Lys Leu Phe Thr Gly Leu Thr Gln		
[0410]		445	450 455
[0411]	tta aaa gtt aac gct agt ttc ttg ata act tca gga tat gaa ggt aca		1564
[0412]	Leu Lys Val Asn Ala Ser Phe Leu Ile Thr Ser Gly Tyr Glu Gly Thr		
[0413]		460	465 470
[0414]	tgg tgt act att ata ttt att atc ctg ttt tat ttc agt caa geg agt		1612
[0415]	Trp Cys Thr Ile Ile Phe Ile Ile Leu Phe Tyr Phe Ser Gln Ala Ser		
[0416]		475	480 485
[0417]	tat tta tgg tgg gtt atg tta get acc tca tgg ttt tta tca get tcc		1660
[0418]	Tyr Leu Trp Trp Val Met Leu Ala Thr Ser Trp Phe Leu Ser Ala Ser		
[0419]		490	495 500 505
[0420]	tgt aag tgg gga tgt gaa ggc ata gag get gtc agt agt atc ttc cat		1708
[0421]	Cys Lys Trp Gly Cys Glu Gly Ile Glu Ala Val Ser Ser Ile Phe His		
[0422]		510	515 520
[0423]	atg ctc gcc tgg gca tta cca gca tta aaa act att atc ata tta atc		1756
[0424]	Met Leu Ala Trp Ala Leu Pro Ala Leu Lys Thr Ile Ile Ile Leu Ile		
[0425]		525	530 535
[0426]	atg cat cgt att gac gca gat gag tta act gga tta tgc aac gtt ggt		1804
[0427]	Met His Arg Ile Asp Ala Asp Glu Leu Thr Gly Leu Cys Asn Val Gly		
[0428]		540	545 550
[0429]	tat caa aat tca acg agt tta ctc aca ttg ata tta cta cca caa atc		1852

[0430] Tyr Gln Asn Ser Thr Ser Leu Leu Thr Leu Ile Leu Leu Pro Gln Ile
 [0431] 555 560 565
 [0432] atc tac cta ttg ttg ggc atc atg ttt tta ttg gtt ggt ttt cac tcc 1900
 [0433] Ile Tyr Leu Leu Leu Gly Ile Met Phe Leu Leu Val Gly Phe His Ser
 [0434] 570 575 580 585
 [0435] ttg att tcc ctt cga tca aac tta aaa caa cac gct aaa agc ctc ttg 1948
 [0436] Leu Ile Ser Leu Arg Ser Asn Leu Lys Gln His Ala Lys Ser Leu Leu
 [0437] 590 595 600
 [0438] cct gct acc aaa aca aca aca gca cca gca aac tgt tat aat tct cag 1996
 [0439] Pro Ala Thr Lys Thr Thr Thr Ala Pro Ala Asn Cys Tyr Asn Ser Gln
 [0440] 605 610 615
 [0441] tgc aaa aga aat gct tca gca aca gta gca ata cca act atg aca cta 2044
 [0442] Cys Lys Arg Asn Ala Ser Ala Thr Val Ala Ile Pro Thr Met Thr Leu
 [0443] 620 625 630
 [0444] caa aca atg acc aca tct agt cca gtc atc gcc ctc aac aat gga aat 2092
 [0445] Gln Thr Met Thr Thr Ser Ser Pro Val Ile Ala Leu Asn Asn Gly Asn
 [0446] 635 640 645
 [0447] ata cga cgt tta gat aaa cta atg ata aaa ata ggc ata ttt tca atg 2140
 [0448] Ile Arg Arg Leu Asp Lys Leu Met Ile Lys Ile Gly Ile Phe Ser Met
 [0449] 650 655 660 665
 [0450] ctg tat att att cca tgt gca tgt gtt att ggt gtg aat att tac aat 2188
 [0451] Leu Tyr Ile Ile Pro Cys Ala Cys Val Ile Gly Val Asn Ile Tyr Asn
 [0452] 670 675 680
 [0453] tat att aat tat ccg gaa tgg atg aaa aca ttg gat aga cta tct aaa 2236
 [0454] Tyr Ile Asn Tyr Pro Glu Trp Met Lys Thr Leu Asp Arg Leu Ser Lys
 [0455] 685 690 695
 [0456] caa aca aac tgt tta acg gaa tac ggt aca agt tgg tca ctg gtc act 2284
 [0457] Gln Thr Asn Cys Leu Thr Glu Tyr Gly Thr Ser Trp Ser Leu Val Thr
 [0458] 700 705 710
 [0459] cag tgt cta gaa gat aat cac ttc cca tca gct gaa gct aac atg tta 2332
 [0460] Gln Cys Leu Glu Asp Asn His Phe Pro Ser Ala Glu Ala Asn Met Leu
 [0461] 715 720 725
 [0462] caa ata ttc atg tct cta gtc gtc ggt atc act tca gga atg tgg gta 2380
 [0463] Gln Ile Phe Met Ser Leu Val Val Gly Ile Thr Ser Gly Met Trp Val
 [0464] 730 735 740 745
 [0465] tgg tgt aat aga aaa tct tea gaa aca tgg tta aaa tgc ata aag agg 2428
 [0466] Trp Cys Asn Arg Lys Ser Ser Glu Thr Trp Leu Lys Cys Ile Lys Arg
 [0467] 750 755 760
 [0468] aga aag aaa aat tta tct cag cat agt ttg aat tac cgt tat cat att 2476

