

绵羊HTR4基因的生物信息学分析

刘 佳, 王 祯, 代友超, 张小雪

(甘肃农业大学动物科学技术学院, 甘肃 兰州 730070)

摘要: 利用生物信息学数据库及软件, 对绵羊羟色胺受体4基因(hydroxytryptamine receptor 4, *HTR4*)进行生物信息学分析, 以初步了解其结构并进行功能预测。结果表明, 绵羊 *HTR4* 基因所含最大长度的开放阅读框为 1 317 bp, 编码 438 个氨基酸残基。*HTR4* 基因编码的蛋白分子质量为 49 437.09 KDa, 理论等电点(pI)为 8.32。亚细胞定位结果表明其编码产物主要定位于质膜 (60.9%), 且不属于分泌蛋白。*HTR4* 蛋白不存在信号肽序列, 存在 7 段跨膜结构且无低复杂性区域, 该蛋白的二级结构以 α 螺旋为主, 且三级结构主要由 α 折叠缠绕形成。

关键词: 绵羊; *HTR4* 基因; 生物信息学分析

中图分类号: S826 **文献标志码:** A **文章编号:** 1001-1463(2020)10-0035-06

doi: 10.3969/j.issn.1001-1463.2020.10.008

Bioinformatics Analysis of Sheep *HTR4* Gene

LIU Jia, WANG Zhen, DAI Youchao, ZHANG Xiaoxue

(College of Animal Science and Technology, Gansu Agriculture University, Lanzhou Gansu 730070, China)

Abstract: The bioinformatics analysis of sheep *HTR4* gene was carried out by using bioinformatics database and software, and its structure was preliminarily understood and predicted. The results showed that sheep *HTR4* gene contained an open reading frame of 1 317 bp, encoding 438 amino acid residues. The molecular mass of *HTR4* protein was 49 437.09 KDa, and the theoretical isoelectric point pI was 8.32. The subcells were mainly located in plasma membrane and did not belong to secretory protein (60.9%). There are no signal peptide sequences. There are seven segments of transmembrane structure and no segments of low complexity region. The secondary structure is mainly α -helix, and the tertiary structure is mainly formed by α -helix, winding and folding.

Key words: Sheep; *HTR4* gene; Bioinformatics analysis

5-羟色胺(5-hydroxytryptamine, 5-HT)既是一种神经递质, 又是一种血管活性物质, 在中枢神经系统和外周组织中起着多种生理作用^[1]。*HTR4* 是其受体一亚型, *HTR4* 基因在人体内定位于 5q31-33, 编码羟色胺受体 4 型蛋白^[2]。该基因在人和鼠的中枢神经系统和外周组织中都有广泛的表达, 但在不同的组织中表达量差异显著, 主要分布在

脑组织和肠道组织中, 在人类中枢神经系统所有组织中都有表达, 其中脑组织的表达率显著高于外周组织^[3]。外周组织中, 小肠的表达率最高, 参与胃肠道的运动与分泌、调节蠕动反射及影响肠道感觉功能^[4]。*HTR4* 参与胃肠道的运动与分泌、调节蠕动反射及影响肠道感觉功能^[5]。外周 *HTR4* 的作用首先是在豚鼠体内发现的。在用电刺激豚鼠肠

收稿日期: 2020-07-03; 修订日期: 2020-08-20

基金项目: 甘肃农业大学学生科研训练计划项目(202004030)。

作者简介: 刘 佳(1999—), 女, 甘肃白银人, 本科在读, 研究方向为动物科学。联系电话: (0)18893476427。Email: 3200825330@qq.com。

通信作者: 张小雪(1984—), 女, 湖北武汉人, 讲师, 主要从事动物遗传育种与繁殖研究及教学工作。联系电话: (0931)7631225。Email: zhangxx@gsau.edu.cn。

纵行平滑肌丛时, HT 作用于一个在药理学上与 *HTR4* 类似的受体, 从而增强胆碱能介导的收缩。其在人和鼠的中枢神经系统和外周组织中都有广泛的表达, 但不同组织内的表达量却有显著差异, 其主要分布在脑组织和胃肠道组织中。

