

[19] 中华人民共和国国家知识产权局



[12] 发明专利说明书

专利号 ZL 03129585.1

[51] Int. Cl.

C07K 14/47 (2006.01)

C07K 16/18 (2006.01)

C12N 15/12 (2006.01)

C12N 15/63 (2006.01)

[45] 授权公告日 2007 年 9 月 5 日

[11] 授权公告号 CN 100335500C

[22] 申请日 2003.6.27 [21] 申请号 03129585.1

[73] 专利权人 中国科学院上海生命科学研究院
地址 200031 上海市岳阳路 320 号

[72] 发明人 景乃禾 唐珂 王晨 沈承勇

[56] 参考文献

WO0150829 A2 2001.7.19

WO0183811 A1 2001.11.8

WO0202769 A1 2002.1.10

Homology of the amyloid beta protein precursor
in monkey and human supports a primate model for
beta amyloidosis in Alzheimer's disease Podlisny,
M. B. 等, NCBI 中序列 A49795 1999

审查员 汪波莉

[74] 专利代理机构 上海智信专利代理有限公司
代理人 吴林松

权利要求书 1 页 说明书 15 页 附图 5 页

[54] 发明名称

人淀粉样蛋白前体蛋白 639、其编码序列及用途

[57] 摘要

本发明公开了一种新的多肽——人淀粉样蛋白前体蛋白 639 (APP639 蛋白) 的氨基酸序列及其编码序列, 抗此多肽的抗体, 以及该多肽和核苷酸在筛选治疗老年性痴呆症的药物和制备检测肝脏发育状况的试剂盒等方面的应用。

1. 一种分离的人淀粉样蛋白前体蛋白 639, 其特征在于, 为 SEQ ID NO:1 氨基酸序列的多肽。
2. 一种分离的多核苷酸, 其特征在于, 为选自下组的一种核苷酸序列:
 - (a) 编码如权利要求 1 所述多肽的多核苷酸;
 - (b) 与多核苷酸 (a) 完全互补的多核苷酸。
3. 如权利要求 2 所述的多核苷酸, 其特征在于, 该多核苷酸编码氨基酸序列为 SEQ ID NO:1 的多肽。
4. 一种载体, 其特征在于, 它含有权利要求 2 所述的多核苷酸。
5. 一种遗传工程化的宿主细胞, 其特征在于, 它含有权利要求 4 所述的载体。
6. 一种具有活性的人淀粉样蛋白前体蛋白 639 的制备方法, 其特征在于, 该方法包含:
 - (a) 在适合表达人淀粉样蛋白前体蛋白 639 的条件下, 培养权利要求 6 所述的宿主细胞;
 - (b) 从培养物中分离出具有人淀粉样蛋白前体蛋白 639 活性的多肽。
7. 一种能与权利要求 1 所述的人淀粉样蛋白前体蛋白 639 特异性结合的抗体。

人淀粉样蛋白前体蛋白 639、其编码序列及用途

技术领域

本发明属于生物技术领域，具体的本发明涉及一种新的多肽——人淀粉样蛋白前体蛋白 639 (APP639 蛋白) 的氨基酸序列及其编码序列和它们在筛选治疗老年性痴呆症的药物和制备肝脏发育检测试剂盒等方面的应用。

背景技术

老年性痴呆症(Alzheimer's disease, AD)是目前世界上最严重的神经退行性疾病。AD 患者脑中有三种异常结构：老年斑 (senile plaques, SP)、神经纤维缠结 (neurofibrillary tangles, NFT) 和类淀粉样物质沉积脑血管 (amyloid-laden cerebral vessels)。老年斑和类淀粉样物质沉积脑血管的主要组分是含有 39—43 个氨基酸残基的 β -淀粉样蛋白 ($A\beta$ 蛋白, β -amyloid protein)。 $A\beta$ 蛋白是由淀粉样蛋白前体蛋白 (amyloid precursor protein, APP) 经蛋白水解而来。目前, 已发现 APP 基因在转录过程中可产生 8 种不同的 mRNA 形式。

老年性痴呆症(AD, Alzheimer's disease)与中、老年人群中常见的记忆力衰退密切相关。我国人口现在老龄化趋势严重, 而且 AD 患病率在不断上升。据调查, 65 岁以上人口中痴呆总患病率为 5.22%, 全国共有各类痴呆症病人 500 万, 其中 AD 病人约占 340 万。大量证据表明 $A\beta$ 在 AD 发病起始发展过程中起着中心的作用。

$A\beta$ 蛋白是由淀粉样蛋白前体蛋白 (amyloid precursor protein, APP) 经蛋白水解而来, 淀粉样前体蛋白因此得名。APP 基因位于第 21 号染色体, 是第一个被发现与家族性老年痴呆症 (familial Alzheimer's disease, FAD) 相关的基因。APP 是一个广泛性分布蛋白, 在哺乳动物身体各组织均有表达。APP 蛋白是一个典型的 I-型整合膜蛋白, 含有一个信号肽序列, 一个大的胞外氨基端区域, 一个单次跨膜区和一个小的胞内区含有 N 端信号肽, 一个大的 N 端胞外结构域 (包含一个 Cys 丰富区域、负电区和 N-糖基化区), 单次跨膜区和一个短的 C 端胞内区域。APP 基因位于 21 号染色体上, 其转录产物为 3.2-3.4 kb 的 mRNA, APP 基因在许多组织中均表达, 其表达在肾脏和脑最高。APP 基因有 18 个外显子, 分别是 1-13、13a 和 14-18 外显子。APP 基因经不同的剪接加工, 可产生 8 种不同的转录产物, 这些产物均以其所含有的氨基酸数目而命名, 其中包括 APP770、APP751、APP714、APP695、L-APPs。APP 基因已知的剪接位点主要发生在外显子 7-8 和外显子 15 两处。

APP695 cDNA 是最早分离得到的全长 APP 基因转录产物, 它编码一个含有 695 个氨基酸残基的蛋白质。与 APP695 mRNA 相比, APP751 mRNA 多一个含 168 个核苷酸的片段——外显子 7, 该外显子编码一个与 Kunitz 丝氨酸蛋白酶抑制剂家族 (KPI) 高度同源的含 56 个氨基酸残基的片段。APP770 mRNA 比 APP751 mRNA 多出外显子 8, 该外显子编码一个含有 19 个氨基酸残基的片段, 这 19 个氨基酸残基的序列与在神经元和胸腺细胞发现的 MRC OX-2 抗原具有 47% 的同源性。APP714 比 APP695 多外显子 8。L-APPs 不含第 15 号外显子, 且因它们主要在白细胞 (Leukocytes) 和脑内小胶质细胞中存在而得名。L-APPs 除了不含外显子 15 外,

在外显子 7、8 也发生相应的剪接，而产生一系列剪接产物。PP695、APP751 和 APP770 是三种最常见的 APP 基因转录产物。APP770 和 APP751 都有 KPI 结构域，它们主要在外周组织和成年灵长类动物脑内表达；而 APP695 则主要在神经系统表达。有实验显示，与正常的对照组相比，APP695 mRNA 在 AD 患者脑中的表达降低，而 APP770 mRNA 的表达水平升高。这些结果提示，APP 基因的不同剪接产物及其在转录水平上的调控可能与 AD 相关。而且，APP 基因的不同剪接、表达的调节比其可能的结构变化更重要。

因此，本领域迫切需要研究开发新的与 AD 相关的 APP 基因的剪接产物（蛋白）。

发明内容

本发明的目的就是提供一种与 AD 相关的 APP 基因的新的剪接产物人淀粉样蛋白前体蛋白 639 以及其核苷酸序列。

本发明的另一目的是提供生产这种蛋白及核苷酸序列的方法以及所述蛋白的用途。

本发明的第一方面，提供了一种分离的人淀粉样蛋白前体蛋白 639，其特征在于，它包含：具有 SEQ ID NO: 1 氨基酸序列的多肽、或其保守性变异多肽、或其活性片段、或其活性衍生物。优选的该多肽是具有 SEQ ID NO: 1 氨基酸序列的多肽。