[0469]	Arg Lys Lys Asn Leu Ser Gln His Ser Leu Asn Tyr Arg Tyr His Ile	
[0470]	765	770 775
[0471]	tgt ggg gat gtt gtc gct ggt gat act gtt gga gaa act aat aat gtg	2524
[0472]	Cys Gly Asp Val Val Ala Gly Asp Thr Val Gly Glu Thr Asn Asn Val	
[0473]	780	785 790
[0474]	tcc caa gtt gat caa aag gca caa cta atg aat acg act tgt tgt aat	2572
[0475]	Ser Gln Val Asp Gln Lys Ala Gln Leu Met Asn Thr Thr Cys Cys Asn	
[0476]	795	800 805
[0477]	tca aga gcg gca tgt gga aaa aac aac cag tat cca gga aac acc atg	2620
[0478]	Ser Arg Ala Ala Cys Gly Lys Asn Asn Gln Tyr Pro Gly Asn Thr Met	
[0479]	810	815 820 825
[0480]	aat ttc gat aat aca atg aac aat att aac cag ctg gtg gct gat tcc	2668
[0481]	Asn Phe Asp Asn Thr Met Asn Asn Ile Asn Gln Leu Val Ala Asp Ser	
[0482]	830	835 840
[0483]	act cag tat cat caa caa caa cac caa aat gaa atc gca cga aac aac	2716
[0484]	Thr Gln Tyr His Gln Gln Gln His Gln Asn Glu Ile Ala Arg Asn Asn	
[0485]	845	850 855
[0486]	gta aga aat tta atg ctt tet gga tca ttt aac atg tca tet gga aca	2764
[0487]	Val Arg Asn Leu Met Leu Ser Gly Ser Phe Asn Met Ser Ser Gly Thr	
[0488]	860	865 870
[0489]	gct tca tgt aca agt gct cca cta tta tca cat ggt tta gga aca ata	2812
[0490]	Ala Ser Cys Thr Ser Ala Pro Leu Leu Ser His Gly Leu Gly Thr Ile	
[0491]	875	880 885
[0492]	aat atg tca tta tta aat tea ttt gaa tcc aat tat tgc acc cgt tea	2860
[0493]	Asn Met Ser Leu Leu Asn Ser Phe Glu Ser Asn Tyr Cys Thr Arg Ser	
[0494]	890	895 900 905
[0495]	tca caa cga caa aca ttt aaa act aga cca tet aat acg gga aga cca	2908
[0496]	Ser Gln Arg Gln Thr Phe Lys Thr Arg Pro Ser Asn Thr Gly Arg Pro	
[0497]	910	915 920
[0498]	gac tgt taa tttgtttgct tccttggtaa ataatgtaaa tttttcttga	2957
[0499]	Asp Cys	
[0500]	atacacttac ceatgattac aaatttattt gtaacagctg atatttactc tacattgcat	3017
[0501]	aagccattgc acaaacaaaa ccttgagcga agatcaata tgaaaaataa tcataaatct	3077
[0502]	tgttcatctc attttaattt tgtatcaatt gatcgttcat attacattct gtaatcatt	3137
[0503]	tctaaatgat agttctttcg ttctgcttgt ttcatcagc atttatcaat ttgaacaagt	3197
[0504]	gattacgect atccgcacag aatgctcatt tetgtatata aattatatttg tgacttgat	3257
[0505]	tccggttate tttaacttg aaaataattt catttattct ttatttttta actacaaate	3317
[0506]	aagaagctat tgettgtcgc ttcacctatt atgetttaat ttgatatttt ttacgattaa	3377
[0507]	ttcaattcag actttgttca aacgttttct ttaaacttca cactttcaga aaaagaaaa	3437