生物信息学是利用数学方法处理和分析生物数据, 基于人们现有对分子生物学的认识和构建生物模型, 进而分析其生物学特性的方法^[6]。对 *HTR4* 基因的研究大多为人类 *HTR4* 基因及其编码产物的功能, 而绵羊 *HTR4* 基因方面的研究不足。本研究以生物基因组数据库调取的绵羊 *HTR4* 的序列为基础, 利用生物信息学方法对绵羊 *HTR4* 基因及其编码产物的理化性质、序列特征、蛋白质结构以及生物学功能进行预测和分析, 以期为深入研究 *HTR4* 基因及其编码蛋白的基本结构和生物学功能提供理论基础。

1 材料与方法

1.1 序列来源

数据来源于 NCBI 网站 GenBank 数据库, 包括绵羊 (XM_027970630.1)、狗 (XM_022417831.1)、鸡 (XM_015293658.2)、马 (NM_001163982.1)、猫 (XM_023260627.1)、牛 (NM_001040485.1)、兔 (XM_008255221.2)、豚鼠 (NM_001172963.1)、猪 (NM_001001267.1)、人 (NM_000870.7) 和山羊 (XM_018050181.1) 等 11 个物种的 mRNA 序列。括号内为 GenBank 登录号。

1.2 方法

绵羊 *HTR4* 基因开放阅读框 (Open reading frame, ORF) 采用 NCBI 的 ORF Finder 程序分析, 参照 Kozak 法则^[7]; *HTR4* 基因编

码产物的理化性质采用 Bioedit 及 ExPASy 分析软件预测; 亚细胞定位采用 PSORT II 预测^[8]。蛋白潜在信号肽剪切位点预测采用 Signalp 3.0 软件。跨膜螺旋区域的预测采用 TMHMM 程序。蛋白保守结构域分析采用 Smart 软件。采用 Prot Scale 进行蛋白亲疏水性分析。二级结构采用 Jpred 分析预测。采用 Swiss-model 软件分析蛋白三级结构多序列比对及同源性分析采用 DNAMAN 软件分析^[9]。

2 结果与分析

2.1 绵羊 *HTR4* 基因开放阅读框分析

从开放阅读框分析结果 (图 1) 可以看出, 绵羊 *HTR4* 基因序列中最大长度的 ORF 有 1 317 bp, 起始密码子位于 1 bp 处, 终止密码子位于 1 845 bp 处, 推测其编码 438 个氨基酸残基。

2.2 绵羊 *HTR4* 编码产物的理化性质分析

蛋白质的基本性质包括其相对分子质量、氨基酸组成、等电点基因编码产物半衰期和不稳定指数等^[10]。对绵羊 *HTR4* 基因编码产物理化性质的分析结果表明, 绵羊的 *HTR4* 基因编码 438 个氨基酸残基, 其分子式为 $C_{2253}H_{3497}N_{577}O_{605}S_{34}$, 分子质量为 49 437.09 KDa, 理论等电点 pI 为 8.32。其氨基酸组成如图 2 所示, 其中含量最多的是 Leu (亮氨酸), 所占比例为 10.27%; 含量最少的是 His (组氨酸), 所占比例为 1.83%。负电荷残基总数 (Asp+Glu) 为 32, 正电荷残基总数 (Arg+Lys) 为 37。基因编码产物半衰期和不稳定指数分别为 30 h 和 51.70, 不稳定指数为 41.57 > 40.00, 由此可确定绵羊 *HTR4* 基因编码产物属不稳定蛋白。

