(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2018 年 10 月 4 日 (04.10.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/176808 A1

(51) 国际专利分类号:
G01N 30/88 (2006.01)

(21) 国际申请号: PCT/CN2017/106915

(22) 国际申请日: 2017 年 10 月 19 日 (19.10.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710196769.9 2017年3月29日 (29.03.2017) CN

(71) 申请人: 山 东 大 学 (SHANDONG UNIVERSITY) [CN/CN]; 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。

(72) 发明人: 杨静华(YANG, Jinghua); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 陈新骏(CHEN, Xinjun); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 吕鑫(LV, Xin); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 赵涵(ZHAO, Han); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 李翠玲(LI, Cuiling); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 毕白斌(BI, Baibin); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 巩晶(GONG, Jing); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 王凤芹(WANG, Fengqin); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 孙胜楠(SUN, Shengnan); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 王兴元(WANG, Xingyuan); 中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 陈子

(54) **Title:** SCREENING AND USE OF BIOMARKER RELATED TO SEVERE OLIGOASTHENOSPERMIA

(54) **发明名称:** 与重度少弱精子症相关的生物标志物的筛选与应用

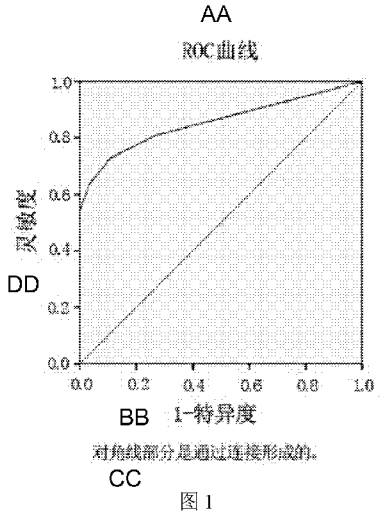

图 1

AA ROC curve
BB 1-specificity
CC The diagonal part is formed by connection.
DD Sensitivity

(57) **Abstract:** Provided is a method for screening a biomarker related to severe oligoasthenospermia. The method comprises firstly carrying out mass spectrometry on non-coding amino acids of multiple groups of sperm proteins in the case of severe oligoasthenospermia by means of a NanoHPLC-MS/MS mass spectrometry system and an non-labelled quantitative proteomics method; then searching for the mass spectrometry data by means of a non-limiting amino acid protein modification analysis method, and carrying out multivariable Gaussian mixture distribution clustering analysis to identify the non-coding amino acids in the sperm protein groups to the greatest possible extent; and finally, comparing the non-coding amino acids of normal and patient sperm protein groups to obtain non-coding amino acid sites of the protein related to severe oligoasthenospermia, wherein same are thereby used as molecular markers of severe

WO 2018/176808 A1

[见续页]

江(CHEN, Zijiang)；中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。 杨静鸣(YANG, Jingming)；中国山东省济南市历下区经十路17923号, Shandong 250061 (CN)。

(74) 代理人：济南圣达知识产权代理有限公司(JINAN SHENGDA INTELLECTUAL PROPERTY AGENCY CO., LTD)；中国山东省济南市历下区经十路17703号华特广场B308, Shandong 250061 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护)：AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：
— 包括国际检索报告(条约第21条(3))。

oligoasthenospermia. The method provides new diagnostic and therapeutic targets for severe oligoasthenospermia.

(57) 摘要：提供一种与重度少弱精子症相关的生物标志物的筛选方法。首先利用NanoHPLC-MS/MS质谱系统和非标记定量蛋白质组学方法对多组重度少弱精疾病的精子蛋白非编码氨基酸进行质谱分析；然后利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常和病人精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非编码氨基酸位点，从而将其作为重度少弱精症的分子标志物，为重度少弱精症提供了新的诊断和治疗靶点。

与重度少弱精子症相关的生物标志物的筛选与应用

技术领域

本发明涉及医学和分子诊断技术领域，具体涉及一种与重度少弱精子症相关的生物标志物的筛选与应用。

背景技术

不孕不育已成为全世界范围内的生殖健康问题，其中由于男性因素造成的不育大概占50%左右，且近年来呈上升的趋势。造成男性不育的主要原因是少精弱精症。根据世界卫生组织标准规定，如果 a 级精子数<25%，(a+b)级精子数<50%，且精子活率低于 60%的话，就可诊断为弱精子症。少精子症是指精液中的精子数目低于正常具有生育能力男性的一种病症，当男性的精子在每毫升低于 2 千万时，就为少精子症。

目前关于精子蛋白质组的研究有很多。Saraswat 等利用 UPLC-MS 的方法在人类精子中定量了 667 个蛋白，并且分析出了 20 个健康人和弱精症患者精子中的差异蛋白（Saraswat M, Joenvaara S, Jain T *et al.* Human Spermatozoa Quantitative Proteomic Signature Classifies Normo- and Asthenozoospermia. *Molecular & cellular proteomics : MCP*, 16(1), 57-72 (2017).）。最新更新的人类精子蛋白质组共 6198 个蛋白（Amaral A, Castillo J, Ramalho-Santos J, Oliva R. The combined human sperm proteome: cellular pathways and implications for basic and clinical science. *Human reproduction update*, 20(1), 40-62 (2014).）。Gaigai Wang 等利用高分辨质谱在人类精子中鉴定出 4675 个蛋白（Wang G, Guo Y, Zhou T *et al.* In-depth proteomic analysis of the human sperm reveals complex protein compositions. *Journal of proteomics*, 79, 114-122 (2013).）。络氨酸磷酸化对于精子的运动、获能、超激运动等过程重要作用。Chying-Chyuan Chan 等通过对 20 组正常人和弱精症患者的精子进行蛋白质组学分析发现有 12 种包括 TUBGCP2 在内的蛋白发生了过磷酸化（Chan CC, Shui HA, Wu CH *et al.* Motility and Protein Phosphorylation in Healthy and Asthenozoospermic Sperm. *Journal of proteome research*, 8(11), 5382-5386 (2009)）。

非编码氨基酸包括翻译后修饰和氨基酸突变，是调控蛋白功能和结构的重要方式，因此将疾病状态下异常的或者数量变化极大的非编码氨基酸作为疾病的生物标志物，进而用于诊断疾病的进程具有重要意义。但目前还未有关于非编码氨基酸作为与重度少弱精子疾病相关的生物标志物的报道。

发明内容

针对上述现有技术，本发明的目的提供一种与重度少弱精子症相关的生物标志物的筛选与应用。本发明首先利用 NanoHPLC-MS/MS 质谱系统和非标记定量蛋白质组学方法对多组重度少弱精疾病的精子蛋白非编码氨基酸进行了深度的质谱分析；然后利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常和病人精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非编码氨基酸位点，从而将其作为重度少弱精症的分子标志物。

为实现上述目的，本发明采用如下技术方案：

本发明的第一方面，提供了一种与重度少弱精子症相关的生物标志物的筛选方法，包括如下步骤：

（1）提取精子细胞全蛋白；

（2）将精子细胞全蛋白采用凝胶电泳分离，切胶酶解，对酶解后的肽段进行脱盐，制备得到样品；

（3）将步骤（2）的样品采用纳流液相色谱分离，经纳流液相色谱分离后的样品再进行质谱检测，采集质谱数据；

（4）利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常个体和重度少弱精个体精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非编码氨基酸位点，即为与重度少弱精子症相关的生物标志物。

步骤（1）中，提取精子细胞全蛋白采用的方法为：将精子样本采用 DPBS 洗涤，加入 RIPA 裂解液超声 1~2min，置于冰上孵育 30min 裂解，离心，取上清。

优选的，在 4℃的条件下离心，离心转速为 14,000 g，离心时间为 20min。

步骤（2）中，优选的，采用 10%聚丙烯酰氨凝胶电泳（SDS-PAGE）对蛋白进行分离。

步骤（2）中，优选的，采用 ziptip 对酶解后的肽段进行脱盐。

步骤（3）中，纳流液相色谱分离的色谱条件为：流动相 A：含有 0.1%甲酸的水，流动相 B：含有 0.1%甲酸的乙腈；纳流液相质谱分析系统为 Orbitrap Elite (Thermo Scientific)

洗脱条件为：0-100min，95-68%流动相 A，5-32%流动相 B；100-120min，68-20%流动相 A，32-80%流动相 B；120-150min，20%流动相 A，80%流动相 B；

流速为 300 nL/min。

步骤（3）中，质谱检测的条件为：350-1800 m/z 的全扫描，分辨率为 60,000 (m/z 200)。

二级图谱扫描时，活化时间为 10ms，隔离宽度为 2m/z；碎裂方式为诱导碰撞解离(collision-induced dissociation, CID)，归一化碰撞能量设定为 35%，动态排出时间为 90 s。

步骤（4）中，对质谱数据进行搜索的参数设置为：蛋白酶为胰蛋白酶，漏切位点设置为2，母离子质量偏差为 10ppm，碎片离子的质量偏差为 0.6Da，盲搜上限设为 1000，盲搜下限设为-200，蛋白 FDR 为 0.01；

选择肽段分数>200 和 FDR<0.01 的肽段作为 Wildcard Search™ 搜索到的未知修饰数据，组成质量变化的一维数据矩阵（-200Da-400Da），再将数据按照 1Da 的变化范围，0.5Da 为界限，分割成 601 个数据窗口。

步骤（4）中，多变量高斯混合分布聚类分析的方法为：针对每一个数据窗口，用 R 语言中的 mclust 程序包做高斯混合分布聚类分析，根据 BIC 取最优值，再对每一个峰进行合并分析，然后用高斯分布拟合每一个峰，确定峰值；聚类之后的每个峰中所包含的肽段位点数据，根据位点氨基酸分布，选择分布大于 5%的数据作为一类非编码氨基酸。