[0508]	caaaccatac tagaacagtg gtacgggtta gtaatgagga ataaatgaac aataaataac	3497
[0509]	tcaatttgat cacaaagctc attaattcca ttttcttgag tcaccttgta tatgagtaac	3557
[0510]	tgtagacttg tatttccttc agttttcatt tgtaattta ttcaccttg ccaataattt	3617
[0511]	aaagacacta attatagcct tgaaaaaaaa aaaaaaaaa aaaaaaaaa aaa	3670
[0512]	<210>3	
[0513]	<211>26	
[0514]	<212>DNA	
[0515]	<213> 人工序列	
[0516]	<220>	
[0517]	<221>misc_feature	
[0518]	<222>(1).. (26)	
[0519]	<223> 引物	
[0520]	<220>	
[0521]	<221>misc_feature	
[0522]	<222>(1).. (26)	
[0523]	<223>N = A/C/G/T ;R = A/G ;M = A/C ;Y = C/T	
[0524]	<400>3	
[0525]	nnngaattct aycengarmg nccnat	26
[0526]	<210>4	
[0527]	<211>24	
[0528]	<212>DNA	
[0529]	<213> 人工序列	
[0530]	<220>	
[0531]	<221>misc_feature	
[0532]	<222>(1).. (24)	
[0533]	<223> 引物	
[0534]	<220>	
[0535]	<221>misc_feature	
[0536]	<222>(1).. (24)	
[0537]	<223>N = A/C/G/T ;R = A/G ;Y = C/T	
[0538]	<400>4	
[0539]	nnnaagcttn arnaynaccc acca	24
[0540]	<210>5	
[0541]	<211>26	
[0542]	<212>DNA	
[0543]	<213> 人工序列	
[0544]	<220>	
[0545]	<221>misc_feature	
[0546]	<222>(1).. (26)	

- [0547] <223> 引物
[0548] <220>
[0549] <221>misc_feature
[0550] <222>(1)..(26)
[0551] <223>N = A/C/G/T ;R = A/G ;M = A/C ;Y = C/T
[0552] <400>5
[0553] nnngaattct ayccngarmg nccnat 26
[0554] <210>6
[0555] <211>21
[0556] <212>DNA
[0557] <213> 人工序列
[0558] <220>
[0559] <221>misc_feature
[0560] <222>(1)..(21)
[0561] <223> 引物
[0562] <220>
[0563] <221>misc_feature
[0564] <222>(1)..(21)
[0565] <223>N = A/C/G/T ;D = A/G/T
[0566] <400>6
[0567] nnnaagcttn gtdtancnt c 21
[0568] <210>7
[0569] <211>25
[0570] <212>DNA
[0571] <213> 人工序列
[0572] <220>
[0573] <221>misc_feature
[0574] <222>(1)..(25)
[0575] <223> 引物
[0576] <400>7
[0577] ctatccggag cgaccatta tttac 25
[0578] <210>8
[0579] <211>25
[0580] <212>DNA
[0581] <213> 人工序列
[0582] <220>
[0583] <221>misc_feature
[0584] <222>(1)..(25)
[0585] <223> 引物