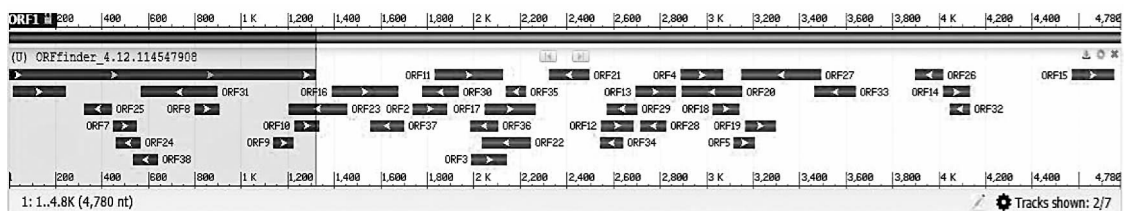


图 1 绵羊 *HTR4* 基因序列的 ORF 分析

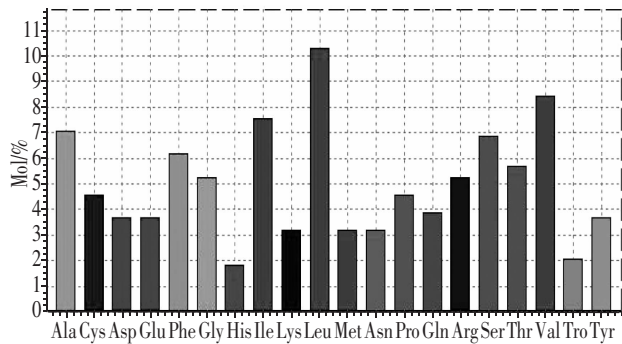


图2 绵羊 HTR4 基因编码的蛋白氨基酸组成

2.3 绵羊 HTR4 蛋白亚细胞定位

从绵羊 HTR4 基因编码蛋白的亚细胞定位结果(表1)可以看出, 绵羊 HTR4 蛋白分布于质膜的可能性为 60.9%, 分布于囊泡分泌系统的可能性为 17.4%, 分布于内质网的可能性为 8.7%, 分布于细胞核和线粒体的可能性均为 4.3%。由此推断, 绵羊 HTR4 基因主要在质膜中发挥生物学作用, 其次在囊泡分泌系统中发挥作用。

表1 HTR4 编码产物的亚细胞定位预测结果

亚细胞定位	概率 /%
质膜	60.9
囊泡分泌系统	17.4
内质网	8.7
细胞核	4.3
线粒体	4.3

2.4 绵羊 HTR4 编码产物的同源性比对及系统发育分析

HTR4 基因在多个物种中都有表达。采用 DNAMAN 对兔、牛、绵羊、山羊、猫、马、人、豚鼠、猪、狗和鸡 HTR4 蛋白多序列比对的结果如图 3 所示, 可以看出, HTR4 基因在这 11 个物种中均有表达, 且绵羊与牛和山羊的 HTR4 氨基酸序列同源性较高, 这也说明了它们在进化过程具有较近的亲缘关系。HTR4 编码产物同源树证明, 在分析中选取的 11 个物种中, 绵羊、山羊