步骤（4）中，将正常个体和重度少弱精个体的非编码氨基酸按照其检测频率的 T 检验（$p<0.05$）和比值（ratio>2）进行筛选，从而得到差异非编码氨基酸。

上述筛选方法是用于获得生物标志物，且不是以获得疾病的诊断和治疗结果为目的；经上述筛选方法获得的生物标志物可用于重度少弱精子症的理论研究或者新药物的开发。

本发明的第二方面，提供根据上述筛选方法筛选得到的与重度少弱精子症相关的生物标志物，所述生物标志物包括但不限于：

AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸（标记为 S+79.96685；根据质量偏移值，确定该位置的丝氨酸发生了磷酸化修饰）；

AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺（标记为 N-113.05347）；

AKAP4 蛋白 186 位发生-114.04278 质量偏移的天门冬酰胺（标记为 N-114.04278）；

AKAP4 蛋白 617 位发生-17.02660 质量偏移的谷氨酰胺（标记为 Q-17.02660）；

AKAP4 蛋白 733 位发生+211.09682 质量偏移的赖氨酸（标记为 K+211.09682）；

ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸（标记为 K+42.01108；根据质量偏移值，确定该位置的赖氨酸发生了乙酰化修饰）；

COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸（标记为 K+42.01108；根据质量偏移值，确定该位置的赖氨酸发生了乙酰化修饰）；

GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸（标记为 T+79.96685；根据质量偏移值，确定该位置的苏氨酸发生了磷酸化修饰）；

和 KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸（标记为 S+79.96685；根据质量偏移值，确定该位置的丝氨酸发生了磷酸化修饰）。

本发明的第三方面，提供 AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

优选的，AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸还可以作为重度少弱精治疗的靶标，从而用于重度少弱精的治疗。

进一步的，本发明还提供 AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸作为生物标志物在制备重度少弱精的治疗药物中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸）的试剂。

本发明还提供一种治疗重度少弱精的药物，所述药物中含有能够使 AKAP3 蛋白 208 位丝氨酸进行磷酸化修饰的组分。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 AKAP3 蛋白 208 位丝氨酸发生+79.96685 质量偏移的频次，若检测频次小于 1.5 时，被判为少弱精患者。

本发明的第四方面，提供 AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺作为生物标志物在制备重度少弱精的诊断试剂中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺）的试剂。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 AKAP4 蛋白 186 位天门冬酰胺发生-113.05347 质量偏移的频次，检测频次小于 0.5 时，被判为少弱精患者。

本发明的第五方面，提供 AKAP4 蛋白 186 位发生-114.04278 质量偏移的天门冬酰胺作为生物标志物在制备重度少弱精的诊断试剂中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（AKAP4 蛋白 186 位发生-114.04278 质量偏移的天门冬酰胺）的试剂。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 AKAP4 蛋白 186 位天门冬酰胺发生-114.04278 质量偏移的频次，检测频次小于 0.5 时，被判为少弱精患者。

本发明的第六方面，提供 AKAP4 蛋白 617 位发生-17.02660 质量偏移的谷氨酰胺作为生物标志物在制备重度少弱精的诊断试剂中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（AKAP4 蛋白 617 位发生-17.02660 质量偏移的谷氨酰胺）的试剂。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 AKAP4 蛋白 617 位谷氨酰胺发生-17.02660 质量偏移的频次，检测频次小于 0.5 时，被判为少弱精患者。

本发明的第七方面，提供 AKAP4 蛋白 733 位发生+211.09682 质量偏移的赖氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（AKAP4 蛋白 733 位发生+211.09682 质量偏移的赖氨酸）的试剂。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 AKAP4 蛋白 733 位赖氨酸发生+211.09682 质量偏移的频次，检测频次小于 3.5 时，被判为少弱精患者。

本发明的第八方面，提供 ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

优选的，ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸还可以作为重度少弱精治疗的靶标，从而用于重度少弱精的治疗。

进一步的，本发明还提供 ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸作为生物标志物在制备重度少弱精的治疗药物中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸）的试剂。

本发明还提供一种治疗重度少弱精的药物，所述药物中含有能够使 ATP5A1 蛋白 531 位赖氨酸进行乙酰化修饰的组分。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 ATP5A1 蛋白 531 位赖氨酸发生+42.01108 质量偏移的频次，若检测频次小于 0.5 时，被判为少弱精患者。

本发明的第九方面，提供 COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

优选的，COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸还可以作为重度少弱精治疗的靶标，从而用于重度少弱精的治疗。

进一步的，本发明还提供 COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸作为生物标志物在制备重度少弱精的治疗药物中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸）的试剂。

本发明还提供一种治疗重度少弱精的药物，所述药物中含有能够使 COX4I1 蛋白 87 位赖氨酸进行乙酰化修饰的组分。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 COX4I1 蛋白 87 位赖氨酸发生+42.01108 质量偏移的频次，若检测频次小于 0.5 时，被判为少弱精患者。

本发明的第十方面，提供 GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

优选的，GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸还可以作为重度少弱精治疗的靶标，从而用于重度少弱精的治疗。

进一步的，本发明还提供 GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸作为生物标志物在制备重度少弱精的治疗药物中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸）的试剂。

本发明还提供一种治疗重度少弱精的药物，所述药物中含有能够使 GAPDHS 蛋白 64 位苏氨酸进行磷酸化修饰的组分。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 GAPDHS 蛋白 64 位苏氨酸发生+79.96685 质量偏移的频次，若检测频次小于 3.5 时，被判为少弱精患者。

本发明的第十一方面，提供 KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸作为生物标志物在制备重度少弱精的诊断试剂中的用途。

优选的，KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸还可以作为重度少弱精治疗的靶标，从而用于重度少弱精的治疗。

进一步的，本发明还提供 KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸作为生物标志物在制备重度少弱精的治疗药物中的用途。

本发明还提供一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物（KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸）的试剂。

本发明还提供一种治疗重度少弱精的药物，所述药物中含有能够使 KIAA1683 蛋白 692 位丝氨酸进行磷酸化修饰的组分。

本发明还提供一种重度少弱精的诊断方法，步骤为：检测待测样本 KIAA1683 蛋白 692 位丝氨酸发生+79.96685 质量偏移的频次，若检测频次小于 0.5 时，被判为少弱精患者。

本发明的有益效果：

（1）本发明首次建立了一种与重度少弱精子症相关的生物标志物的筛选方法，通过对大量样本精子蛋白的质谱数据进行分析处理，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常和病人精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非

编码氨基酸位点，从而将其作为重度少弱精症的分子标志物。

（2）本发明进一步的对上述筛选方法得到的生物标志物进行研究，发现可以通过上述生物标志物的检验频次来诊断重度少弱精症，为重度少弱精症提供了新的诊断和治疗靶点。

附图说明

构成本申请的一部分的说明书附图用来提供对本申请的进一步理解，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。

图 1：AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 检测频率的 ROC 曲线。

图 2：健康和少弱精样本中 AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 检测频率比较。

图 3：AKAP4 蛋白 186 位非编码氨基酸 N-113.05347 检测频率的 ROC 曲线。

图 4：健康和少弱精样本的 AKAP4 蛋白 186 位非编码氨基酸 N-113.05347 检测频率比较。

图 5：AKAP4 蛋白 186 位非编码氨基酸 N-114.04278 检测频率的 ROC 曲线。

图 6：健康和少弱精样本的 AKAP4 蛋白 186 位非编码氨基酸 N-114.04278 检测频率比较。

图 7：AKAP4 蛋白 617 位非编码氨基酸 Q-17.02660 检测频率的 ROC 曲线。

图 8：健康和少弱精样本的 AKAP4 蛋白 617 位非编码氨基酸 Q-17.02660 检测频率比较。

图 9：AKAP4 蛋白 733 位非编码氨基酸 K+211.09682 检测频率的 ROC 曲线。

图 10：健康和少弱精样本的非编码氨基酸 K+211.09682 检测频率比较。

图 11：ATP5A1 蛋白 531 位赖氨酸乙酰化修饰 K+42.01108 检测频率的 ROC 曲线。

图 12：健康和少弱精样本的 ATP5A1 蛋白 531 位赖氨酸乙酰化修饰 K+42.01108 检测频率比较。

图 13：COX4I1 蛋白 87 位赖氨酸乙酰化修饰 K+42.01108 检测频率的 ROC 曲线。

图 14：健康和少弱精样本的 COX4I1 蛋白 87 位赖氨酸乙酰化修饰 K+42.01108 检测频率比较。

图 15：GAPDHS 蛋白 64 位苏氨酸上的磷酸化修饰 T+79.96685 检测频率的 ROC 曲线。

图 16：健康和少弱精样本中 GAPDHS 蛋白 64 位苏氨酸上的磷酸化修饰 T+79.96685 检测频率比较。

图 17：KIAA1683 蛋白 692 位丝氨酸磷酸化修饰 S+79.96685 检测频率的 ROC 曲线。

图 18：健康和少弱精样本中 KIAA1683 蛋白 692 位丝氨酸磷酸化修饰 S+79.96685 检测频率比较。

具体实施方式

应该指出，以下详细说明都是例示性的，旨在对本申请提供进一步的说明。除非另有指明，本文使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解的相同含义。

需要注意的是，这里所使用的术语仅是为了描述具体实施方式，而非意图限制根据本申请的示例性实施方式。如在这里所使用的，除非上下文另外明确指出，否则单数形式也意图包括复数形式，此外，还应当理解的是，当在本说明书中使用术语“包含”和/或“包括”时，其指明存在特征、步骤、操作、器件、组件和/或它们的组合。

正如背景技术所介绍的，现有技术中还未有关于非编码氨基酸作为与重度少弱精子症相关的生物标志物的报道。基于此，本发明提出了一种与重度少弱精子症相关的生物标志物的筛选方法与应用。

在本申请的一种实施方案中，提出了一种与重度少弱精子症相关的生物标志物的筛选方法，包括如下步骤：

（1）提取精子细胞全蛋白；

（2）将精子细胞全蛋白采用凝胶电泳分离，切胶酶解，对酶解后的肽段进行脱盐，制备得到样品；

（3）将步骤（2）的样品采用纳流液相色谱分离，经纳流液相色谱分离后的样品再进行质谱检测，采集质谱数据；

（4）利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常个体和重度少弱精个体精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非编码氨基酸位点，即为与重度少弱精子症相关的生物标志物。