[0586] <400>8
[0587] acactctctg attgggcgtg atttg 25
[0588] <210>9
[0589] <211>25
[0590] <212>DNA
[0591] <213> 人工序列
[0592] <220>
[0593] <221>misc_feature
[0594] <222>(1).. (25)
[0595] <223> 引物
[0596] <400>9
[0597] tgaccacca taaataactc gcttg 25
[0598] <210>10
[0599] <211>25
[0600] <212>DNA
[0601] <213> 人工序列
[0602] <220>
[0603] <221>misc_feature
[0604] <222>(1).. (25)
[0605] <223> 引物
[0606] <400>10
[0607] caccatgtac cttcatatcc tgaag 25
[0608] <210>11
[0609] <211>25
[0610] <212>DNA
[0611] <213> 人工序列
[0612] <220>
[0613] <221>misc_feature
[0614] <222>(1).. (25)
[0615] <223> 引物
[0616] <400>11
[0617] taacggaata cggtaacaagt tggtc 25
[0618] <210>12
[0619] <211>25
[0620] <212>DNA
[0621] <213> 人工序列
[0622] <220>
[0623] <221>misc_feature
[0624] <222>(1).. (25)

[0625] <223> 引物
[0626] <400>12
[0627] tatcacttca ggaatgtggg tatgg 25
[0628] <210>13
[0629] <211>25
[0630] <212>DNA
[0631] <213> 人工序列
[0632] <220>
[0633] <221>misc_feature
[0634] <222>(1)..(25)
[0635] <223> 引物
[0636] <400>13
[0637] taagttcgat tggccatgga aagct 25
[0638] <210>14
[0639] <211>25
[0640] <212>DNA
[0641] <213> 人工序列
[0642] <220>
[0643] <221>misc_feature
[0644] <222>(1)..(25)
[0645] <223> 引物
[0646] <400>14
[0647] accaacattg atgagggaac gatag 25
[0648] <210>15
[0649] <211>25
[0650] <212>DNA
[0651] <213> 人工序列
[0652] <220>
[0653] <221>misc_feature
[0654] <222>(1)..(25)
[0655] <223> 引物
[0656] <400>15
[0657] tatcacttca ggaatgtggg tatgg 25
[0658] <210>16
[0659] <211>25
[0660] <212>DNA
[0661] <213> 人工序列
[0662] <220>
[0663] <221>misc_feature

- [0664] <222>(1).. (25)
[0665] <223> 引物
[0666] <400>16
[0667] ttagtttctc caacagtatc accag 25
[0668] <210>17
[0669] <211>25
[0670] <212>DNA
[0671] <213> 人工序列
[0672] <220>
[0673] <221>misc_feature
[0674] <222>(1).. (25)
[0675] <223> 引物
[0676] <400>17
[0677] ttagttggtg gagcgatttg tctgg 25
[0678] <210>18
[0679] <211>25
[0680] <212>DNA
[0681] <213> 人工序列
[0682] <220>
[0683] <221>misc_feature
[0684] <222>(1).. (25)
[0685] <223> 引物
[0686] <400>18
[0687] ctgttattgc tcaatttcgt gcgac 25
[0688] <210>19
[0689] <211>33
[0690] <212>DNA
[0691] <213> 人工序列
[0692] <220>
[0693] <221>misc_feature
[0694] <222>(1).. (33)
[0695] <223> 引物
[0696] <220>
[0697] <221>misc_signal
[0698] <222>(3).. (8)
[0699] <223>BamH I 酶切位点
[0700] <400>19
[0701] gtggatccac agtcatgtta actcaatgtc tct 33
[0702] <210>20

- [0703] <211>33
[0704] <212>DNA
[0705] <213>人工序列
[0706] <220>
[0707] <221>misc_feature
[0708] <222>(1)..(33)
[0709] <223>引物
[0710] <220>
[0711] <221>misc_signal
[0712] <222>(3)..(8)
[0713] <223>Xho I 酶切位点
[0714] <400>20
[0715] gtctcgagtt cacactttga ttgtacataa cga 33