本发明的第二方面，提供了一种分离的多核苷酸，其特征在于，它包含选自下组的一种核苷酸序列：

- (a) 编码如权利要求 1 和 2 所述多肽的多核苷酸；
- (b) 与多核苷酸 (a) 互补的多核苷酸。

优选的该多核苷酸编码具有所示 SEQ ID NO: 1 氨基酸序列的多肽。

本发明的第三方面，提供了一种载体，它含有本发明上述的多核苷酸。

本发明的第四方面，提供了一种遗传工程化的宿主细胞，它含有本发明上述的载体。较佳的，所述宿主细胞为真核细胞。更佳的，所述宿主细胞为哺乳动物细胞。最佳的，所述宿主细胞为人的细胞。

一种具有活性的人淀粉样蛋白前体蛋白 639 的制备方法，其特征在于，该方法包含：

- (a) 在适合表达人淀粉样蛋白前体蛋白 639 的条件下，培养权利要求 6 所述的宿主细胞；
- (b) 从培养物中分离出具有人淀粉样蛋白前体蛋白 639 活性的多肽。

本发明的第六方面，提供了一种能与本发明上述人淀粉样蛋白前体蛋白 639 特异性结合的抗体。较佳的，所述抗体为单克隆抗体。

本发明的第七方面，提供了一种检测肝脏发育状况的试剂盒，其特征在于，它包括特异性扩增或转录本的引物，或者与人淀粉样蛋白前体蛋白 639 特异性结合的抗体。较佳地，检测的是人淀粉样蛋白前体蛋白 639 基因或转录本，并与正常人淀粉样蛋白前体蛋白 639 核苷酸序列比较差异。

本发明的第八方面，提供了一种本发明上述的人淀粉样蛋白前体蛋白 639 的用途，其特征在于，它用于筛选治疗老年性痴呆症的药物，或制备神经营养保健品。

本发明的其他方面由于本文的技术的公开，对本领域的技术人员而言是显而易见的。

附图说明

图1 APP639 cDNA 全序列和推测的氨基酸序列。

箭头所指为 exon 1 和 exon 3 连接部位。方框中包含的是 A β -淀粉样蛋白氨基酸序列。Poly(A) 信号为下横线所示序列。

图2 RT-PCR 和测序分析 exon 2 处发生剪接的 APP 基因转录产物在人胚胎组织中存在。

A: PCR 引物和扩增片段图示。

B: RT-PCR 分析在人胚胎大脑和肝脏组织中 exon 2 处发生剪接的 APP 基因转录产物。

C: 对 RT-PCR 扩增产物的 DNA 测序分析。箭头所指分别为 exon 1 与 exon 2 (上)、exon 1 与 exon 3 (下) 连接位点。

图3 RT-PCR 和 Southern 杂交分析 APP639 mRNA 在人胚胎各种组织中的表达

图4 RT-PCR 分析 APP639 mRNA 在成年人大脑皮层和肝脏组织的表达。

A: RT-PCR 分析 APP639 mRNA 在成年人大脑皮层的表达。

B: RT-PCR 分析 APP639 mRNA 在成年人肝脏组织的表达。

图5 Western blot 分析 APP639 融合蛋白在瞬时转染的 HEK293 细胞中的表达

图6 人 APP 基因发生剪接的三个可能位点。

A: APP 基因发生剪接的三个可能位点和 APP770。

B: 与 exon 7 和 exon 8 相关的 I 位点及其剪接产物, 其中包括 APP751、APP714 和 APP695。

C: 与 exon 15 相关的 II 位点及其剪接产物 L-APPs, L-APPs 除了不含外显子 15 外, 在外显子 7、8 也发生相应的剪接, 而产生四个分别与 APP770、APP751、APP714 和 APP695 相对应的剪接产物。

D: 与 exon 2 相关的 III 位点及 APP639。

具体实施方式

本发明人经过广泛而深入的研究, 首次筛选到人 APP 基因的一个新剪接形式 APP639。测序结果显示, 新的剪接位点为第 2 外显子, 造成外显子 1 与 3 直接连接; APP639 cDNA 不含外显子 2、7 和 8, 编码蛋白含 639 个氨基酸残基。RT-PCR 和 Southern 杂交分析证实, APP639 mRNA 在胚胎组织中存在。发明人在成年人大脑皮层中未能检测到 APP639 mRNA; 然而, APP639 mRNA 在成年人肝脏有表达。Western blot 结果显示, APP639 蛋白的分子量大约为 110 kD。

APP639 cDNA 全长含 3011 个核苷酸, 核苷酸序列按 5' 到 3' 的方向排列, 序列中包括一个有 639 氨基酸残基的开放阅读框。在此基础上完成了本发明。

如本文所用, 术语“人淀粉样蛋白前体蛋白 639”和“人淀粉样蛋白前体多肽 639”可互换使用。

如本文所用, “分离的”是指物质从其原始环境中分离出来(如果是天然物质, 原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的, 但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开, 则为分离纯化的。

如本文所用, “分离的人淀粉样蛋白前体蛋白 639 或多肽”是指人淀粉

样蛋白前体蛋白 639 基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化人淀粉样蛋白前体蛋白 639。

本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。

本发明还包括人淀粉样蛋白前体蛋白 639 的片段、衍生物和类似物。

本发明的多核苷酸可以是 DNA 形式或 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与图 1 所示的编码区序列相同或者是简并的变异体。如本文所用，“简并的变异体”在本发明中是指编码具有 SEQ ID NO:1 的蛋白质，但与图 1 所示的编码区序列有差别的核酸序列。

本发明还涉及与上述的序列杂交的核酸片段。如本文所用，“核酸片段”的长度至少含 15 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码淀粉样蛋白前体蛋白 639 的多聚核苷酸。

人淀粉样蛋白前体蛋白 639 核苷酸全长序列或其片段通常可以用 PCR 扩增法、重组法或人工合成的方法获得。对于 PCR 扩增法，可根据本发明所公开的大鼠淀粉样蛋白前体蛋白 639 的核苷酸序列，尤其是开放阅读框序列来设计引物，并用市售的 cDNA 库或按本领域技术人员已知的常规方法所制备的 cDNA 库作为模板，扩增而得有关序列。当序列较长时，常常需要进行两次或多次 PCR 扩增，然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列，就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外，还可用人工合成的方法来合成有关序列，尤其是片段长度较短时。通常，通过先合成多个小片段，然后再进行连接可获得序列很长的片段。

用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。获得的转化子可以用常规方法培养，表达本发明的基因所编码的多肽。根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法(如温度转换或化学诱导)诱导选择的启动子，将细胞再培养一段时间。

在上面的方法中的重组多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但并不限于：常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

具体而言，本发明的人淀粉样蛋白前体蛋白 639 蛋白或多肽有多方面的用途。这些用途包括(但不限于)：用于筛选治疗老年性痴呆症的药物，或制备神经营养保健品，和制备人淀粉样蛋白前体蛋白 639 的抗体，可以制备检测器官尤其是肝脏发育的试剂盒，可以作为肝再生医学研究模型，以及检测老年性痴呆症

中早晚期 APP639 蛋白的表达变化。用表达的重组人淀粉样蛋白前体蛋白 639 蛋白筛选多肽库可用于寻找有治疗价值的能抑制或刺激人淀粉样蛋白前体蛋白 639 蛋白功能的多肽分子。

另一方面, 本发明还包括对完整的单克隆抗体或多克隆抗体, 而且还包括具有免疫活性的抗体片断, 如 Fab' 或 (Fab)₂ 片段; 抗体重链; 抗体轻链; 或嵌合抗体, 如具有鼠抗体结合特异性但仍保留来自人的抗体部分的抗体。

本发明的抗体可以通过本领域内技术人员已知的各种技术进行制备。例如, 纯化的人淀粉样蛋白前体蛋白 639 基因产物或者其具有抗原性的片段, 可被施用于动物 (如家兔、小鼠等) 以诱导多克隆抗体的产生。可使用多种佐剂增强免疫反应, 其中包括 (但不限于) 弗氏佐剂等。本发明的单克隆抗体可以利用杂交瘤技术来制备。抗人淀粉样蛋白前体蛋白 639 蛋白的抗体可用于免疫组织化学技术中, 检测活检标本中的淀粉样蛋白前体蛋白 639 蛋白。本发明实验已证实了人淀粉样蛋白前体蛋白 639 可作为的检测 APP639 蛋白的表达水平的分子标记物。