由于精子样本的易得性，精子成熟后其转录和翻译处于停滞状态，这也为我们在蛋白质水平上研究少弱精子症提供了方便，本申请首先利用 NanoHPLC-MS/MS 质谱系统和非标记定量蛋白质组学方法对多组重度少弱精疾病的精子蛋白非编码氨基酸进行了深度的质谱分析；然后利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸。最后将正常与患病组的非编码氨基酸按照其检测频率的 T 检验（$p<0.05$）和比值（ratio>2）进行筛选，从而得到差异非编码氨基酸。然后利用 SPSS 软件作出差异非编码氨基酸 ROC 曲线，并计算其曲线下面积（AUC），进而判断其诊断价值。

采用本申请的上述筛选方法，得到了系列与重度少弱精子症相关的生物标志物，具体如

下：

AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸（标记为 S+79.96685；根据质量偏移值，确定该位置的丝氨酸发生了磷酸化修饰）；

AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺（标记为 N-113.05347）；

AKAP4 蛋白 186 位发生-114.04278 质量偏移的天门冬酰胺（标记为 N-114.04278）；

AKAP4 蛋白 617 位发生-17.02660 质量偏移的谷氨酰胺（标记为 Q-17.02660）；

AKAP4 蛋白 733 位发生+211.09682 质量偏移的赖氨酸（标记为 K+211.09682）；

ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸（标记为 K+42.01108；根据质量偏移值，确定该位置的赖氨酸发生了乙酰化修饰）；

COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸（标记为 K+42.01108；根据质量偏移值，确定该位置的赖氨酸发生了乙酰化修饰）；

GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸（标记为 T+79.96685；根据质量偏移值，确定该位置的苏氨酸发生了磷酸化修饰）；

和 KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸（标记为 S+79.96685；根据质量偏移值，确定该位置的丝氨酸发生了磷酸化修饰）。

在本申请的另一种实施方案中，提出了一种用于重度少弱精诊断的试剂盒，所述试剂盒中包括特异性检测上述生物标志物的试剂。

通过对上述生物标志物进行检测，可以实现对重度少弱精症的诊断。

在本申请的另一种实施方案中，提出了一种治疗重度少弱精的药物，所述药物中含有能够使 AKAP3 蛋白 208 位丝氨酸、GAPDHS 蛋白 64 位苏氨酸或 KIAA1683 蛋白 692 位丝氨酸进行磷酸化修饰的组分，或者含有能够使 ATP5A1 蛋白 531 位赖氨酸或 COX4I1 蛋白 87 位赖氨酸进行乙酰化修饰的组分。

经研究发现，磷酸化修饰的 AKAP3 蛋白 208 位丝氨酸、GAPDHS 蛋白 64 位苏氨酸和 KIAA1683 蛋白 692 位丝氨酸在重度少弱精样本中显著性的下调，乙酰化修饰的 ATP5A1 蛋白 531 位赖氨酸或 COX4I1 蛋白 87 位赖氨酸也在重度少弱精样本中显著性的下调。由此可以合理预期，以上述生物标志物作为靶标，通过对多重度少弱精患者该靶标处的氨基酸进行磷酸化或乙酰化修饰，可以起到对重度少弱精的治疗作用。

为了使得本领域技术人员能够更加清楚地了解本申请的技术方案，以下将结合具体的实施例详细说明本申请的技术方案。

本发明实施例中所用的试验材料均为本领域常规的试验材料，均可通过商业渠道购买得

到。

实施例 1：与重度少弱精子症相关的生物标志物的筛选

具体筛选方法如下：

一、样本处理及实验分析

1. 精子细胞全蛋白的提取：等量的重度少弱精和正常精子样本分别用 DPBS 洗三次，加入等量 RIPA 裂解液超声 1~2min，置于冰上孵育 30min 裂解，4℃离心 14,000 g×20min 取上清。利用 Bradford 方法测定蛋白浓度。

2. 蛋白酶解：取约重度少弱精和正常精子样本各 150μg 精子蛋白，使用 10%聚丙烯酰氨凝胶电泳（SDS-PAGE）对蛋白进行分离，各分成 5 份进行切胶酶解。使用 ziptip 对肽段进行脱盐。

3. 质谱分析：纳流液相色谱分离：A 相：含有 0.1%甲酸的水；B 相：含有 0.1%甲酸的乙腈

每个样品分别用 13.5 μL A 相溶解,样品进样量为 4 μL，纳流液相质谱分析系统为 Orbitrap Elite (Thermo Scientific)。样品分离之前分别用 4 μL A 相平衡自制的预柱和分析柱。预柱和分析柱的规格分别为：预柱(4 cm×150 μm I.D., C18 填料粒径 5 μm, 100 Å)，分析柱(30 cm× 75 μm I.D., C18 填料填充，粒径 3μm, 100 Å, Dr. Maisch GmbH, Germany)。平衡之后样品在 A 相的带动下首先载样于预柱，然后在不同梯度下进行液相分离。150 min 色谱梯度变化如下：5-32% 流动相 B 100 min； 32-80%流动相 B，20 min； 80%流动相 B，30 min。流速始终保持在 300 nL/min。经过纳流液相分离的样品直接进入 ESI 离子喷雾源并进入 Orbitrap Elite 质谱仪中进行质谱检测。

质谱数据采集：350 - 1800 m/z 的全扫描，分辨率为 60,000 (m/z 200)。二级图谱扫描时，活化时间为 10ms，隔离宽度为 2m/z。碎裂方式为诱导碰撞解离(collision-induced dissociation, CID)，归一化碰撞能量设定为 35%，动态排出时间为 90 s。

二、质谱数据分析

Byonic 分析：为了鉴定出精子蛋白的非编码氨基酸，我们用 Byonic™分析 21 对正常与重度少弱精患者的精蛋白质谱数据。搜索参数如下：蛋白酶为胰蛋白酶，漏切位点设置为 2，母离子质量偏差为 10ppm，碎片离子的质量偏差为 0.6Da,盲搜上限设为 1000，盲搜下限设为 -200.蛋白 FDR 为 0.01。

选择肽段分数>200 和 FDR<0.01 的肽段作为 Wildcard Search™ 搜索到的未知修饰数据，组成质量变化的一维数据矩阵（-200Da-400Da），再将数据按照 1Da 的变化范围，0.5Da 为界

限，分割成 601 个数据窗口。针对每一个数据窗口，用 R 语言中的 mclust 程序包做高斯混合分布聚类分析，根据 BIC 取最优值，再对每一个峰进行合并分析，然后用高斯分布拟合每一个峰，确定峰值。聚类之后的每个峰中所包含的肽段位点数据，根据位点氨基酸分布，选择分布大于 5%的数据作为一类非编码氨基酸。

将正常与患病组的非编码氨基酸按照其检测频率的 T 检验（$p<0.05$）和比值（ratio>2）进行筛选，从而得到差异非编码氨基酸。然后利用 SPSS 软件作出差异非编码氨基酸 ROC 曲线，并计算其曲线下面积（AUC），进而判断其诊断价值。

三、实验结果：

经质谱数据分析和将正常与患病组的非编码氨基酸比较，从而得到 9 个差异非编码氨基酸，可以作为与重度少弱精子症相关的生物标志物，具体如下：

1．AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸（标记为 S+79.96685；根据质量偏移值，确定该位置的丝氨酸发生了磷酸化修饰）

我们发现 AKAP3 蛋白 208 位丝氨酸上发生了磷酸化修饰 S+79.96685，经过比较发现这一磷酸化修饰在重度少弱精样本显著性下调了 10.7 倍，p 值为 7.49E-15<0.05。

为评价 AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 检测频率对重度少弱精的诊断效能，本发明采用了 ROC 曲线分析，AUC 为 ROC 曲线下的面积，是最常用的评价 ROC 曲线特征的参数，是重要的试验准确度指标。若 AUC 在 0.7 以下，则表示诊断的准确率较低；AUC 在 0.7 以上，则可以满足临床诊断的要求。

图 1 为 AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 检测频率的 ROC 曲线，ROC 分析显示这一磷酸化修饰的 AUC 为 0.856>0.7，说明具有较好的诊断效果，即 AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 可以作为重度少弱精的诊断标志物。

在检验频次为 1.5 时，灵敏度为 73%，特异度为 88.9%。当进行个体检测时，检测频次小于 1.5 时，被判为少弱精患者（假阳性率为 11.1%）。

健康和少弱精样本中 AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 检测频率比较结果见图 2，由图 2 可以看出这一非编码氨基酸在健康人样本中平均发生了 4.6 次而在病理样本中发生了 0.4 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，AKAP3 蛋白 208 位丝氨酸上的磷酸化修饰 S+79.96685 可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

2．AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺（标记为 N-113.05347）

经过质谱数据分析，我们发现 AKAP4 蛋白的 186 位的天门冬酰胺有-113.05347 的质量偏移（N-113.05347），经过比较发现这一非编码氨基酸 N-113.05347 在重度少弱精样本显著性下调了 9.2 倍，p 值为 4.60E-13<0.05。

图 3 为 AKAP4 蛋白 186 位非编码氨基酸 N-113.05347 检测频率的 ROC 曲线，ROC 分析显示这一非编码氨基酸 N-113.05347 的 AUC 为 0.852>0.7，说明具有较好的诊断效果。在检验频次为 1.5 时，灵敏度为 65.1%，特异度为 93.7%。当进行个体检测时，检测频次小于 1.5 时，被判为少弱精患者（假阳性率为 6.3%）。

图 4 比较了非编码氨基酸 N-113.05347 在健康和少弱精样本的检测频率，可以看出这一非编码氨基酸在健康人样本中平均发生了 2.9 次而在病理样本中发生了 0.3 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，AKAP4 蛋白的 186 位点的天门冬酰胺 N-113.05347 这一非编码氨基酸可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