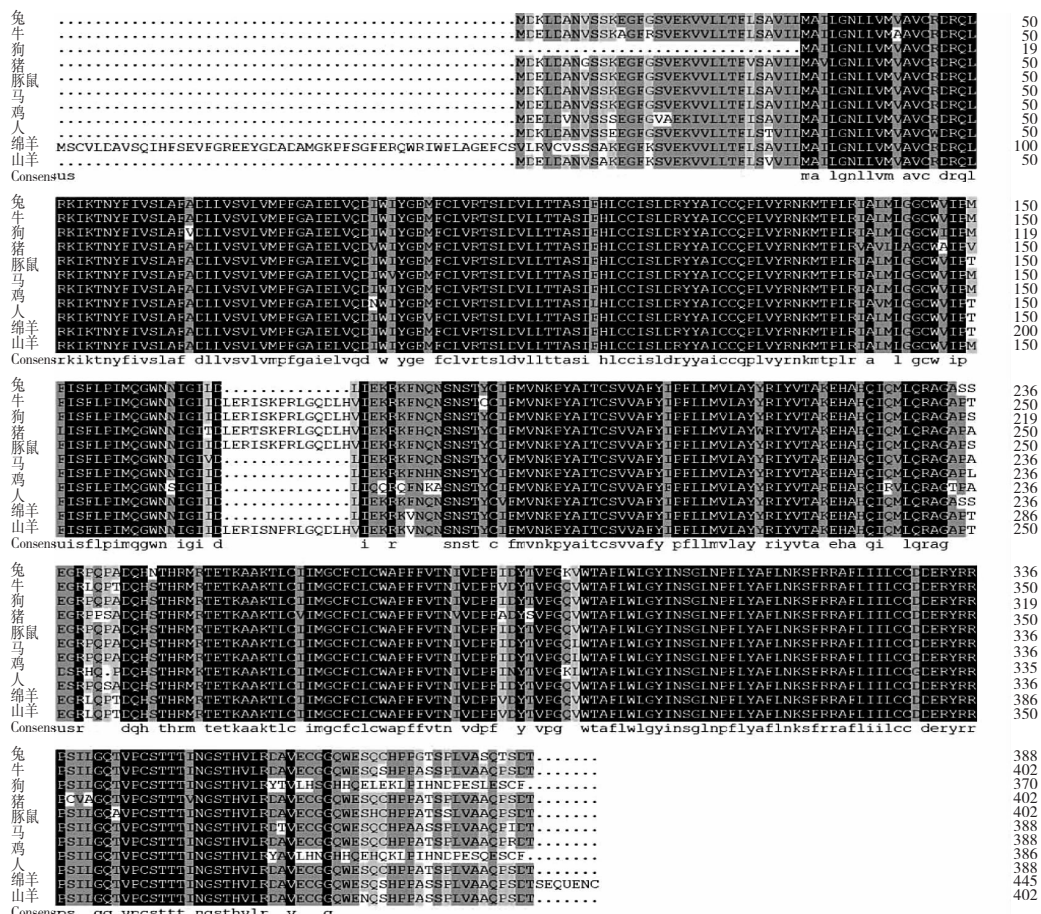


图3 11个物种的 HTR4 基因编码产物序列的同源性分析

和牛的同源性最高(图 4)。

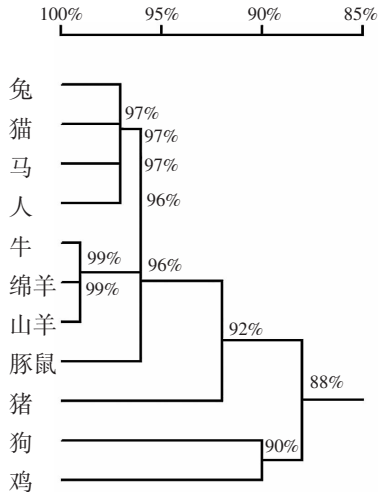


图 4 11 个物种的 HTR4 基因编码产物序列的同源树

2.5 绵羊 HTR4 蛋白潜在信号肽剪切位点预测

信号肽序列是存在于分泌蛋白基因编码序列中、在起始密码子之后的 1 段富含疏水氨基酸多肽的序列。蛋白质的合成是在核糖体上进行，信号肽是蛋白质的 1 个片段，引导核糖体并定位于内质网上的一个通道上，使核糖体附着在内质网上，并将不断伸长的蛋白质链穿透通过通道，随后信号肽被切割下来，合成的蛋白质被释放进内质网腔，最后被转运到胞外^[11]。通过检测绵羊 HTR4 蛋白潜在信号肽的存在情况可以得知该基因编的产物是否是分泌蛋白和跨膜蛋白以及跨膜蛋白的基本信息。从图 5 可以看出，绵羊 HTR4 基因编码产物的 C 值、Y 值和 S 值分别为 0.619、0.058 和 0.463。因此推断 HTR4 基因的编码产物不存在信号肽，但发现跨膜区，故该蛋白既是非分泌性蛋白又是跨膜蛋白。