利用本发明蛋白, 通过各种常规筛选方法, 可筛选出与人淀粉样蛋白前体蛋白 639 发生相互作用的物质, 如抑制剂、激动剂或拮抗剂等。

本发明还提供了一种药物组合物, 它含有安全有效量的本发明人淀粉样蛋白前体蛋白 639 蛋白、其编码核酸、和/或反义核酸, 以及药学上可接受的载体或赋形剂。这类载体包括 (但并不限于): 盐水、缓冲液、葡萄糖、水、甘油、乙醇、及其组合。药物制剂应与给药方式相匹配。本发明的药物组合物可以被制成针剂形式, 例如用生理盐水或含有葡萄糖和其他辅剂的水溶液通过常规方法进行制备。诸如片剂和胶囊之类的药物组合物, 可通过常规方法进行制备。药物组合物如针剂、溶液、片剂和胶囊宜在无菌条件下制造。活性成分的给药量是治疗有效量, 例如每天约 0.01 微克/千克体重-约 0.5 毫克/千克体重。此外, 本发明的多肽还可与其他治疗剂一起使用。如果 APP639 蛋白比 APP695 产生的 A β 能力低, 而它的正常生理学功能没有变化, 那么, 可以通过用 APP639 代替 APP695 的方法, 来防止或延缓 AD 病症。

在发明人的一具体实施例中, 发明人发现了一种 APP 基因的新剪接产物 APP639, 它缺乏外显子 2, 7 和 8。利用 RT-PCR 和 Southern 杂交方法, 发明人发现 APP639 mRNA 在胚胎许多组织内表达, 尤其是肝脏中高表达; 在正常成人和 AD 病人大脑皮层中均没有考察到它的存在, 但在肝脏中依然表达。利用 Western 杂交技术, 细胞外源表达的 APP639 蛋白分子量约为 110 kD。

Muller 等在 1994 年报道, 他们在 APP 基因突变小鼠体内检测到 APP 基因的一种特殊形式的 mRNA, 其中缺少了第 2 外显子。他们在小鼠的 APP 基因的第 2 外显子中插入了一段含有 neomycin 抗性基因 DNA 序列组件和一个转录终止序列, 以达到破坏 APP 基因功能的目的。RT-PCR 结果显示, 在他们获得的 APP 基因突变纯合子小鼠的脑和其它组织中, 仍然存在 APP-特异的 RNA 转录产物, 只不过检测到的 APP mRNA 表达水平比野生型小鼠低 5-10 倍, 而且该转录产物不含有被破坏的第 2 外显子。这一结果提示, APP 基因的第 2 外显子在小鼠体内可能发生低概率的剪接。在本实验中, 发明人不仅从人胎脑 cDNA 文库中筛选到同样缺少第 2 外显子的 APP639 cDNA, RT-PCR 和 Southern 杂交分析还证实了 APP639 mRNA 在人体组织中存在。因而, 发明人和 Muller 等人的实验结果表明, 在哺乳动物体内 APP 基因在第 2 外显子发生剪接的机制的确存在。

人 APP 基因位于第 21 号染色体上, 共含有 18 个外显子。较早的研究表明 APP 基因在转录水平上的剪接主要发生在两个位点。一个位点是外显子 7、8 处,

选择性的剪接产生四种不同的转录形式：包含两个外显子的 APP770 (图 6A)，只含第 7 外显子的 APP751，只含第 8 外显子的 APP714，以及两个外显子均没有的 APP695 (图 6B)。另一个位点与第 15 外显子相关 (图 6C)，不含外显子 15 的 APP mRNA 亦有四种形式，分别对应于前面提及的四种在第 7、8 外显子处发生剪接的产物。因为它们主要在白细胞 (Leukocyte) 和脑小胶质细胞中表达，因而被命名为 L-APPs (图 6C)。发明人首次发现了人 APP 基因的一个新剪接产物 APP639，APP639 除了不含外显子 7 和 8，还没有外显子 2。发明人的实验证实 APP 基因在转录水平存在与第 2 外显子相关的第 3 个可能的剪接位点 (图 6D)。

发明人从人胎脑 cDNA 文库中克隆到缺少第 2 外显子的 APP639 (图 1)，RT-PCR 和 Southern 杂交的结果证实了这种剪接形式在胎脑和其它组织中存在 (图 3)。然而，在非痴呆老年人以及 AD 患者大脑皮层发明人均未能检测到 APP639 mRNA 的表达 (图 4)。可能有两个原因，一是缺少第 2 外显子的 APP 剪接形式在成年人脑组织里本身就不表达；另一个可能是，这种剪接形式在成年人脑组织中的表达很低，而且获得人体组织是一个冗长的过程，在组织长期保存过程中 RNA 也会自然降解，当 RNA 含量过低时，用常规手段很难检测到。RT-PCR 和 Southern 杂交的结果还显示，在胚胎期，APP639 在肝脏的表达高于脑组织 (图 3)。利用 RT-PCR 方法，发明人考察了 APP639 在成年人脑组织和肝脏的表达，发现尽管在大脑皮层中不能检测到 APP639 mRNA 的存在，但是，在肝脏组织中有 APP639 的表达 (图 4)。以上的实验证据提示，APP639 可能在肝脏中具有某种功能。

根据同源性比较，APP 蛋白 N 端区域相对保守，传统观点认为 APP 蛋白 N 端可能有保守的功能区域。已发现 APP 蛋白 N 端有一个肝素结合区，能够促进神经轴突的生长，调节神经突触的发生。APP 蛋白 N 端富含半胱氨酸，能够形成二硫键，保证 APP 蛋白 N 端的立体结构，N 端结构类似于富半胱氨酸类生长因子，这与体外发现的 APP 蛋白 N 端具有神经营养因子的功能相符。晶体结构学方法显示 APP 蛋白 N 端除肝素结合区外，还具有一个保守的疏水区域，这两个区域的存在可能与 N 端已知神经营养功能密切相关。APP639 蛋白缺少 N 端第 20 到 75 位氨基酸，但这 56 个氨基酸的缺失可能并不影响 N 端功能区域，因为从氨基酸序列来看，主要的功能域肝素结合区和疏水域并没有受很大的影响。所以，发明人推测，APP639 蛋白 N 端神经营养功能与 APP695 等其他蛋白相似。但半胱氨酸的缺失可能会造成 N 端结构的改变，从而影响它的功能。发明人实验室推测外显子 2 的缺失可能会影响 Aβ 的分泌，目前正在筛选 APP639 稳定表达细胞株。发明人相信，对 APP639 蛋白功能的进一步研究，有助于发明人阐明 APP 蛋白家族的功能。APP639 这个新剪接形式的发现，有利于发明人了解 APP 基因的功能和代谢。

但遗憾的是，发明人没有在胚胎大脑，肝脏，肾脏以及 AD 病人和正常成人脑中检测到内源性的 APP639 蛋白的表达。这很可能是因为 APP639 蛋白在体内的表达量很低。总之，根据发明人克隆的 APP639cDNA，能表达相应的缺少外显子 2 的 APP639 蛋白，而且能够被不同的 APP 抗体识别证明。

发明人发现了一种 APP 基因的新剪接产物 APP639，它缺乏外显子 2, 7 和 8。利用 RT-PCR 和 Southern 杂交方法，发明人发现 APP639 mRNA 在胚胎许多组织内表达，尤其是肝脏中高表达。在正常成人和 AD 病人大脑皮层中均没有考察到它的存在，但在肝脏中依然表达。利用 Western 杂交技术，细胞外源表达的 APP639 蛋白分子量约为 110 kD。总之，APP639 这种新剪接形式确实在体内存在。

为了进一步确证扩增得到的 460 bp 和 300 bp 左右的片段与 APP 基因转录产物的联系，PCR 产物经纯化，克隆到 pGEM-T-easy 载体 (购自 Promega 公司, USA)，

进行测序。测序结果显示, 对应于正对照 APP695 的大约 460bp 的片段含有 APP 基因第 2 外显子, 外显子 1 和外显子 2 相连接 (图 2C); 对应于正对照 APP639 的 300 bp 左右的片段不含第 2 外显子, 使外显子 1 与外显子 3 直接相连(图 2C)。APP 基因基因组序列分析表明, 上面的两种剪接过程均通过保守的 RNA 编辑 GT-AG 机制完成。