3．AKAP4蛋白186位发生-114.04278质量偏移的天门冬酰胺（标记为N-114.04278）

经过质谱数据分析，我们发现 AKAP4 蛋白的 186 位的天门冬酰胺有-114.04278 的质量偏移（N-114.04278），经过比较发现这一非编码氨基酸 N-114.04278 在重度少弱精样本显著性下调了 7.3 倍，p 值为 1.11E-11<0.05。

图 5 为 AKAP4 蛋白 186 位非编码氨基酸 N-114.04278 检测频率的 ROC 曲线，ROC 分析显示这一非编码氨基酸 N-114.04278 的 AUC 为 0.817>0.7，说明具有较好的诊断效果。在检验频次为 0.5 时，灵敏度为 74.6%，特异度为 84.1%。当进行个体检测时，检测频次小于 0.5 时，被判为少弱精患者（假阳性率为 15.9%）。

健康和少弱精样本的 AKAP4 蛋白 186 位非编码氨基酸 N-114.04278 检测频率比较见图 6，由图 6 可以看出这一非编码氨基酸在健康人样本中平均发生了 1.6 次而在病理样本中发生了 0.2 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，AKAP4蛋白的186位点的天门冬酰胺N-114.04278这一非编码氨基酸可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

4．AKAP4蛋白617位发生-17.02660质量偏移的谷氨酰胺（标记为Q-17.02660）

经过质谱数据分析，我们发现 AKAP4 蛋白的 617 位的谷氨酰胺有-17.02660 的质量偏移（Q-17.02660），经过比较发现这一非编码氨基酸 Q-17.02660 在重度少弱精样本显著性下调了

4.8 倍，p 值为 2.21E-09<0.05。

图 7 为 AKAP4 蛋白 617 位非编码氨基酸 Q-17.02660 检测频率的 ROC 曲线，ROC 分析显示这一非编码氨基酸 Q-17.02660 的 AUC 为 0.802>0.7，说明具有较好的诊断效果。在检验频次为 0.5 时，灵敏度为 77.8%，特异度为 79.4%。当进行个体检测时，检测频次小于 0.5 时，被判为少弱精患者（假阳性率为 20.6%）。

健康和少弱精样本的 AKAP4 蛋白 617 位非编码氨基酸 Q-17.02660 检测频率比较见图 8，由图 8 可以看出这一非编码氨基酸在健康人样本中平均发生了 2.7 次而在病理样本中发生了 0.6 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，AKAP4蛋白的617位点的谷氨酰胺Q-17.02660这一非编码氨基酸可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

5．AKAP4蛋白733位发生+211.09682质量偏移的赖氨酸（标记为K+211.09682）

经过质谱数据分析，我们发现 AKAP4 蛋白的 733 位的赖氨酸有 211.09682 的质量偏移（K+211.09682），经过比较发现这一非编码氨基酸 K+211.09682 在重度少弱精样本显著性下调了 4.4 倍，p 值为 5.79E-11<0.05。

图 9 为 AKAP4 蛋白 733 位非编码氨基酸 K+211.09682 检测频率的 ROC 曲线，ROC 分析显示这一非编码氨基酸 K+211.09682 的 AUC 为 0.804>0.7，说明具有较好的诊断效果。在检验频次为 3.5 时，灵敏度为 74.6%，特异度为 71.4%。当进行个体检测时，检测频次小于 3.5 时，被判为少弱精患者（假阳性率为 28.6%）。

健康和少弱精样本的非编码氨基酸 K+211.09682 检测频率比较见图 10，由图 10 可以看出这一非编码氨基酸在健康人样本中平均发生了 14 次而在病理样本中发生了 3 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，AKAP4蛋白的733位点的赖氨酸K+211.09682这一非编码氨基酸可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

6．ATP5A1蛋白531位发生+42.01108质量偏移的赖氨酸

经过质谱数据分析，我们发现 ATP5A1 蛋白 531 位赖氨酸上发生了乙酰化修饰 K+42.01108，经过比较发现这一乙酰化修饰在重度少弱精样本显著性下调了 10.5 倍，p 值为 3.48E-16<0.05。

图 11 为 ATP5A1 蛋白 531 位赖氨酸乙酰化修饰 K+42.01108 检测频率的 ROC 曲线，ROC

分析显示这一乙酰化修饰的 AUC 为 0.848>0.7，说明具有较好的诊断效果。在检验频次为 0.5 时，灵敏度为 76.2%，特异度为 85.7%。当进行个体检测时，检测频次小于 0.5 时，被判为少弱精患者（假阳性率为 14.3%）。

健康和少弱精样本的 ATP5A1 蛋白 531 位赖氨酸乙酰化修饰 K+42.01108 检测频率比较见图 12，可以看出这一非编码氨基酸在健康人样本中平均发生了 1.7 次而在病理样本中发生了 0.2 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，ATP5A1蛋白531位赖氨酸上的乙酰化修饰K+42.01108可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

7．COX4I1蛋白87位发生+42.01108质量偏移的赖氨酸（标记为K+42.01108）

经过质谱数据分析，我们发现COX4I1 蛋白 87 位赖氨酸上发生了乙酰化修饰 K+42.01108，经过比较发现这一乙酰化修饰在重度少弱精样本显著性下调了 11.3 倍，p 值为 5.06E-13<0.05。

图 13 为 COX4I1 蛋白 87 位赖氨酸乙酰化修饰 K+42.01108 检测频率的 ROC 曲线，ROC 分析显示这一乙酰化修饰的 AUC 为 0.803>0.7，说明具有较好的诊断效果。在检验频次为 0.5 时，灵敏度为 66.7%，特异度为 95.2%。当进行个体检测时，检测频次小于 0.5 时，被判为少弱精患者（假阳性率为 4.8%）。

图 14 为健康和少弱精样本的 COX4I1 蛋白 87 位赖氨酸乙酰化修饰 K+42.01108 检测频率比较，由图 14 可以看出这一非编码氨基酸在健康人样本中平均发生了 1.1 次而在病理样本中发生了 0.1 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，COX4I1蛋白87位赖氨酸上的乙酰化修饰K+42.01108可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

8．GAPDHS蛋白64位发生+79.96685质量偏移的苏氨酸（标记为T+79.96685）

经过质谱数据分析，我们发现 GAPDHS 蛋白的 64 位苏氨酸上发生了磷酸化修饰 T+79.96685，经过比较发现这一磷酸化修饰在重度少弱精样本显著性下调了 4.1 倍，p 值为 1.82E-14<0.05。

图 15 为 GAPDHS 蛋白 64 位苏氨酸上的磷酸化修饰 T+79.96685 检测频率的 ROC 曲线，ROC 分析显示这一磷酸化修饰的 AUC 为 0.868>0.7，说明具有较好的诊断效果。在检验频次为 3.5 时，灵敏度为 71.4%，特异度为 92.1%。当进行个体检测时，检测频次小于 3.5 时，被判为少弱精患者（假阳性率为 7.9%）。

图 16 为健康和少弱精样本中 GAPDHS 蛋白 64 位苏氨酸上的磷酸化修饰 T+79.96685 检测频率比较，由图 16 可以看出这一非编码氨基酸在健康人样本中平均发生了 5.8 次而在病理样本中发生了 1.4 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，GAPDHS蛋白64位苏氨酸上的磷酸化修饰T+79.96685可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

9．KIAA1683蛋白692位发生+79.96685质量偏移的丝氨酸（标记为S+79.96685）

经过质谱数据分析，我们发现 KIAA1683 蛋白 692 位丝氨酸上发生了磷酸化修饰 S+79.96685，经过比较发现这一磷酸化修饰在重度少弱精样本显著性下调了 22 倍，p 值为 2.73E-10<0.05。

图 17 为 KIAA1683 蛋白 692 位丝氨酸磷酸化修饰 S+79.96685 检测频率的 ROC 曲线，ROC 分析显示这一磷酸化修饰的 AUC 为 0.808>0.7，说明具有较好的诊断效果。在检验频次为 0.5 时，灵敏度为 65.1%，特异度为 95.2%。当进行个体检测时，检测频次小于 0.5 时，被判为少弱精患者（假阳性率为 4.8%）。

图 18 为健康和少弱精样本中 KIAA1683 蛋白 692 位丝氨酸磷酸化修饰 S+79.96685 检测频率比较，由图 18 可以看出这一非编码氨基酸在健康人样本中平均发生了 2.1 次而在病理样本中发生了 0.1 次（图中实线），并且其中位数（图中虚线）相差较远，说明在少弱精样本中这一非编码氨基酸有较大程度地丢失。

鉴于上述结果，KIAA1683蛋白692位丝氨酸上的磷酸化修饰S+79.96685可以作为少弱精子症的潜在生物标志物，从而对这一病症进行预测。

实施例2：临床检测验证

以 4 例健康样本、8 例已临床确诊的重度少弱精样本作为研究对象进行验证，分别检测上述样本的 AKAP3 蛋白 208 位发生+79.96685 质量偏移的丝氨酸、AKAP4 蛋白 186 位发生-113.05347 质量偏移的天门冬酰胺、AKAP4 蛋白 186 位发生-114.04278 质量偏移的天门冬酰胺、AKAP4 蛋白 617 位发生-17.02660 质量偏移的谷氨酰胺、AKAP4 蛋白 733 位发生+211.09682 质量偏移的赖氨酸、ATP5A1 蛋白 531 位发生+42.01108 质量偏移的赖氨酸、COX4I1 蛋白 87 位发生+42.01108 质量偏移的赖氨酸、GAPDHS 蛋白 64 位发生+79.96685 质量偏移的苏氨酸和 KIAA1683 蛋白 692 位发生+79.96685 质量偏移的丝氨酸各自的检测频次，并按实施例 1 中各生物标志物进行个体检测时的判断标准对待测样本进行诊断。