2.6 绵羊 HTR4 蛋白跨膜螺旋结构预测

从 TMHMM 2.0 软件的分析结果(图6)可以看出，HTR4 基因编码的蛋白质是跨膜蛋白，该基因编码的蛋白有 7 段跨膜区间，分别是 145 ~ 167、180 ~ 202、217 ~ 239、259 ~ 281、314 ~ 336、383 ~ 401 和 416 ~ 438。

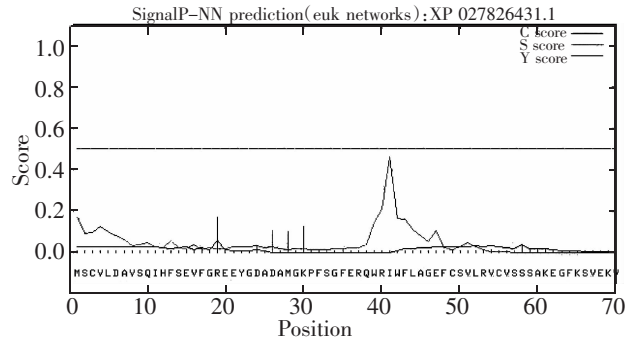


图 5 绵羊 HTR4 基因蛋白潜在信号肽剪切位点分析结果

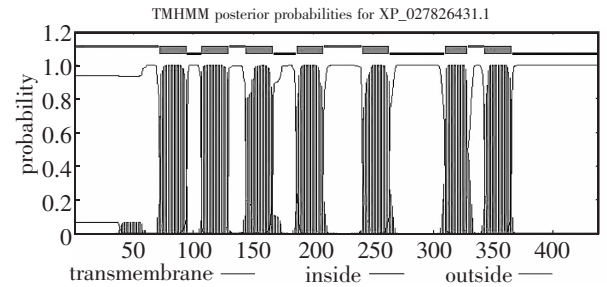


图 6 绵羊 HTR4 基因蛋白跨膜螺旋结构分析结果

2.7 绵羊 HTR4 蛋白保守结构域分析

由绵羊 HTR4 蛋白潜在信号肽剪切位点预测已知绵羊 HTR4 基因编码的蛋白是跨膜蛋白，通过 Smart 软件分析，绵羊 HTR4 有 7 段跨膜结构，分别是 72 ~ 94、107 ~ 129、144 ~ 166、186 ~ 208、241 ~ 263、310 ~ 328、343 ~ 365(图 7、图 8)。

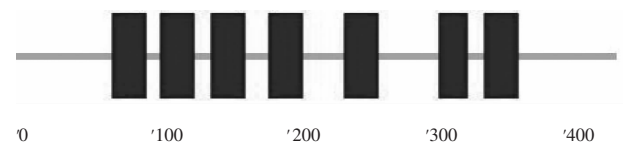


图 7 绵羊 HTR4 蛋白保守结构域分析数据

Name	Start ▲	End	E-value
transmembrane region	72	94	N/A
transmembrane region	107	129	N/A
transmembrane region	144	166	N/A
transmembrane region	186	208	N/A
transmembrane region	241	263	N/A
transmembrane region	310	328	N/A
transmembrane region	343	365	N/A