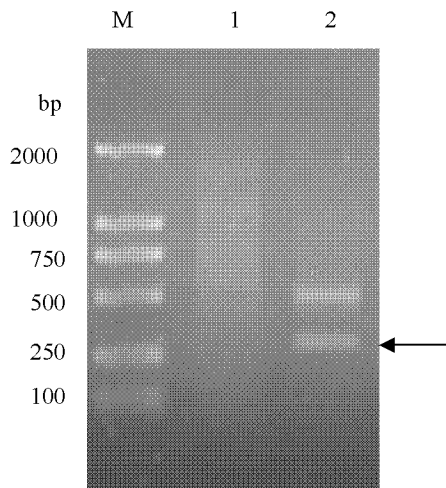


图 1

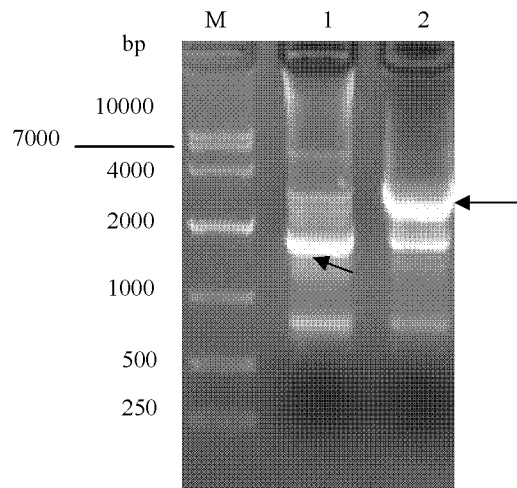


图 2

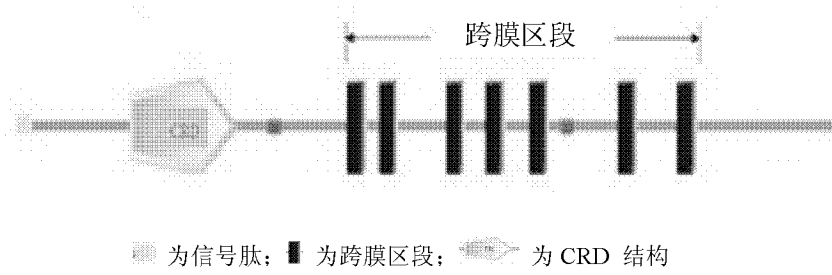


图 3

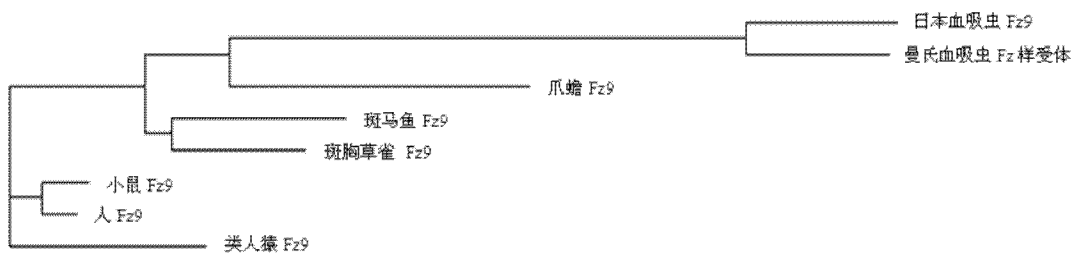


图 4

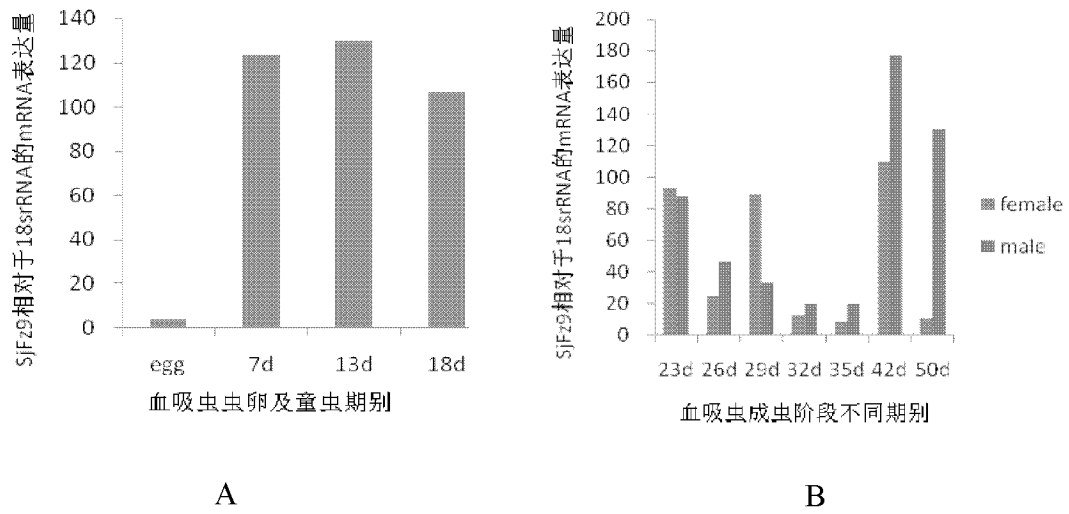


图 5

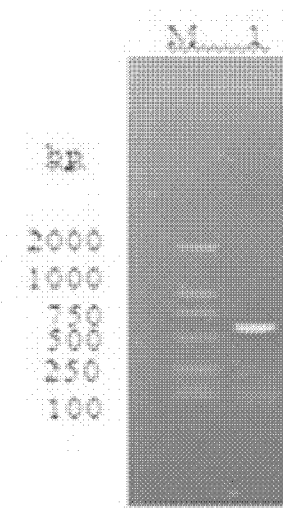


图 6

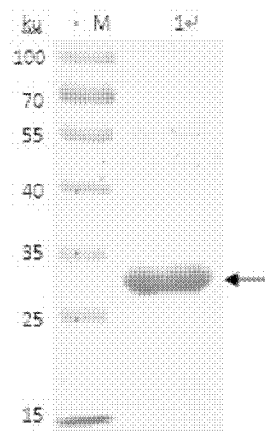


图 7

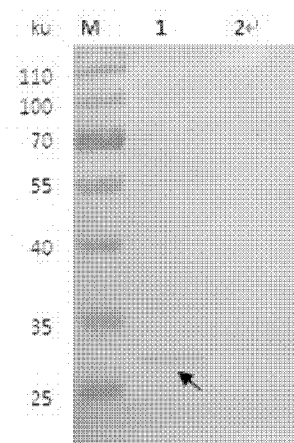


图 8

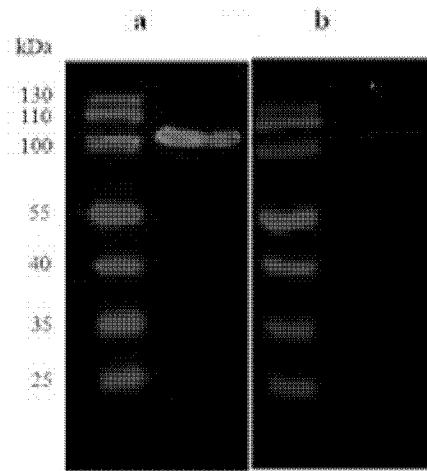