而且, 由于外显子 1 和外显子 2 分别含有 57 和 168 个核苷酸, 正好都是 3 的整倍数, 所以, 外显子 2 的缺失并不会改变转录产物编码的相应 APP 蛋白的氨基酸顺序; 由此发明人推测这可能也不会影响该 APP 蛋白行使其正常功能。以上的证据表明, APP639 这个 APP 基因新的转录形式确实存在于人体胎儿组织中。

荷兰国际脑库 (Netherlands Brain Bank, NBB) 提供来源于有良好临床记录的神经病理特征确定病例的死亡后样品。痴呆患者通过临床分析确定; 对“可能的 AD 患者”的临床诊断基于家族、医学检查和实验室测试, 而与其它痴呆病症分开。临床诊断参照 NINCDS-ADRDA 标准进行, 重症痴呆患者参照 Global Deterioration Scale 评估。非痴呆老年正常对照组成员在家族历史上和表征上均没有神经和精神病症。

由于 APP 基因是第一个被确定在遗传上与 AD 相关的基因, APP 基因的突变可导致家族遗传性 AD。除此之外, 有实验显示 APP 基因在外显子 7、8 处发生剪接后产生的不同的 mRNA 剪接产物 (APP695、APP751 和 APP770) 在脑里的表达变化也可能与 AD 相关。这些研究提示, APP 基因的不同剪接形式及其转录水平的调控可能与 AD 产生相关。这促使发明人去考察 APP639 mRNA 在 AD 患者和非痴呆老年人脑组织里的表达。

在一个具体实施例中, 发明人从荷兰国际脑库一共获得 17 大脑皮层样品, 其中有 8 个样品来源于 AD, 9 个来源于非痴呆老年对照。发明人使用引物对 P3 / P5 进行 PCR 扩增, 在所有样品中只能获得与 APP695 对应的产物, 却没能检测到与缺少 exon 2 的 APP639 相符合的扩增条带。图 4A 仅仅显示了部分样品的 PCR 分析结果, 泳道 4-8 的大脑皮层样品来源于 5 个非痴呆老年对照; 泳道 9-13 的样品来源于 5 个 AD 患者。如箭头所示, 在样品中仅能扩增得到与 APP695 对应的条带, 而没有与 APP639 对应的产物 (图 4A)。

如图 5 所示, APP639-myc 蛋白的分子量略小于 APP695-myc 蛋白, 大约在 120 kD 左右。22C11 不能识别 APP639-myc 蛋白, 而 6E10, myc 抗体均可以识别 APP695-myc 和 APP639-myc 蛋白。由于其 C 端含有大约 60 个氨基酸残基的 myc 重复序列, 所以, APP639 蛋白的分子量大约为 110 kD。

由于 APP639 mRNA 在胚胎肝脏的表达较高, 因而, 在另一具体实施例中发明人考察了它在成年人肝脏组织的表达。如图 2-7B 所示, 利用引物对 P3 / P5 进行 PCR 扩增, 发明人在一个成年人的肝脏组织中得到分别与 APP695 和 APP639 对应的两个扩增条带 (图 4 箭头指示)。这表明 APP639 的 mRNA 在成年人肝脏组织中有表达。APP639 分布表达具有发育时段以及组织的特异性, 因此可以作为器官尤其是肝脏发育检测试剂盒, 还可以作为肝再生医学研究模型。

APP639 蛋白的 N 端立体构象发生变化, 可能影响 A β 的生成和分泌。APP 蛋白也可能作为老年性痴呆症的新药物靶点。APP639 蛋白可以作为老年性痴呆症早期诊断指标, 利用 APP639 特异性抗体, 检测老年性痴呆症中早晚期 APP639 蛋白表达变化。分泌型 APP639 蛋白片断可能呈现除营养作用外的新生物学功能, 可以作为新颖的神经营养保健品。APP639 新剪接形式的发现有利于发明人研究不同剪接对老年性痴呆症的发生发展的调控机制。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如 Sambrook 等人，分子克隆：实验室手册 (New York: Cold Spring Harbor Laboratory Press, 1989) 中所述的条件，或按照制造厂商所建议的条件。

实施例 1 cDNA 文库筛选与 Southern 杂交所用探针模板的制备

- 1) cDNA 文库筛选探针模板 (Probe 1) 的制备
 - a) 使用 TRIZOL 试剂，从 SK-N-SH 细胞 (购自中国科学院上海生命科学研究院细胞库) 中提取总 RNA。
 - b) 反转录。
 - c) 以 RT-PCR 产物为模板，使用引物对 P1 / P2 进行 PCR 扩增。
 - d) 电泳，回收纯化 PCR 扩增产物，克隆进 pGEM-T-easy 载体，酶切，测序。
 - e) Xho I 和 Sac I 酶切，电泳回收纯化得到的 682 bp 片段即是 Probe 1。
- 2) Southern 杂交所用探针模板 (Probe 2 和 Probe 3) 的制备
 - a) Probe 2
以筛库获得的 APP695 cDNA 为模板，使用引物 P3 / P4 进行 PCR 扩增，所得到的 459 bp 片段即为 Probe 2。
 - b) Probe 3
用 Kpn I 和 Acc I 酶切含有 APP695 cDNA 的质粒，电泳，回收纯化得到的 167 bp 片段即为 Probe 3。

实施例 2 APP639 cDNA 的获得

发明人用 Probe 1 筛选人胎脑 cDNA 文库 (购自 GIBCO 公司)，经过两轮杂交，共获得 17 个阳性克隆。利用引物 P1 / P6 (表一) 进行 PCR 分析和酶切鉴定这些克隆。图 2-2 显示了部分阳性克隆的 PCR 结果。在 17 个阳性克隆中，有 14 个克隆含有 APP 基因的编码序列，其中一个克隆对应于 APP770，一个克隆对应于 APP751，其余 12 个克隆属于 APP695 型。发明人挑选了几个对应于 APP695 的克隆进行测序 (由基康公司测序)，结果发现 Clone I3 除了不含有第 7、8 外显子外，还缺少第 2 外显子，造成第 1 和第 3 外显子直接连接。全序列测定显示，Clone I3 全长为 3011 bp，阅读框内含有 1920 bp，推测编码的蛋白含 639 个氨基酸。因此，发明人将 Clone I3 命名为 APP639 (图 1)。

表一 PCR 引物

Primer	Region	Direction	Sequence
P1	Exon 6	Sense	5' -GAAGAGGCTGAGGAACCCTACG-3'
P2	Exon 14	Antisense	5' -TCCATTACGGAAGGAGCTCC-3'
P3	Exon 1	Sense	5' -GTTTGGCACTGCTCCTGCTG-3'
P4	Exon 4	Antisense	5' -TCTCTTTGGCGACGGTGTGC-3'
P5	Exon 6	Antisense	5' -CACCATCCTCATCGTCCTCG-3'
P6	Exon 11	Antisense	5' -TGAGCATGGCTTCCACTCTGGC-3'

实施例3 组织的获得和保存

组织来源于2个流产胎儿(由复旦大学基础医学院提供),8个具有明显病理特征偶发性AD患者和9个非痴呆老年正常对照。来自2个3个月流产胎儿的组织,在流产后的3小时以内即获得,采集的组织分装入干净1.5 ml离心管后立即投入盛有液氮的容器中,然后置于-80℃环境保存。来自AD患者和非痴呆老年正常对照组的脑组织通过荷兰国际脑库(Netherlands Brain Bank, NBB)快速尸体解剖系统获得。一旦采集到,脑组织经显微检查,立即按照标准的操作进行分离。组织样品在液氮中速冻后,保存于-80℃。成年人肝脏组织cDNA为中国科学院上海生命科学研究院生物化学与细胞生物学研究所赵慕钧研究员惠赠。

实施例4 RT-PCR和Southern杂交分析APP 639 mRNA在人体胚胎不同组织中的表达

分离实施例3所述流产胚胎的不同组织,分别抽取RNA,利用引物P3/P5进行RT-PCR(图3A),以实施例1所述探针2、探针3与PCR扩增产物进行Southern杂交,分析APP mRNA在胚胎脑、肾脏、肝脏、肌肉、心脏、胰腺、肺和小肠等组织中的表达。以APP639 cDNA和APP695 cDNA分别作为正对照;双蒸水作为负对照,进行Southern杂交。图3中APP639 cDNA(泳道1),APP695 cDNA(泳道2),双蒸水(泳道3),总RNA分别来源于大脑(泳道4)、肾脏(泳道5)、脾脏(泳道6)、肺(泳道7)、心脏(泳道8)、肌肉(泳道9)、肝脏(泳道10)、小肠(泳道11)。