图 8 绵羊 HTR4 蛋白保守结构域

2.8 绵羊 HTR4 蛋白亲疏水性分析

研究发现，该基因编码蛋白疏水性最大值为 2.878(第 82 位)，最小值为 -3.078(第 384 位)，由此得出该基因编码的蛋白属于亲水蛋白(图 9)。

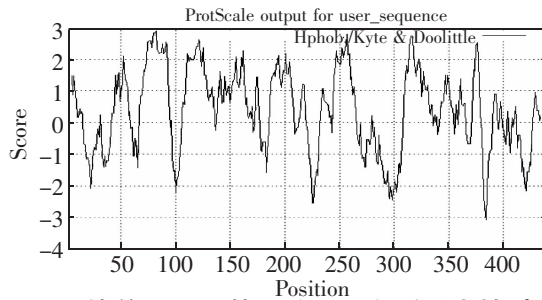


图 9 绵羊 *HTR4* 基因编码蛋白质疏水性/亲水性预测分析

2.9 绵羊 *HTR4* 蛋白二级结构的预测

二级结构(secondary structure)是指蛋白质第一个水平的折叠,即部分蛋白质链折叠形成一些通用结构,这些通用结构在所有蛋白质中都能找到^[12]。通过 Jpred 软件分析可知(图 10)绵羊 *HTR4* 蛋白二级结构 α 螺旋(α -helix, Hh)、 β 折叠(β -sheet, Ee)和无规卷曲(random coil, Cc)分别有 243、9 和 188 段,分别占比 55.22%、2.05%和42.73%,可见绵羊 *HTR4* 蛋白二级结以 α 卷曲为主。



图 10 绵羊 *HTR4* 基因编码蛋白质二级结构预测

2.10 绵羊 *HTR4* 蛋白三级结构预测与分析

三级结构(tertiary structure)是指蛋白质在二级结构基础上的进一步折叠,通过将二级结构元素组装在一起形成每个蛋白质特有的三维构象^[12]。由图 11 可知,*HTR4* 基因编码蛋白的三级结构主要由 α 卷曲折叠缠绕形成。

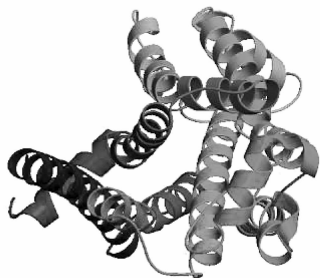


图 11 *HTR4* 基因编码蛋白的三级结构的分析结果

3 小结

绵羊 *HTR4* 基因的最长 ORF 长度为 1 317 bp, 编码 438 个氨基酸残基; 亮氨酸所占比例最多, 为 10.27%, 分子质量为 49 437.09 KDa, 理论等电点(pI)为 8.32。*HTR4* 基因编码的产物为不稳定性蛋白, 其亚细胞定位的结果显示在质膜的可能性最大, 为 60.9%。*HTR4* 基因编码产物在多种物种中氨基酸序列高度相似, 绵羊与牛、山羊在系统发育树中距离最近。*HTR4* 基因的编码产物中不存在信号肽, 为非分泌性蛋白但发现了跨膜区, 故该蛋白是跨膜蛋白。该基因编码的蛋白有 7 段跨膜结构。*HTR4* 基因编码的蛋白为亲水蛋白, 富含亲水氨基酸。绵羊 *HTR4* 基因编码产物的二级结构主要以 α 螺旋为主, 三级结构主要由 α 螺旋折叠缠绕形成。

参考文献:

- [1] CICHON S, KESPER K, PROPPING P, et al. Assignment of the human serotonin 4 receptor gene (*HTR4*) to the long arm of chromosome 5 (5q31-q33)[J]. Molecular Membrane Biology, 1998, 15(2): 75-78.
- [2] 刘丽波, 鞠桂芝, 史杰萍, 等. 5-HTR4 和 GABRG2 基因多态性与精神分裂症关系[J]. 中国公共卫生, 2007, 23(12): 1461-1463.
- [3] 于平, 尹文娟, 于萍. 5-羟色胺受体对记忆的调节作用[J]. 首都师范大学学报(社会科学版), 2009(S4): 141-145.
- [4] 李晓丽, 吴博威. 5-羟色胺-4 受体激动剂的促肠道动力作用研究[J]. 中国药物与临床, 2018, 18(8): 1279-1281.
- [5] 魏兰福. 5-羟色胺₄受体对消化道运动的作用及临床意义[J]. 医学综述, 2002, 8(11): 29-30.
- [6] 隋景巍. 绵羊 PTGER3 基因生物信息学分析[J]. 乡村科技, 2017, 4(11): 52-55.
- [7] 罗轶. 鸡 FATP1 基因 cDNA 的克隆、组织表达及其生物信息学分析[D]. 雅安: 四川农业大学动物科学技术学院, 2008: 23-24.
- [8] NAKAI KHORTON P. PSORT: a program for