图 9

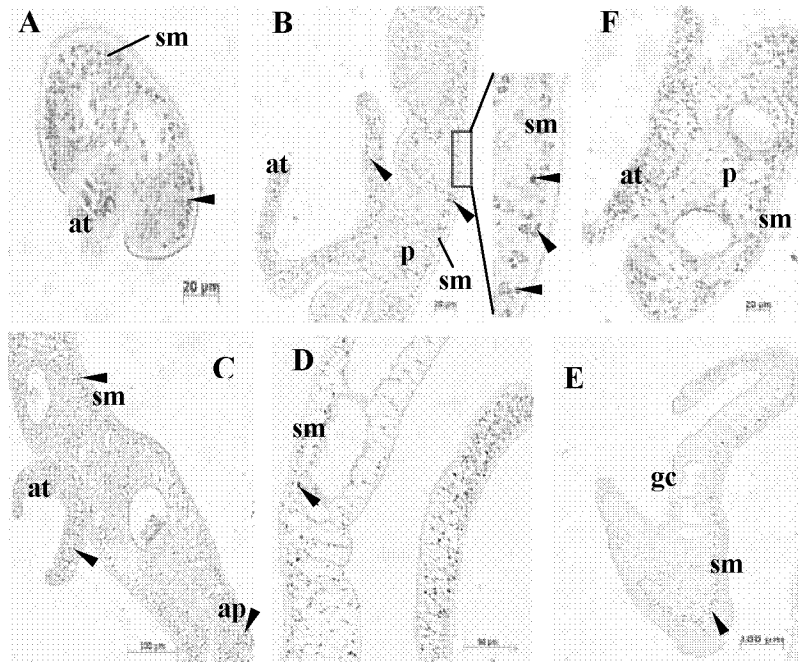


图 10

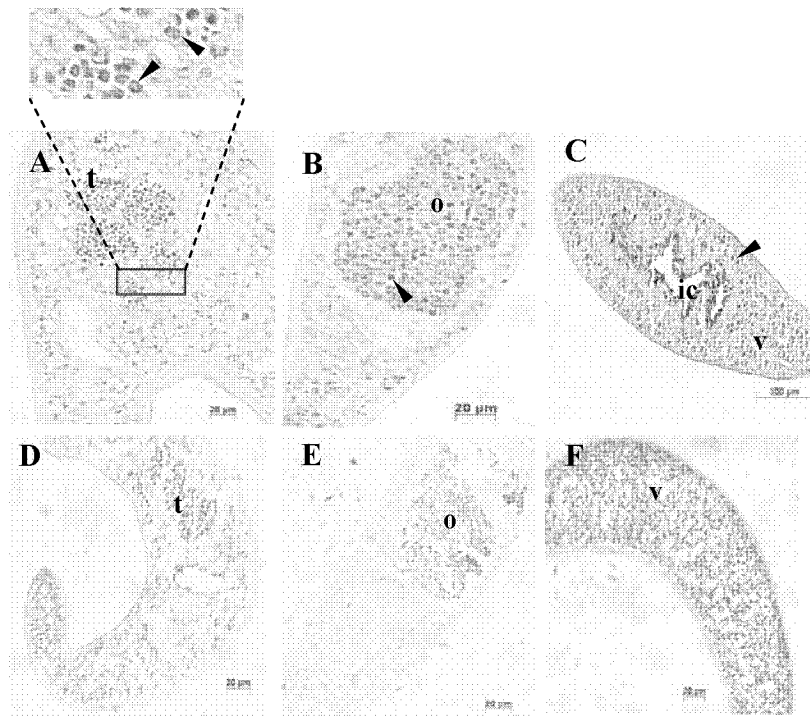


图 11

专利名称(译)	日本血吸虫Frizzled9基因、蛋白及用途		
公开(公告)号	CN101892239B	公开(公告)日	2012-05-23
申请号	CN201010224645.5	申请日	2010-07-15
[标]申请(专利权)人(译)	中国农业科学院上海兽医研究所		
申请(专利权)人(译)	中国农业科学院上海兽医研究所		
当前申请(专利权)人(译)	中国农业科学院上海兽医研究所		
[标]发明人	王孝波 苑纯秀 李红飞 冯新港 杨健美		
发明人	王孝波 苑纯秀 李红飞 冯新港 杨健美		
IPC分类号	C12N15/12 C07K14/705 C07K16/28 C12N15/63 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12Q1/68 G01N33/53 A61K48/00 A61K38/17 A61P33/12		
CPC分类号	Y02A50/423		
其他公开文献	CN101892239A		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种日本血吸虫基因，所述基因具有编码Frizzled(Fz)家族蛋白的核苷酸序列，该核苷酸序列包含：编码SEQ ID NO.1氨基酸序列的DNA序列，或编码SEQ ID NO.1所示氨基酸序列中缺失、添加、插入或取代一个或多个氨基酸所得到的具有SEQ ID NO.1相同活性的氨基酸序列的DNA序列。本发明还公开了上述基因编码的蛋白质。本发明的日本血吸虫基因和蛋白，是日本血吸虫Fz受体家族的新成员，参与调控血吸虫的生长发育及生殖系统发育，对开发阻断血吸虫生长发育及生殖系统发育的新药具有很高的应用价值。