图3B显示APP639 mRNA在人胚胎各种组织中的表达。在这些所有组织中,均扩增得到分子量在744 bp和576 bp左右的两个片段,它们分别对应于APP695和APP639。744 bp的片段应该含有第2外显子,576 bp片段则不含第2外显子。图3C为使用Probe 2为探针进行的Southern杂交。Probe 2包含exon 2及其两侧的序列,与两个扩增片段均能杂交。图3D为使用Probe 3为探针进行的Southern杂交。Probe 3仅仅含有exon 2的序列,只能和含exon 2的扩增片段杂交,因而只能识别744 bp的大片段。发明人使用Probe 2和Probe 3进行Southern杂交。如图3A所示,Probe 2含有完整的第2外显子及其两侧的邻近序列,对应于APP695 cDNA的第158到第616 bp之间的核苷酸序列;因而,其能与外显子2以及外显子2外的邻近区域杂交。而Probe 3仅含有第2外显子的序列,因而,它只能与含有外显子2序列的核酸片段杂交。由此可见,图3B中扩增得到的两个片段无论包含或不包含外显子2,都能与Probe 2杂交;而只有含第2外显子的PCR扩增片段方能被Probe 3标记。Southern杂交的结果证实了发明人的推测。以APP695和APP639 cDNA为正对照,每个泳道中的两条PCR扩增产物片段都能被Probe 2识别(图3C);然而,Probe 3却只能标记分子量较大的与APP695对应的扩增条带(图3D)。以上RT-PCR和Southern杂交分析显示,APP639 mRNA在发明人所检测的所有人胚胎组织中均有表达,其表达在肝脏和肌肉比较丰富,在其它组织中较低(图3C)。发明人使用引物对P3/P4重复以上的实验,获得了相同的结果。

在人胚胎大脑和肝脏组织中均能得到相应的两个扩增片段,分子量大约为300 bp的片段对应于APP639,460 bp左右的片段相应于APP695(图2-4A)。而且,在胎脑中460 bp片段的表达高于300 bp片段,在胎肝中则正好相反,300 bp

片段的表达高于 460 bp 片段 (图 2)。

为了进一步确证扩增得到的 460 bp 和 300 bp 左右的片段与 *APP* 基因转录产物的联系, PCR 产物经纯化, 克隆到 pGEM-T-easy 载体(购自 Promega 公司, USA), 进行测序。测序结果显示, 对应于正对照 APP695 的大约 460bp 的片段含有 *APP* 基因第 2 外显子, 外显子 1 和外显子 2 相连接 (图 2C); 对应于正对照 APP639 的 300 bp 左右的片段不含第 2 外显子, 使外显子 1 与外显子 3 直接相连(图 2C)。

实施例 5 RT-PCR 分析 APP639 mRNA 在成年人大脑皮层和肝脏组织的表达。从实施例 3 所述大脑皮层中抽取 RNA, 利用引物 P3 / P5 进行 RT-PCR, 进行 DNA 聚丙烯酰胺凝胶电泳 (浓度为 5%) 及银染。图 4A 中 APP639 cDNA (泳道 1) 和 APP695 cDNA (泳道 2) 分别作为正对照; 双蒸水 (泳道 3) 作为负对照。发明人在获得的所有大脑皮层样品中 (来源于 AD 患者: n=9; 来源于非痴呆老年对照: n=10) 都没能检测到缺少 exon 2 的 *APP* 基因转录产物。图中仅仅显示了部分样品的 PCR 扩增结果, 泳道 4-8 的样品来源于 5 个非痴呆老年对照; 泳道 9-13 的样品来源于 5 个 AD 患者; 如图中箭头所示, 在所有大脑皮层组织中, 仅能扩增得到与 APP695 对应的条带。图 4A 显示 APP639 mRNA 在成年人大脑皮层的表达。

发明人在仅有的一个肝脏样品反转录产物中 (中国科学院上海生命科学研究院生物化学与细胞生物学研究所赠送), 扩增得到了分别与 APP695 和 APP639 对应的条带 (箭头所示)。图 4 B: 显示 APP639 mRNA 在成年人肝脏组织的表达。

实施例 6 myc-APP639 / pcDNA3 真核表达载体的构建

- 1) 以 APP639 cDNA 为模板, 使用 5' 引物 (5' -TTT **TCG TGA** GAT GCT GCC CGG TTT G-3') 和 3' 引物 (5' -GAG **TCG TGA** CTA GTT CTG CAT CTG C-3') 进行 PCR 扩增。(在引物中导入 **Xba I** 酶切位点)
- 2) 纯化 PCR 扩增片段, 使用 Xba I 酶切, 37°C > 4 小时; 同时对载体 myc / pcDNA3 (购于美国 Invitrogen 公司) 亦使用 Xba I 酶切, 37°C 2 小时, CIP 处理 1 小时。
- 3) 电泳, 纯化回收插入片段和载体片段, 定量。
- 4) 连接, 16°C, 过夜。
- 5) 电转化细菌, 涂 AP 板, 37°C 培养过夜。
- 6) 挑克隆, 酶切鉴定。

实施例 7 APP639 融合蛋白在 HEK293 细胞中的表达

从得到的 APP639 cDNA 全序列, 发明人推测其编码蛋白含 639 个氨基酸残基。为了进一步证实该 cDNA 序列确实编码相应的蛋白质, 发明人将 APP639 cDNA 克隆进真核细胞表达载体 myc / pcDNA3, 并把 APP639 分别瞬时转染入空载体的 HEK293 细胞 (空白对照, 图 5 中 1, 4, 7 泳道, HEK293 细胞购自中国科学院上海生命科学研究院细胞库), APP695-myc / pcDNA3 的 HEK293 细胞 (图 5 中 2, 5, 8 泳道) 及 APP639-myc / pcDNA3 的 HEK293 细胞 (图 5 中 3, 6, 9 泳道), 利用 Western Blot 的手段检测了 APP639 蛋白的表达情况。利用 myc 抗体 (购于美国 Invitrogen 公司), 成功地检测到 APP695-myc 融合蛋白 (图 5 泳道 2, 圆点所示) 和 APP639-myc 融合蛋白 (图 5 泳道 3, 菱形所示), 而在对照组没有检测到相应条带。同时通过 APP 抗体 22C11 (购于德国 Roche 公司, 识别 APP695 蛋白的第 66-81 位氨基酸, 识别 APP 蛋白 N 端第 66-81 位氨基酸) 和 6E10 (购于美国 Signet 公司, 识别 Aβ

第 1-17 位氨基酸，对应于 APP695 蛋白的第 597-613 位氨基酸)进一步确认了相关蛋白。

22C11 能检测到内源性 APP695 蛋白(三角形所示)和外源 APP695-myc 蛋白，但不能检测到 APP639-myc 蛋白。而 6E10 由于识别的是 A β 第 1-17 位氨基酸，所以不但能识别内源性 APP695 蛋白和外源 APP695-myc 蛋白，而且能检测到 APP639-myc 蛋白，这进一步指认了 APP639-myc 蛋白条带，而且说明 APP639-myc 蛋白确实缺少外显子 2。如图 5，APP639-myc 蛋白的分子量略小于 APP695-myc 蛋白，大约在 120 kD 左右。由于其 C 端含有大约 60 个氨基酸残基的 myc 重复序列，所以，发明人推测 APP639 蛋白的分子量大约为 110 kD。文献报道，APP 蛋白的分子量在 90 kD-120 kD 不等。发明人的结果与之符合。

实施例 8 分析外显子 2 的缺失对 APP 蛋白代谢的影响

如实施例 6 的方法构建 APP639/pcDNA3, APP695/ pcDNA3 载体，建立 APP639, APP695 稳定表达细胞株，建立 A β 检测模型，分析外显子 2 的缺失对 APP 蛋白代谢的影响，分析 A β 生成的变化。