结果表明：分别以上述各生物标志物进行单独诊断时，诊断结果与已知结果一致。说明

本发明筛选得到的 9 个生物标志物可以各自作为重度少弱精的诊断标志物。

以上所述仅为本申请的优选实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

1. 一种与重度少弱精子症相关的生物标志物的筛选方法，其特征在于，包括如下步骤：

（1）提取精子细胞全蛋白；

（2）将精子细胞全蛋白采用凝胶电泳分离，切胶酶解，对酶解后的肽段进行脱盐，制备得到样品；

（3）将步骤（2）的样品采用纳流液相色谱分离，经纳流液相色谱分离后的样品再进行质谱检测，采集质谱数据；

（4）利用非限定氨基酸蛋白质修饰分析方法对质谱数据进行搜索，再经过多变量高斯混合分布聚类分析，尽可能大量的鉴定出精子蛋白组中非编码氨基酸；最后通过正常个体和重度少弱精个体精子蛋白组中非编码氨基酸的比较，得到与重度少弱精症相关的蛋白非编码氨基酸位点，即为诊断和/或治疗重度少弱精子疾病的生物标志物。

2. 如权利要求 1 所述的筛选方法，其特征在于，步骤（1）中，提取精子细胞全蛋白采用的方法为：将精子样本采用 DPBS 洗涤，加入 RIPA 裂解液超声 1~2min，置于冰上孵育 30min 裂解，离心，取上清；

3. 如权利要求 2 所述的筛选方法，其特征在于，在 4℃的条件下离心，离心转速为 14,000 g，离心时间为 20min。

4. 如权利要求 1 所述的筛选方法，其特征在于，步骤（2）中，采用 10%聚丙烯酰氨凝胶电泳对蛋白进行分离。

5. 如权利要求 1 所述的筛选方法，其特征在于，步骤（2）中，采用 ziptip 对酶解后的肽段进行脱盐。

6. 如权利要求 1 所述的筛选方法，其特征在于，步骤（3）中，纳流液相色谱分离的色谱条件为：流动相 A：含有 0.1%甲酸的水，流动相 B：含有 0.1%甲酸的乙腈；

洗脱条件为：0-100min，95-68%流动相 A，5-32%流动相 B；100-120min，68-20%流动相 A，32-80%流动相 B；120-150min，20%流动相 A，80%流动相 B。

7. 如权利要求 1 所述的筛选方法，其特征在于，步骤（3）中，质谱检测的条件为：350-1800 m/z 的全扫描，分辨率为 60,000 (m/z 200)。二级图谱扫描时，活化时间为 10ms，隔离宽度为 2m/z；碎裂方式为诱导碰撞解离，归一化碰撞能量设定为 35%，动态排出时间为 90 s。

8. 如权利要求 1 所述的筛选方法，其特征在于，步骤（4）中，对质谱数据进行搜索的参数设置为：蛋白酶为胰蛋白酶，漏切位点设置为 2，母离子质量偏差为 10ppm，碎片离子的质量偏差为 0.6Da，盲搜上限设为 1000，盲搜下限设为-200，蛋白 FDR 为 0.01；

选择肽段分数>200 和 FDR<0.01 的肽段作为 Wildcard Search™ 搜索到的未知修饰数据，

组成质量变化的一维数据矩阵（-200Da-400Da），再将数据按照 1Da 的变化范围，0.5Da 为界限，分割成 601 个数据窗口。

9. 如权利要求 1 所述的筛选方法，其特征在于，步骤（4）中，多变量高斯混合分布聚类分析的方法为：针对每一个数据窗口，用 R 语言中的 mclust 程序包做高斯混合分布聚类分析，根据 BIC 取最优值，再对每一个峰进行合并分析，然后用高斯分布拟合每一个峰，确定峰值；聚类之后的每个峰中所包含的肽段位点数据，根据位点氨基酸分布，选择分布大于 5% 的数据作为一类非编码氨基酸。

10．权利要求 1-9 任一项所述的筛选方法筛选得到的与重度少弱精子症相关的生物标志物。

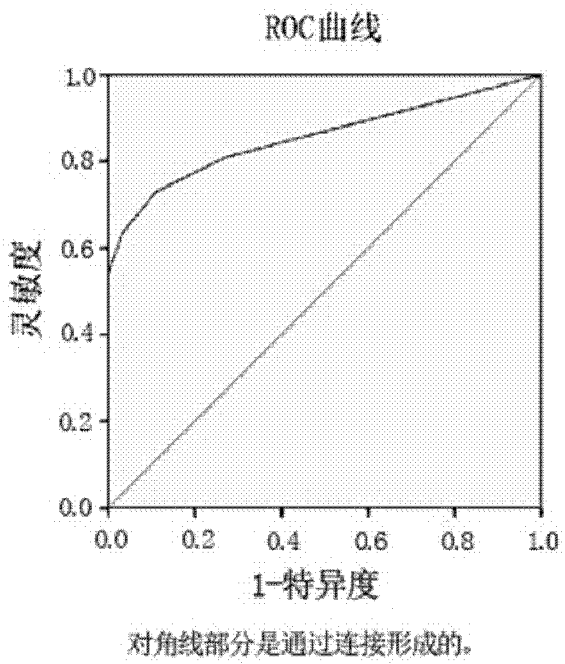
ROC曲线
灵敏度
1-特异度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
对角线部分是通过连接形成的。

图 1

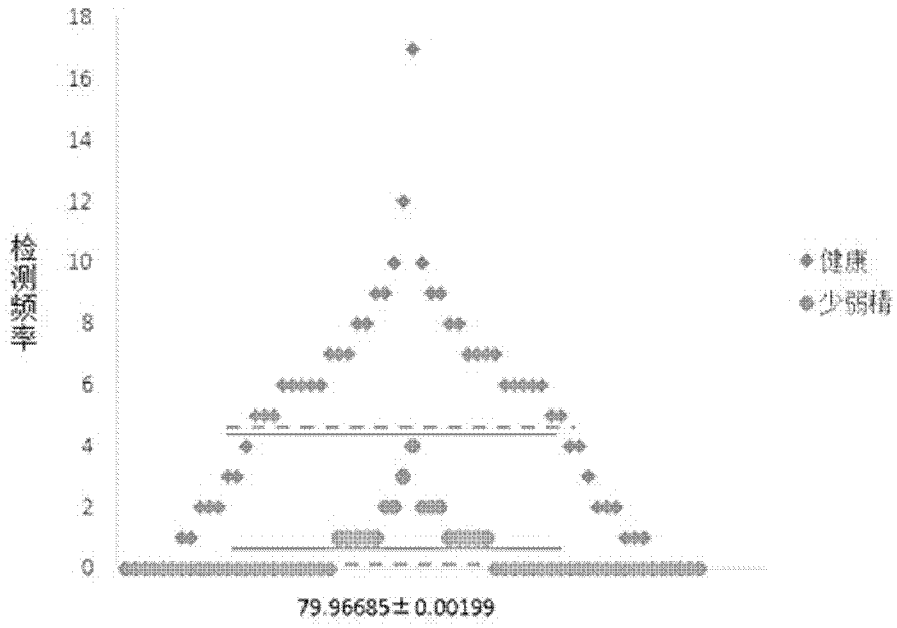
检测频率
18
16
14
12
10
8
6
4
2
0
健康
少弱精
79.96685±0.00199

图 2

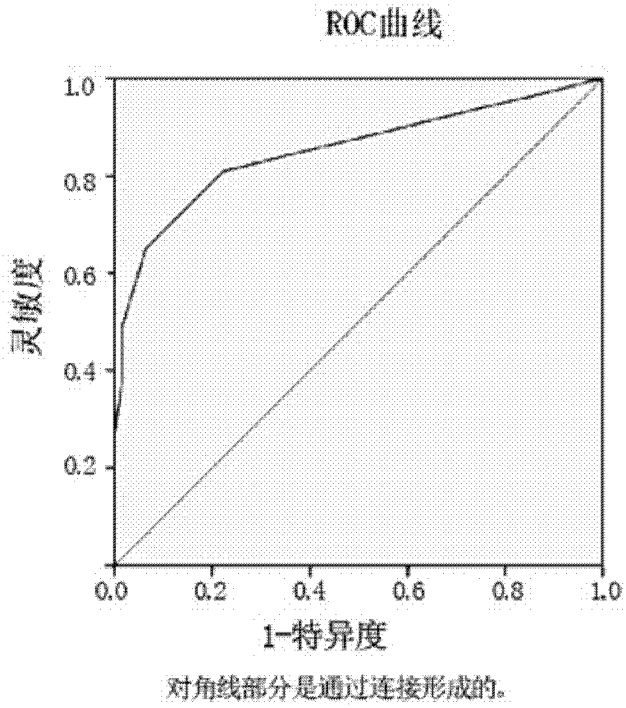
ROC曲线
灵敏度
1-特异度
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
对角线部分是通过连接形成的。

图 3

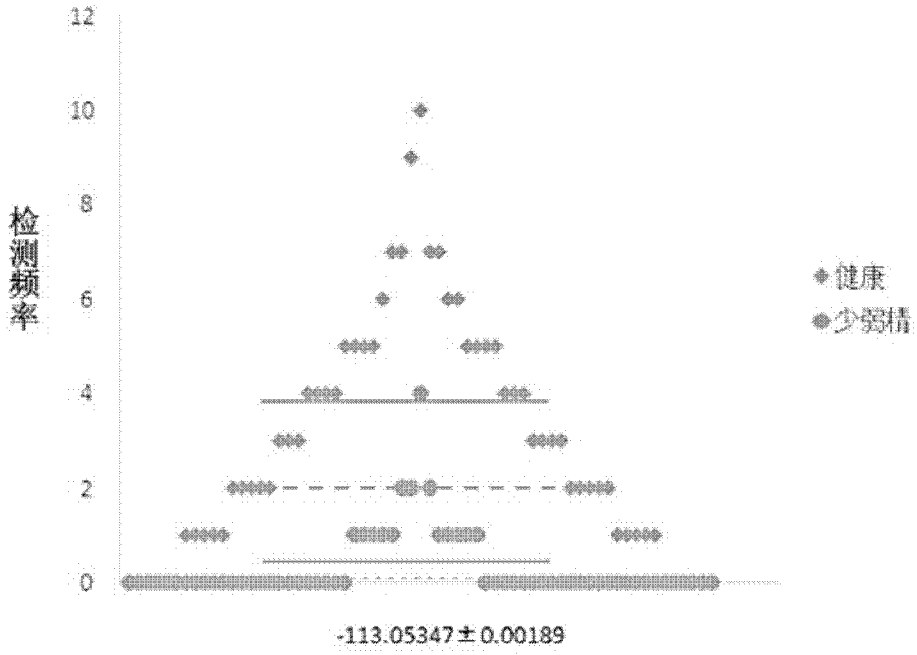
检测频率
12
10
8
6
4
2
0
健康
少弱精
-113.05347±0.00189