定西市马铃薯块茎和土壤中 3 种农药的残留分析

杨 阳^{1,2}, 解 钧^{1,2}, 吴 迟^{1,2}, 刘新刚^{1,2}, 吴小虎^{1,2}, 董丰收^{1,2}, 徐 军^{1,2}, 郑永权^{1,2}

(1. 中国农业科学院植物保护研究所, 北京 100193; 2. 植物病虫害生物学国家重点实验室, 北京 100193)

摘要: 对定西市马铃薯示范区以及周围农户 2019 年马铃薯块茎和土壤中啞菌酯、吡唑醚菌酯、吡虫啉的残留进行了测定。结果表明, 啞菌酯、吡唑醚菌酯、吡虫啉在马铃薯块茎中的残留量分别为 0.044、ND、ND mg/kg, 均低于对照组的浓度 0.067、0.218、0.009 3 mg/kg, 并低于国际食品法典委员会(CAC)规定的最高残留量 0.10、0.02、0.50 mg/kg。此外, 在示范区马铃薯土壤中未检出啞菌酯、吡唑醚菌酯、吡虫啉。表明示范区啞菌酯、吡唑醚菌酯以及吡虫啉的使用, 在马铃薯上是安全的, 且对土壤无污染。

关键词: 马铃薯; 啞菌酯; 吡唑醚菌酯; 吡虫啉; 残留

中图分类号: S532 **文献标志码:** A **文章编号:** 1001-1463(2020)10-0040-05

[doi:10.3969/j.issn.1001-1463.2020.10.009](https://doi.org/10.3969/j.issn.1001-1463.2020.10.009)

Residue Analysis of Three Pesticides in Potato Tuber and Soil in Dingxi

YANG Yang^{1,2}, XIE Jun^{1,2}, WU Chi^{1,2}, LIU Xingang^{1,2}, WU Xiaohu^{1,2}, DONG Fengshou^{1,2}, XU Jun^{1,2}, ZHENG Yongquan^{1,2}

(1. Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China; 2. State Key Laboratory for Biology of Plant Diseases and Insect Pests, Beijing 100193, China)

Abstract: In this study, the residues of azoxystrobin, pyrazolide and imidacloprid in potato tuber and soil in Dingxi demonstration and surrounding farms in 2019 were measured. The results showed that the residue of azoxystrobin, pyrazolide and imidacloprid in potato tuber were 0.044, ND and ND mg/kg, respectively, which were lower than the concentrations of 0.067 mg/kg, 0.218 mg/kg and 0.009 3 mg/kg in the surrounding farms, and also lower than the maximum residue limit (MRL) of 0.1 mg/kg, 0.02 mg/kg and 0.5 mg/kg as determined in

收稿日期: 2020-07-03

基金项目: 国家重点研发计划(2018YFD0800806)。

作者简介: 杨 阳(1987—), 女, 安徽淮北人, 助理研究员, 博士, 主要从事农药风险评估工作。联系电话: (010)62815938。Email: 80mars@163.com。

通信作者: 郑永权(1987—), 男, 广东高州人, 研究员, 主要从事农药科学使用、农药残留、农药环境毒理、农药污染生物修复、植物源农药和化学生态学等研究工作。联系电话: (010)62815938。Email: zhengyongquan@ippcaas.cn。

- detecting sorting signals in proteins and predicting their subcellular localization[J]. Trends in Biochemical Sciences, 1999, 24(1): 34-36.
- [9] 张小雪, 潘香羽, 李发弟, 等. 绵羊 ESR 基因生物信息学分析[J]. 甘肃农业科技, 2014 (9): 30-33.
- [10] 张小雪, 李发弟, 王维民. 绵羊 ANXA10 基因生物信息学分析[J]. 甘肃农业科技, 2016(6): 5-8.
- [11] 叶方寅. 信号肽假说的提出及证实[J]. 国外医学分子生物学分册. 1999, 21(6): 377-379.
- [12] 孔繁良. 基于二级结构的蛋白质三级结构预测[D]. 济南: 济南大学, 2016: 3-27.

(本文责编: 郑立龙)