实施例 9 考察 APP639 对动物各器官尤其是肝脏的发育及功能影响

过表达 APP639 于动物胚胎中，考察 APP639 的存在对动物各器官尤其是肝脏的发育及功能影响；建立 APP639 knock out 动物模型，考察 APP639 的抑制对动物各器官尤其是肝脏的发育及功能影响，如肝脏的表型变化。

实施例 10 观察 APP639 蛋白对神经细胞的分化影响以及可能的轴突突触生长作用

培养 APP639 细胞，收集条件培养液，加到神经细胞培养液中，观察对神经细胞的分化影响以及可能的轴突突触生长作用。

SEQUENCE LISTING

<110> 中国科学院上海生命科学研究院

<120> 人淀粉样蛋白前体蛋白 639、其编码序列及用途

<130> 030626

<160> 1

<170> PatentIn version 3.1

<210> 1

<211> 639

<212> PRT

<213> 人淀粉样蛋白前体蛋白 639

<400> 1

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
1 5 10 15

Ala Leu Glu Val Tyr Pro Glu Leu Gln Ile Thr Asn Val Val Glu Ala
 20 25 30

Asn Gln Pro Val Thr Ile Gln Asn Trp Cys Lys Arg Gly Arg Lys Gln
 35 40 45

Cys Lys Thr His Pro His Phe Val Ile Pro Tyr Arg Cys Leu Val Gly
 50 55 60

Glu Phe Val Ser Asp Ala Leu Leu Val Pro Asp Lys Cys Lys Phe Leu
65 70 75 80

His Gln Glu Arg Met Asp Val Cys Glu Thr His Leu His Trp His Thr
 85 90 95

Val Ala Lys Glu Thr Cys Ser Glu Lys Ser Thr Asn Leu His Asp Tyr
 100 105 110

Gly Met Leu Leu Pro Cys Gly Ile Asp Lys Phe Arg Gly Val Glu Phe
 115 120 125

Val Cys Cys Pro Leu Ala Glu Glu Ser Asp Asn Val Asp Ser Ala Asp
 130 135 140

Ala Glu Glu Asp Asp Ser Asp Val Trp Trp Gly Gly Ala Asp Thr Asp
 145 150 155 160

Tyr Ala Asp Gly Ser Glu Asp Lys Val Val Glu Val Ala Glu Glu Glu
 165 170 175

Glu Val Ala Glu Val Glu Glu Glu Glu Ala Asp Asp Asp Glu Asp Asp
 180 185 190

Glu Asp Gly Asp Glu Val Glu Glu Glu Ala Glu Glu Pro Tyr Glu Glu
 195 200 205

Ala Thr Glu Arg Thr Thr Ser Ile Ala Thr Thr Thr Thr Thr Thr Thr
 210 215 220

Glu Ser Val Glu Glu Val Val Arg Val Pro Thr Thr Ala Ala Ser Thr
 225 230 235 240

Pro Asp Ala Val Asp Lys Tyr Leu Glu Thr Pro Gly Asp Glu Asn Glu
 245 250 255

His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala Lys His Arg
 260 265 270

Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala Glu Arg Gln
 275 280 285

Ala Lys Asn Leu Pro Lys Ala Asp Lys Lys Ala Val Ile Gln His Phe

290	295	300
Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala Asn Glu Arg Gln 305	310	315
Gln Leu Val Glu Thr His Met Ala Arg Val Glu Ala Met Leu Asn Asp 325	330	335
Arg Arg Arg Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu Gln Ala Val 340	345	350
Pro Pro Arg Pro Arg His Val Phe Asn Met Leu Lys Lys Tyr Val Arg 355	360	365
Ala Glu Gln Lys Asp Arg Gln His Thr Leu Lys His Phe Glu His Val 370	375	380
Arg Met Val Asp Pro Lys Lys Ala Ala Gln Ile Arg Ser Gln Val Met 385	390	395
Thr His Leu Arg Val Ile Tyr Glu Arg Met Asn Gln Ser Leu Ser Leu 405	410	415
Leu Tyr Asn Val Pro Ala Val Ala Glu Glu Ile Gln Asp Glu Val Asp 420	425	430
Glu Leu Leu Gln Lys Glu Gln Asn Tyr Ser Asp Asp Val Leu Ala Asn 435	440	445
Met Ile Ser Glu Pro Arg Ile Ser Tyr Gly Asn Asp Ala Leu Met Pro 450	455	460
Ser Leu Thr Glu Thr Lys Thr Thr Val Glu Leu Leu Pro Val Asn Gly 465	470	475
		480

Glu Phe Ser Leu Asp Asp Leu Gln Pro Trp His Ser Phe Gly Ala Asp
 485 490 495

Ser Val Pro Ala Asn Thr Glu Asn Glu Val Glu Pro Val Asp Ala Arg
 500 505 510

Pro Ala Ala Asp Arg Gly Leu Thr Thr Arg Pro Gly Ser Gly Leu Thr
 515 520 525

Asn Ile Lys Thr Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe
 530 535 540

Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe
 545 550 555 560

Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val
 565 570 575

Gly Gly Val Val Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu
 580 585 590