图 4

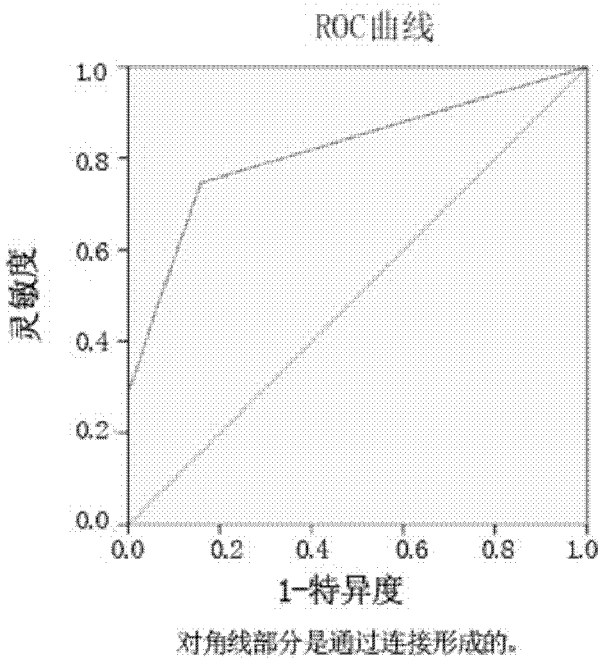
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 5

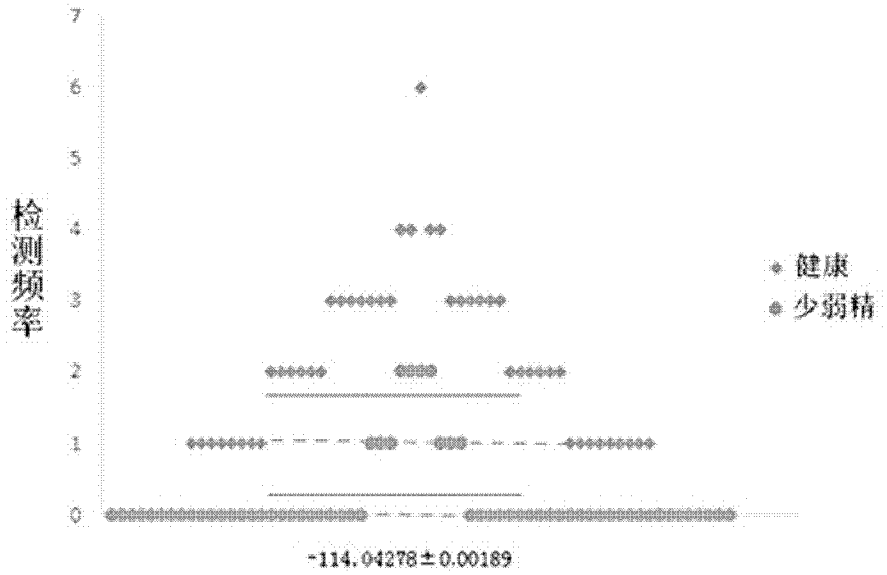
检测频率
7
6
5
4
3
2
1
0
健康
少弱精
-114.04278±0.00189

图 6

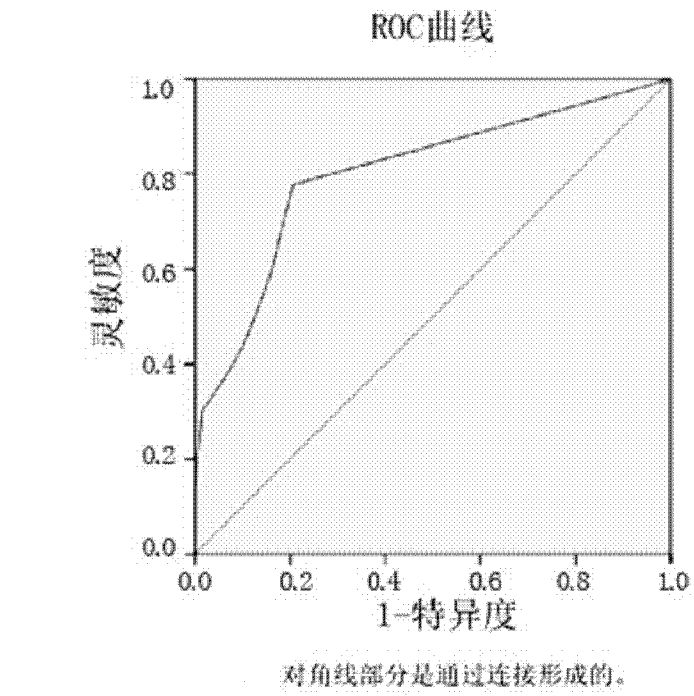
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 7

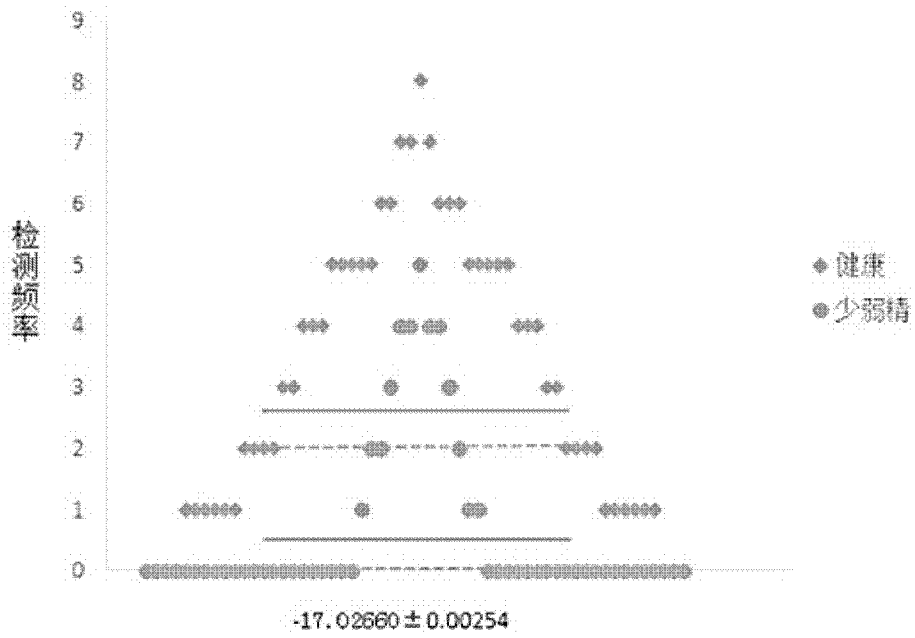
检测频率
9
8
7
6
5
4
3
2
1
0
健康
少弱精
-17.02660±0.00254

图 8

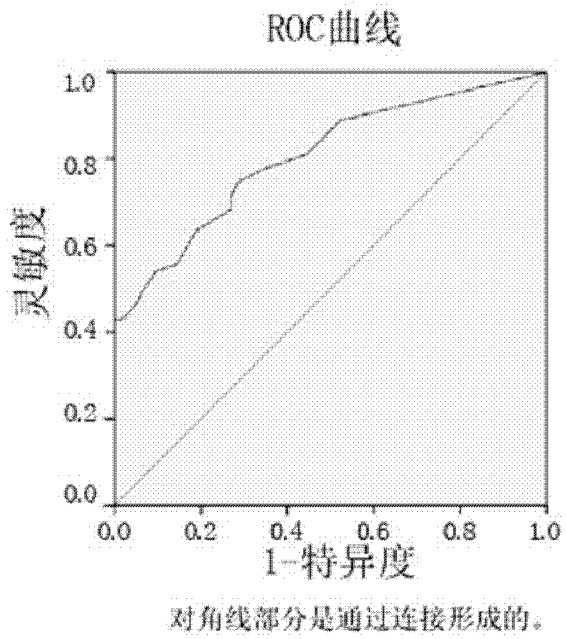
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 9

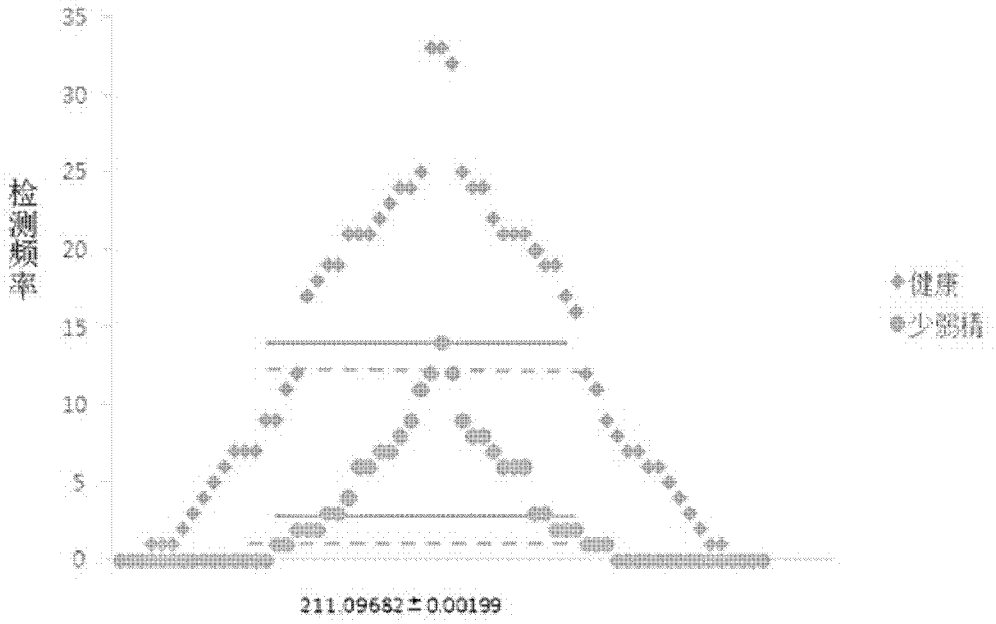
检测频率
35
30
25
20
15
10
5
0
211.09682±0.00199
健康
少弱精