Lys Lys Lys Gln Tyr Thr Ser Ile His His Gly Val Val Glu Val Asp
 595 600 605

Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn
 610 615 620

Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn
 625 630 635

	GCGAGG	7	
CACCCGAGGAGCGGTGCGGGGGCCCGGGAGACGGCGCGGTGCGGGCGGGCGAGG		67	TTGAGCATGTGCGCATGGTGGATCCCAAGAAAGCCGCTCAGATCCGGTCCAGGTTATG
CAAGGACGCGCGGATCCCACTCGCACAGCAGCGCACTCGGTGCCCGCGCAGGGTCCGG		127	F E N V R M V D P K K A A Q I R S Q V M
			390
ATGCTGCCCGGTGGCACTGCTCCTGCTGGCCGCTGGACGGCTCGGGCCCTGGAGTTC		187	ACACACCTCCGTGTATTATGAGCGCATGAATCAGTCTCTCCCTGCTACAACTG
M L P G L A L L L L A A W T A R A L E V			T H L R V I Y E R M N Q S L S L L Y N V
			410
TACCTGACTGCAGTACCAATGTGTAGAGCCACCAACCACTGACCATCCAGAAC		247	CCTGCAGTGGCCGAGGATTGAGGATGAAGTTGATGAGCTGCTCAGNAGAGCAAAAC
Y P E L Q I T N V V E A N Q P V T I Q N			P A V A E E I Q D E V D E L L Q K E Q N
			430
TGGTCAGACGGGGCCGCAAGCACTGCAGACCCATCCOCACCTTTGTGATCCCTACCGC		307	TATTCAGTACGCTCTTGGCCACATGATTAGTGAACCAAGCATCAGTTACGGAAGCAT
W C K R G R K Q C K T H P H F V I P Y R			Y S D D V L A N H I S E P R I S Y G N D
			450
TGCTTAGTTGGTAGTTTGTAGTGTATGCCCTCTCTCTGACAAAGTCAAAATCTTA		367	GCTCTCATGCCATCTTTGACCGAAACGAAACACCGTGGAGCTCCTTCCCGTGAATGA
C L V G E F V S D A L L V P D K C K F L			A L M P S L T E T K T T V E L L P V N G
			470
CACAGGAGAGGATGGATGTTGCGAAATCATCTTCACTGGCACCCCTCGCAAGAG		427	GAGTTCAGCCTGACGATCTCCAGCCGTGGCACTCTTTGGGGCTGACTCTGTGCCAGCC
H Q E R M D V C E T H L H W H T V A K E			E F S L D D L Q P W H S F G A D S V P A
			490
ACATCGATGAGAGAGTACCACTTGCATGACTACCGCATGTTGCTGCCCTCGGAAT		487	AACACAGAAACGAGTTGAGCCTGTTGATGCCCGCCCTGCTGCCGAGGACTGACC
T C S E K S T N L H D Y G H L L P C G I			N T E N E V E P V D A R P A A D R G L T
			510
GACAAATCCGAGGGTAGAGTTTGTGTGTTGCCCACTGGCTGAGAAATGACAAATGTG		547	ACTCGACAGGTTCTGGTTGACAAATATCAAGCGGAGGATCTCTGAATGAAGATG
D K F R G V E F V C C P L A E E S D N V			T R P G S G L T N I K T E E I S E V K M
			530
GATTCCTGATCGGAGGAGGATGACTCGGATCTCTGTGGGGCGGAGCAGACAGAC		607	GATCGCAATCCGACACTGACTCAGGATGAGGTTTATCATCAAAATTTGTTCTTT
D S A D A E E D D S D V W W G G A D T D			<u>D A E F R H D E G Y E V H H Q K L V F F</u>
			550
TATCGAGTGGAGTGAAGCAAGTAGTAGAGTAGCAGAGGAGGAGAAATGGCTGAG		667	GCAGAGATGGGTCAAAACAAGTGCATCATTGGACTATGTTGGCGGTGTTGTC
Y A D G S E D K V V E V A E E E E V A E			<u>A E D V G S N H G A I I G L H V G G V V</u>
			570
GTGGAGAGAGAGCCCGATGATCGAGGAGGATGAGGATGGTGTGATGAGTGAAGAA		727	ATAGCGACAGTGTCTCATCAGCTTGGTGTGCTGGAAGAGAAACATACACATCCATT
V E E E E A D D D E D D E D G D E V E E			<u>I A</u> T V I V I T L V M L K K K Q Y T S I
			590
GAGGCTGAGAACCTCAGAGAGGACACAGAGGACCCAGCATTGCCACCACCACC		787	CATCATGGTGTGGTGGAGTTGACCGCGCTGTCAACCAGAGAGCGCCACCTGTCCAG
E A E E P Y E E A T E R T T S I A T T T			H H G V V E V D A A V T P E E R H L S K
			610
ACCACCCACAGATCTGTGGAGAGTGGTTCGATCTCTACACAGCAGCGATACC		847	ATGCAAGAGCGGCTACGAAATCCCAACTACAGTTCTTTGAGCAGTACGAACTAG
T T T T E S V E E V R V P T T A A S T			M Q Q N G Y E N P T Y K F F E Q M Q N *
			630
CCTGATCCGTTGACAAATATCTCGAGACACTGGGATGAGAAATGAACATGCCATTTTC		907	ACCCCGCCACAGCAGCCTCTGAATGACAGCAAAACCATTTGCTTCACTACCATCGG
P D A V D K Y L E T P G D E N E H A N F			TGTCCATTTATAGAAATATGTGGGAGAAACAACCCGTTTATGATTAATATCG
			210
CAGAAAGCCAAAGAGAGGCTTGAGCCCAAGCAGGAGAGAAATGTCAGGTCTATGAA		967	CCTTTGACAGCTGTCTGTAAACAAGTAGTGGCTGAACTTGAATTAATCCACATC
Q K A K E R L E A K H R E R M S Q V M R			AGTAATGTATCTATCTCTCTTACATTTTGGTCTATACTACATTAATGGTTTT
			270
GATGGAGAGGCGCAACCTCAAGCAAGAACTTGCCTAAAGCTGATAAGAAGCGATT		1027	GTGTACTGTAAAGAAATTAAGCTGTATCAAACTAGTGCATGATGATCTCTCTGATTA
E W E E A E R Q A K N L P K A D K K A V			TTTATCAGTACCCCTTAGCCAGTTGTATATATTCTTGTGTTGTGACCCAAATZAG
			290
ATCCAGCAITTCAGAGAGAAAGTGGAACTTTGGAAAGGAGAGCAGCCACCGAGACAG		1087	TCTACTTTACATAGCTTTAAGATCGTGGGGATGCTTATGTGACGCGGAGTTTC
I Q H F Q E K V E S L E Q E A A N E R Q			AGCTGCTCTCTGCGTAAGTATTCCTTCTGTATCAGTATGCAITTTAAGTTAAACAT
			310
CAGCTGTGGAGACACATGCCAGAGTGGAGCCATGCTCAATGACCGCCGCGCCTG		1147	TTTAAGTATTCAGATGCTTTAGAGAGTTTTTTTTCCATGACTGCAITTTACTGTACA
Q L V E T H M A R V E A H L N D R R L			GATTCGCTCTCTGCTATATTTGTGATAGGAATTAAGAGGATACACAGTTTGTTCCT
			330
GCCCTGAGAACTACATGACCGCTCTGAGGCTGTCTCTCGGCTCTGACGTTCTC		1207	TCGTGCTGTTTTATGTGCACACATTAAGCATTGAGACTTCAAGCTTTCTTTTTTGTG
A L E N Y I T A L Q A V P P R P R H V F			CACGATCTTTGGGCTTTGTAAGAAAGAAATCCCTGTTCAATGTAAGCACTTTTACG
			350
AATATGTAAGAAATATGTCCGCGCAACGAGGAGCAGCAGCAGCAGCAGCCCTAAGCCT		1267	GGCGGGTGGGAGGGGTGCTCTGCTGCTTCAATTAACAAAGATCTCCAAACAAAT
N H L K K Y V R A E Q K D R Q H T L K H			TTCTGCAGGATGATGTACAGAAATCATGCTTATGACATGATCGCTTTCTACACTGAT
			370
			1327
			1387
			1447
			1507
			1567
			1627
			1687
			1747
			1807
			1867
			1927
			1987
			2047
			2107
			2167
			2227
			2287
			2347
			2407
			2467
			2527
			2587
			2647
			2707
			2767
			2827
			2887