图 10

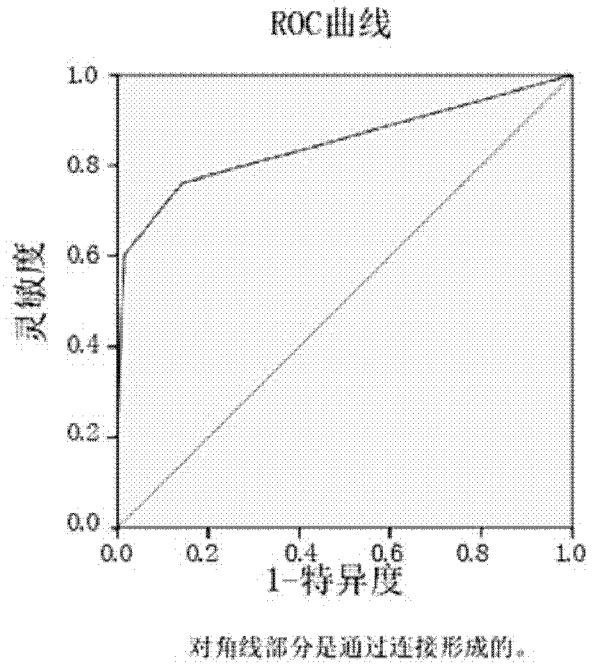
ROC曲线
灵敏度
1-特异度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
对角线部分是通过连接形成的。

图 11

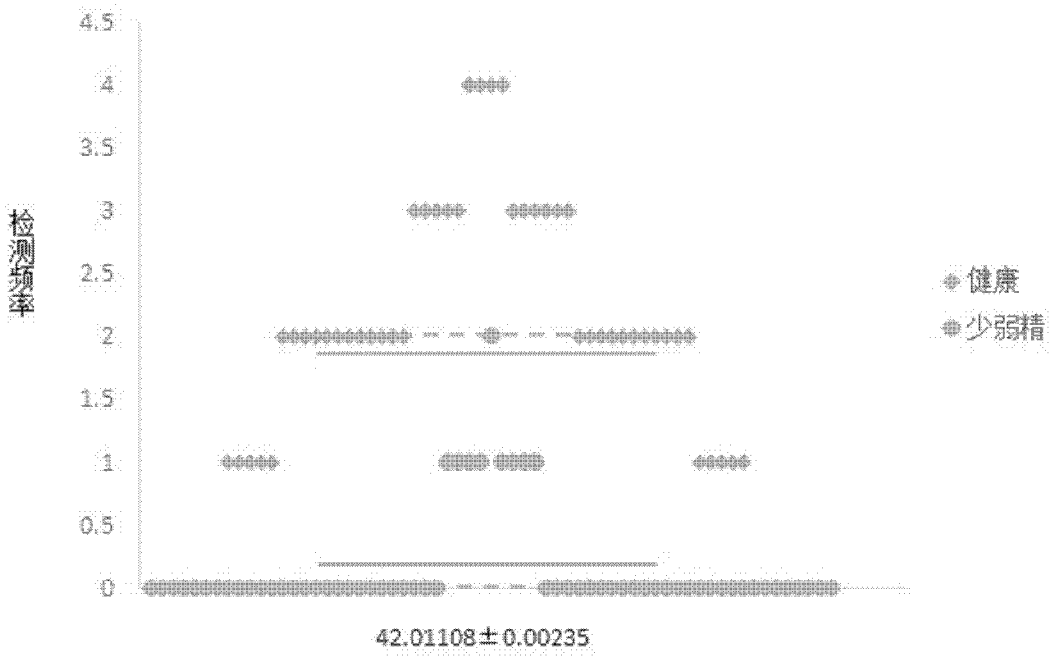
检测频率
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
健康
少弱精
42.01108±0.00235

图 12

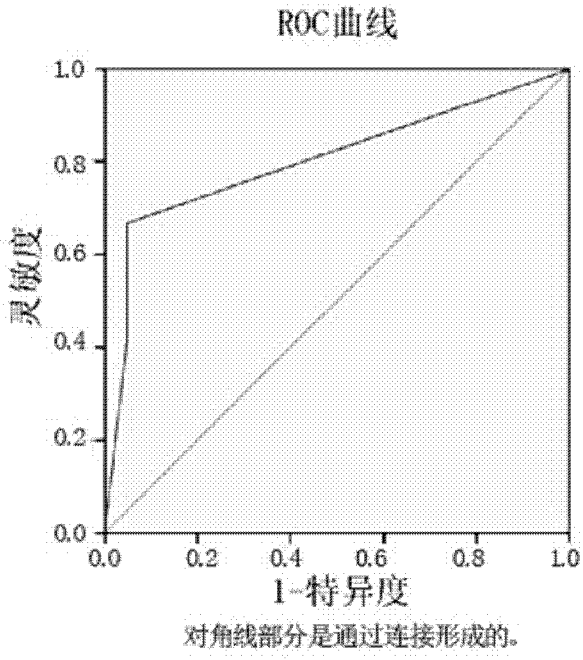
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 13

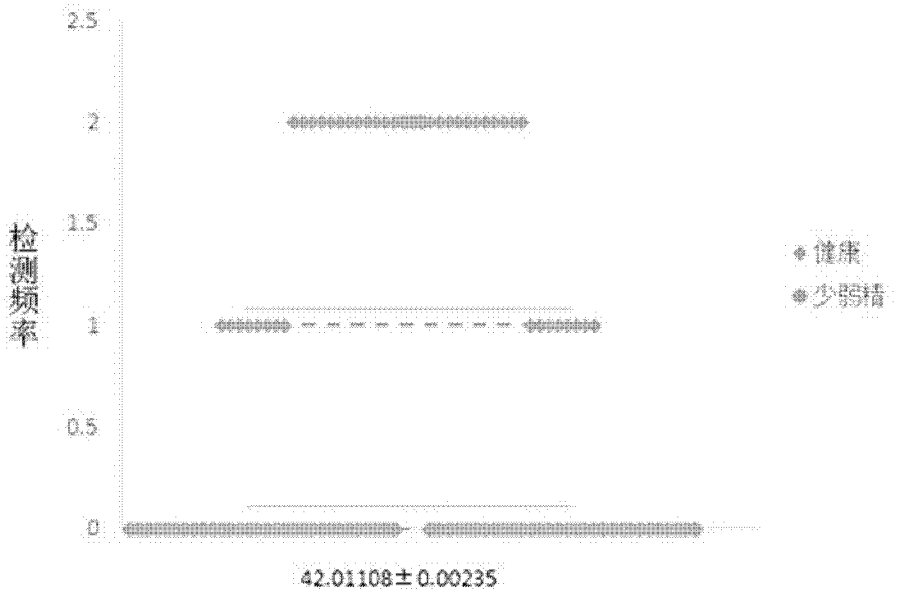
检测频率
2.5
2
1.5
1
0.5
0
健康
少弱精
42.01108±0.00235

图 14

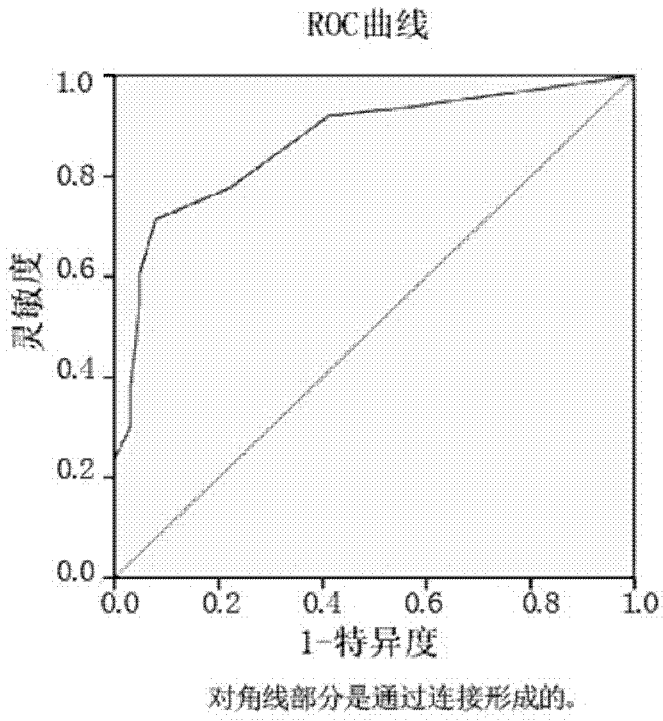
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 15