图 1

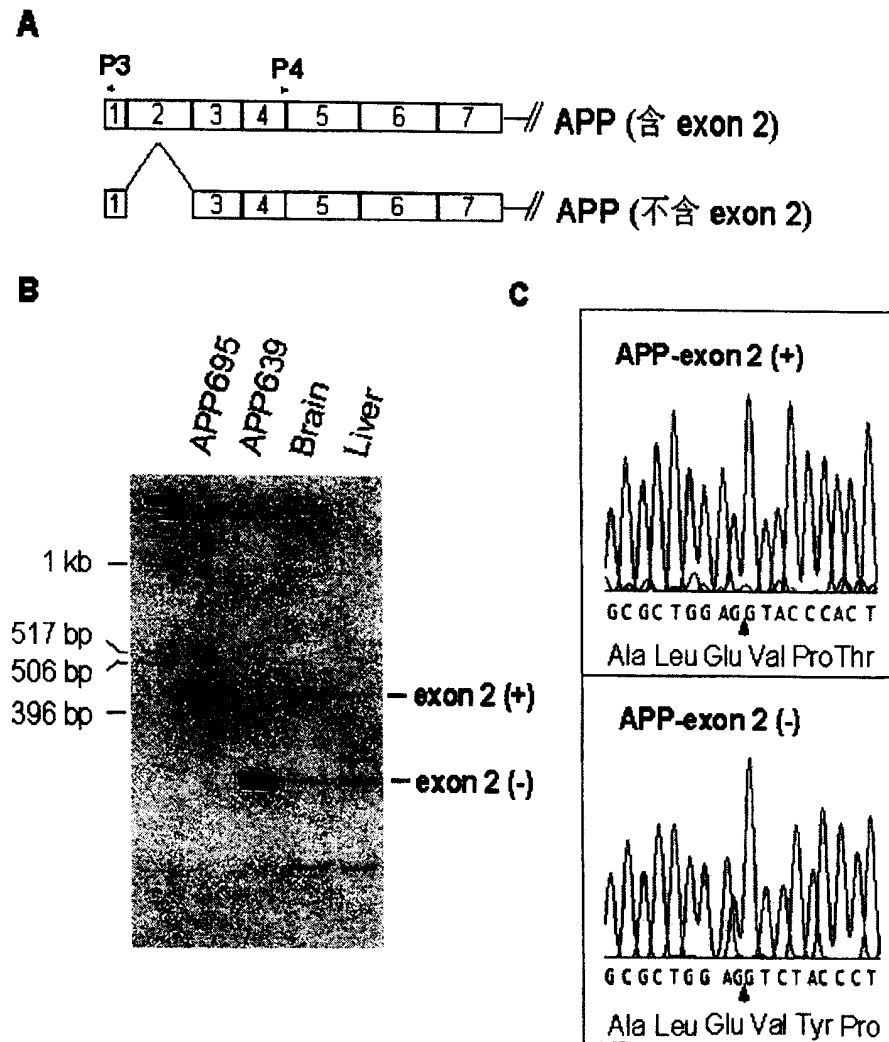


图 2

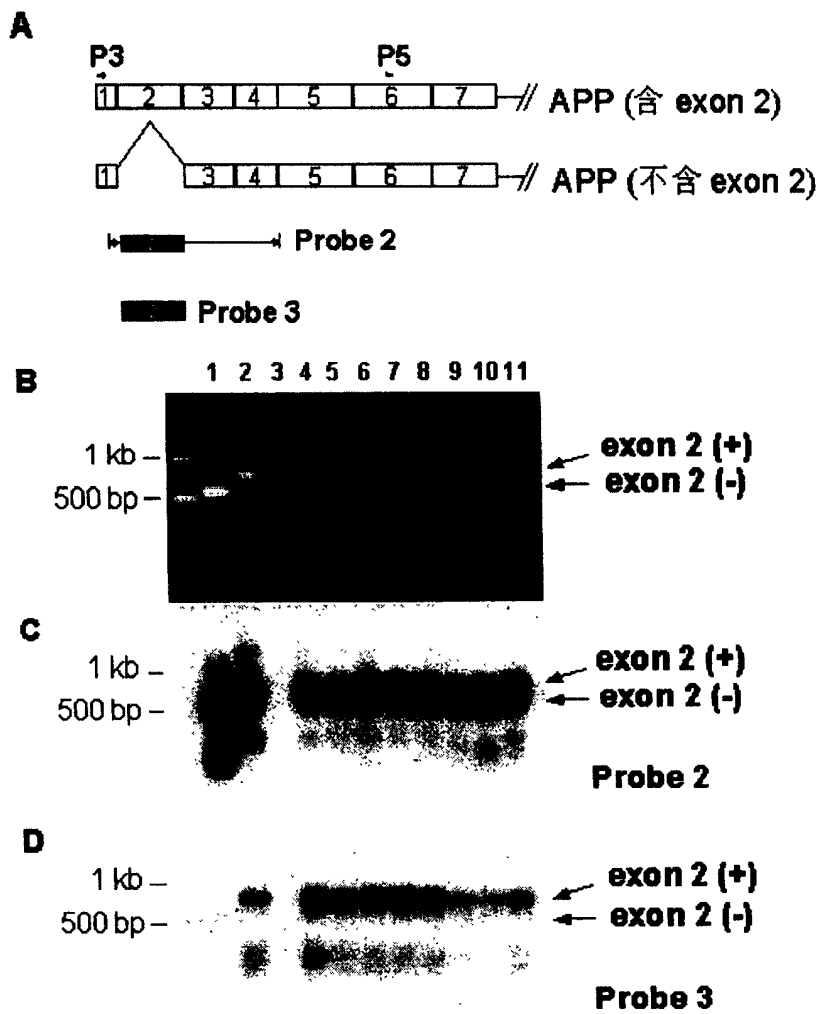


图 3

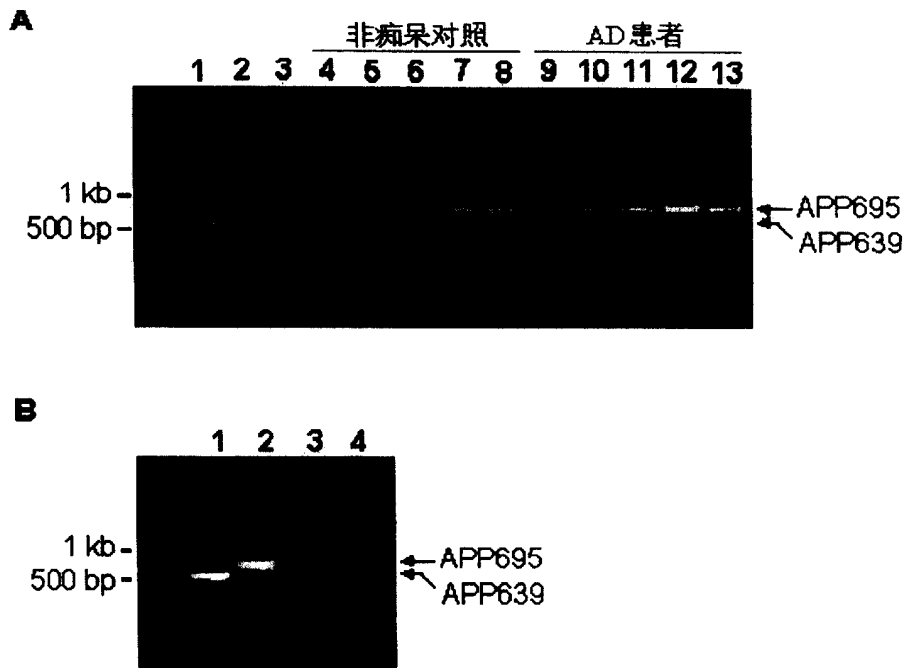


图 4

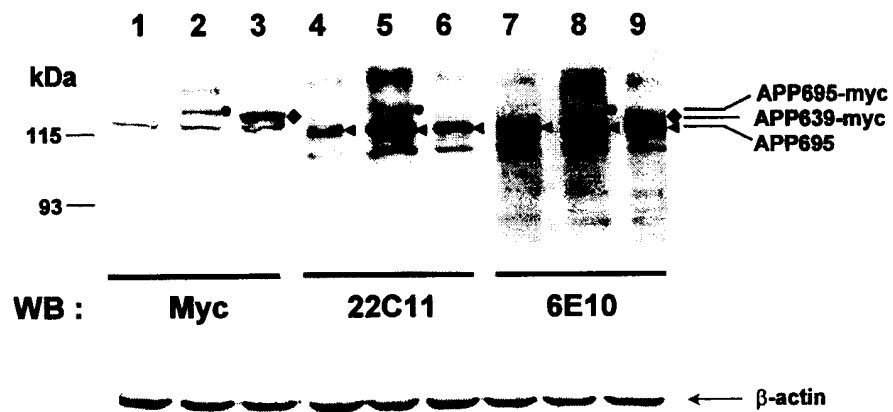


图 5

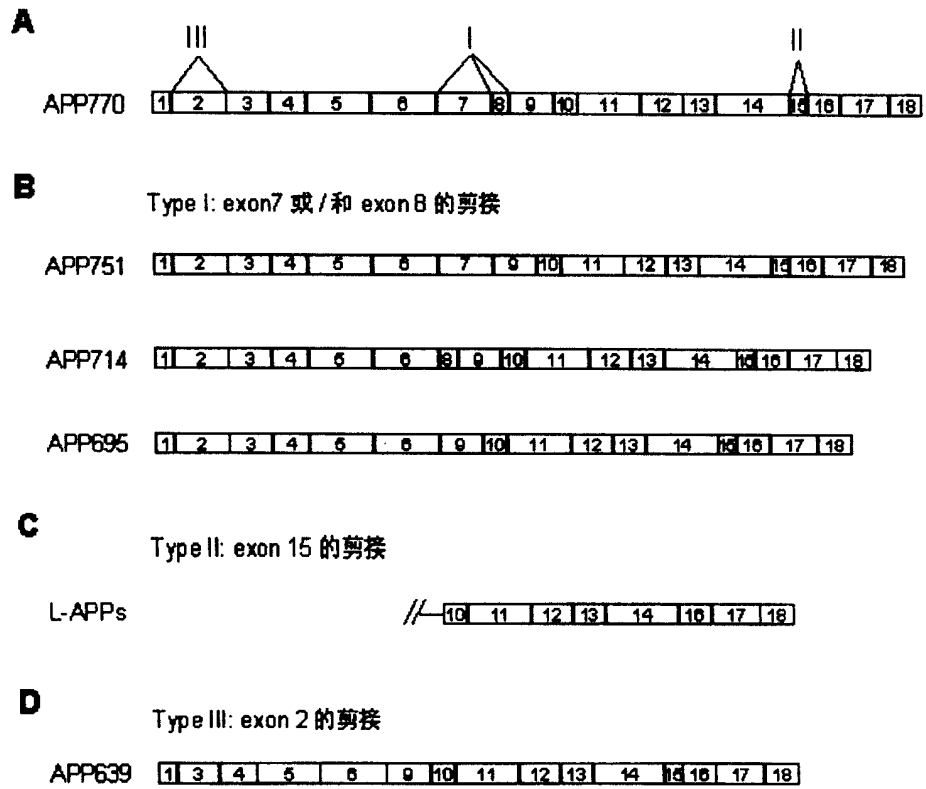


图 6