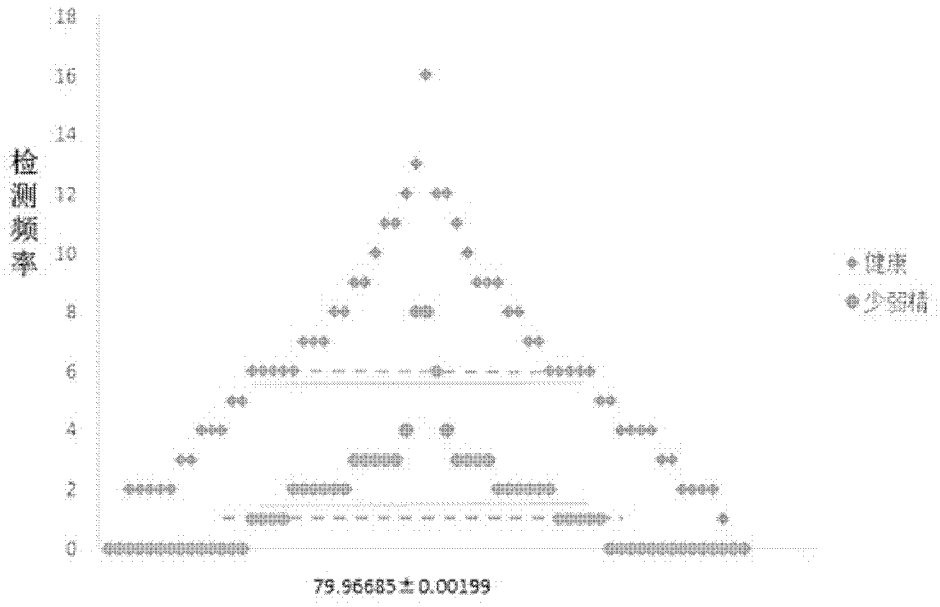
检测频率
18
16
14
12
10
8
6
4
2
0
健康
少弱精
79.96685±0.00199

图 16

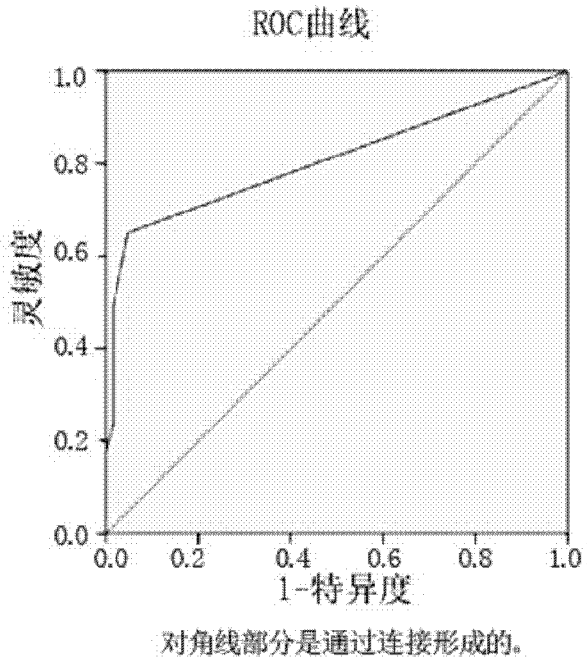
ROC曲线
灵敏度
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1-特异度
对角线部分是通过连接形成的。

图 17

检测频率
9
8
7
6
5
4
3
2
1
0
健康
少弱精
79.96685±0.00199

图 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/106915

A. CLASSIFICATION OF SUBJECT MATTER

G01N 30/88 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, DWPI, SIPOABS, ISI web of knowledge, PubMed, WANFANG, CNKI: 质谱, spectrometry, mass spectra, mass spectrum, MS, 弱精, 少精, 精子, 精液, asthenozoospermia, asthenospermia, oligospermia, oligospermatism, oligozoospermia, sperm, semen, 磷酸化, phosphorylation, 乙酰化, acetylization, l翻译后修饰, posttranslational modification, 标志物, 标记物, indicator, biomarker

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106872630 A (SHANDONG UNIVERSITY et al.), 20 June 2017 (20.06.2017), entire document	1-10
PX	CN 107024553 A (SHANDONG UNIVERSITY), 08 August 2017 (08.08.2017), entire document	1-10
PX	CN 106932597 A (SHANDONG UNIVERSITY), 07 July 2017 (07.07.2017), entire document	1-10
PX	CN 106996979 A (SHANDONG UNIVERSITY), 01 August 2017 (01.08.2017), entire document	1-10
PX	CN 107015005 A (SHANDONG UNIVERSITY), 04 August 2017 (04.08.2017), entire document	1-10
PX	CN 107037172 A (SHANDONG UNIVERSITY), 11 August 2017 (11.08.2017), entire document	1-10
PX	CN 106996981 A (SHANDONG UNIVERSITY), 01 August 2017 (01.08.2017), entire document	1-10

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 January 2018	10 January 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer CAO, Kou Telephone No. (86-10) 62413565

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/106915

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 106996980 A (SHANDONG UNIVERSITY), 01 August 2017 (01.08.2017), entire document	1-10
PX	CN 106990177 A (SHANDONG UNIVERSITY), 28 July 2017 (28.07.2017), entire document	1-10
X	PARTE, P.P. et al., "Sperm Phosphoproteome Profiling by Ultra Performance Liquid Chromatography Followed by Data Independent Analysis (LC-MSE) Reveals Altered Proteomic Signatures in Asthenozoospermia", Journal of Proteomics, 75(18), 13 July 2012 (13.07.2012), pages 5861-5871	1-10
A	MCDONALD, W.H. et al., "Shotgun Proteomics and Biomarker Discovery", Disease Markers,18(2), 31 December 2002 (31.12.2002), pages 99-105	1-10
A	CHAN, C.C. et al., "Motility and Protein Phosphorylation in Healthy and Astheno-zoospermic Sperm", Journal of Proteome Research, 8(11), 13 August 2009 (13.08.2009), pages 5382-5386	1-10
A	夏西洋. "睾丸蛋白质组学数据翻译后修饰再挖掘", 中国优秀硕士学位论文全文数据库医药卫生科技辑, no. 04, 15 April 2016 (15.04.2016), E059-34, (XIA, Xiyang; "Proteomic Analysis of Novel Post-translational Modifications in Human and Mouse Testes", Medicine & Public Health, China Master's Theses Full-Text Database)	1-10
A	CN 101477089 A (INSTITUTE OF COMPUTING TECHNOLOGY, CHINESE ACADEMY OF SCIENCES), 08 July 2009 (08.07.5009), claims 1-17	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/106915

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 106872630 A	20 June 2017	None	
CN 107024553 A	08 August 2017	None	
CN 106932597 A	07 July 2017	None	
CN 106996979 A	01 August 2017	None	
CN 107015005 A	04 August 2017	None	
CN 107037172 A	11 August 2017	None	
CN 106996981 A	01 August 2017	None	
CN 106996980 A	01 August 2017	None	
CN 106990177 A	28 July 2017	None	
CN 101477089 A	08 July 2009	None	

国际检索报告

国际申请号
PCT/CN2017/106915

A. 主题的分类

G01N 30/88(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G01N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

CNABS，DWPI，SIPOABS，ISI web of knowledge，PubMed，万方，CNKI：质谱，spectrometry，mass spectra，mass spectrum，MS，弱精，少精，精子，精液，asthenozoospermia，asthenospermia，oligospermia，oligospermatism，oligozoospermia，sperm，semen，磷酸化，phosphorylation，乙酰化，acetylization，翻译后修饰，post-translational modification，标志物，标记物，indicator，biomarker

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 106872630 A (山东大学 等) 2017年 6月 20日 (2017-06-20) 全文	1-10
PX	CN 107024553 A (山东大学) 2017年 8月 8日 (2017-08-08) 全文	1-10
PX	CN 106932597 A (山东大学) 2017年 7月 7日 (2017-07-07) 全文	1-10
PX	CN 106996979 A (山东大学) 2017年 8月 1日 (2017-08-01) 全文	1-10
PX	CN 107015005 A (山东大学) 2017年 8月 4日 (2017-08-04) 全文	1-10
PX	CN 107037172 A (山东大学) 2017年 8月 11日 (2017-08-11) 全文	1-10
PX	CN 106996981 A (山东大学) 2017年 8月 1日 (2017-08-01) 全文	1-10

☑ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2018年 1月 2日	2018年 1月 10日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 曹扣 电话号码 (86-10)62413565

表 PCT/ISA/210 (第2页) (2009年7月)

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 106996980 A (山东大学) 2017年 8月 1日 (2017-08-01) 全文	1-10
PX	CN 106990177 A (山东大学) 2017年 7月 28日 (2017-07-28) 全文	1-10
X	PARTE, P.P.等. "Sperm phosphoproteome profiling by ultra performance liquid chromatography followed by data independent analysis (LC-MSE) reveals altered proteomic signatures in asthenozoospermia" Journal of Proteomics, 第75卷，第18期，2012年 7月 13日 (2012-07-13), 第5861-5871页	1-10
A	MCDONALD, W.H.等. "Shotgun proteomics and biomarker discovery" Disease Markers, 第18卷，第2期，2002年 12月 31日 (2002-12-31), 第99-105页	1-10
A	CHAN, C.C.等. "Motility and Protein Phosphorylation in Healthy and Asthenozoospermic Sperm" Journal of Proteome Research, 第8卷，第11期，2009年 8月 13日 (2009-08-13), 第5382-5386页	1-10
A	夏西洋. "睾丸蛋白质组学数据翻译后修饰再挖掘" 中国优秀硕士学位论文全文数据库医药卫生科技辑, 第04期，2016年 4月 15日 (2016-04-15), E059-34	1-10
A	CN 101477089 A (中国科学院计算技术研究所) 2009年 7月 8日 (2009-07-08) 权利要求1-17	1-10

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/106915

检索报告引用的专利文件			公布日（年/月/日）	同族专利	公布日（年/月/日）
CN	106872630	A	2017年 6月 20日	无	
CN	107024553	A	2017年 8月 8日	无	
CN	106932597	A	2017年 7月 7日	无	
CN	106996979	A	2017年 8月 1日	无	
CN	107015005	A	2017年 8月 4日	无	
CN	107037172	A	2017年 8月 11日	无	
CN	106996981	A	2017年 8月 1日	无	
CN	106996980	A	2017年 8月 1日	无	
CN	106990177	A	2017年 7月 28日	无	
CN	101477089	A	2009年 7月 8日	无	