

学校编码： 10384

分类号_____ 密级_____

学 号： 20520100153632

UDC_____

厦 门 大 学

博 士 学 位 论 文

基于液质联用技术的膀胱癌肾癌尿液代谢
组学及血清多肽组学研究

LC-MS Based Urine Metabonomics and Serum Peptidomics
on Bladder and Kidney Cancers

黄真真

指 导 教 师： 杭 伟 教 授

专 业 名 称： 分 析 化 学

论 文 提 交 日 期： 2013 年 5 月

论 文 答 辩 时 间： 2013 年 6 月

学 位 授 予 日 期： 年 月

答 辩 委 员 会 主 席： _____

评 阅 人： _____

2013 年 5 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

目录

摘要.....	1
Abstract.....	4
第一章 前言.....	7
1.1 癌症研究背景.....	7
1.1.1 癌症概况.....	7
1.1.2 癌症发病机制.....	7
1.1.3 膀胱癌研究现状.....	10
1.1.4 肾癌研究现状.....	16
1.2 代谢组学.....	20
1.2.1 代谢组学的产生和发展.....	20
1.2.2 代谢组学研究方法.....	22
1.2.3 代谢组学的应用.....	29
1.3 多肽组学.....	322
1.3.1 多肽组学概况.....	32
1.3.2 多肽组学研究方法.....	35
1.4 本论文的研究目的和内容.....	40
参考文献.....	40
第二章 膀胱癌患者尿液代谢组学研究.....	56
2.1 引言.....	56
2.2 实验部分.....	58
2.2.1 试剂与材料.....	58
2.2.2 样品收集和储存.....	59
2.2.3 样品前处理.....	59
2.2.4 实验条件.....	60
2.2.5 数据处理和化学计量学分析.....	61
2.3 结果与讨论.....	62
2.3.1 方法稳定性评估.....	62

2.3.2 多变量统计学分析	65
2.3.3 潜在标志物的筛查与鉴定	69
2.3.4 标志物的临床价值评估	76
2.3.5 潜在标志物的代谢路径分析	80
2.4 本章小结	84
参考文献	85
第三章 膀胱癌和肾癌诊断的全变量有导模型	90
3.1 引言	90
3.2 实验部分	92
3.2.1 试剂与材料	92
3.2.2 实验仪器	92
3.2.3 样品采集	93
3.2.4 样品制备	93
3.2.5 实验条件	94
3.2.6 数据预处理	95
3.2.7 化学计量学分析	96
3.3 结果与讨论	97
3.3.1 分析方法重现性评价	97
3.3.2 膀胱癌患者和健康人志愿者的分类	100
3.3.3 膀胱癌患者和肾癌患者的分类	104
3.3.4 膀胱癌和肾癌患者之间的代谢差异	105
3.4 本章小结	121
参考文献	122
第四章 肾癌患者血清多肽组学研究	128
4.1 引言	128
4.2 实验部分	130
4.2.1 试剂和材料	130
4.2.2 样品收集和储存	131
4.2.3 LC-MS 样品预处理	132

4. 2. 4 Tricine SDS-PAGE	133
4. 2. 5 实验条件.....	133
4. 2. 6 数据预处理	134
4. 2. 7 化学计量学分析	135
4. 2. 8 多肽鉴定.....	135
4. 3 结果和讨论	136
4. 3. 1 多肽提取方法评估	136
4. 3. 2 多变量统计学分析	139
4. 3. 3 多肽标志物的确定与评估.....	142
4. 4 本章小结.....	151
参考文献.....	15151
总结与展望.....	15858
攻读博士期间已发表和待发表论文.....	16161
致谢.....	16163

Contents

Abstract (in Chinese)	1
Abstract	4
Chapter 1 Introduction	7
1.1 Background of cancer research	7
1.1.1 Profiles of cancer.....	7
1.1.2 Pathogenesis of cancer	7
1.1.3 Introduction of bladder cancer.....	10
1.1.4 Introduction of kidney cancer	16
1.2 Introduction of metabonomics	20
1.2.1 Generation of metabonomics	20
1.2.2 Methodology of metabonomics.....	22
1.2.3 Application of metabonomics	29
1.3 Introduction of peptidomics	32
1.3.1 Generation of peptidomics.....	32
1.3.2 Methodology of peptidomics	35
1.4 Objectives of this research	40
Reference	40
Chapter 2 Urinary metabonomics on bladder cancer patients	56
2.1 Introduction	56
2.2 Experimental	58
2.2.1 Reagents and materials.....	58
2.2.2 Sample collection and storage	59
2.2.3 Sample preparation.....	59
2.2.4 Experimental conditions	60
2.2.5 Data treatment and chemometrics analysis.....	61
2.3 Results and discussion	62
2.3.1 Stability assessment	62
2.3.2 Multivariate statistical analyses.....	65
2.3.3 Biomarker discovery and identification.....	69
2.3.4 Biomarker evaluation.....	76

2.3.5 Pathway elucidation of biomarkers.....	80
2.4 Conclusion	84
Reference	85
Chapter 3 Bladder and kidney cancers determination via OPLS-DA model	90
3.1 Introduction	90
3.2 Experimental	92
3.2.1 Reagents and materials.....	92
3.2.2 Apparatus.....	92
3.2.3 Sample collection.....	93
3.2.4 Sample preparation.....	93
3.2.5 Experimental conditions.....	94
3.2.6 Data pretreatment.....	95
3.2.7 Chemometrics analysis.....	96
3.3 Results and discussion	97
3.3.1 Stability assessment.....	97
3.3.2 Classification of bladder cancers and healthy controls.....	100
3.3.3 Classification of bladder and kidney cancers.....	104
3.3.4 Metabolic difference between bladder and kidney cancers.....	105
3.4 Conclusion	121
Reference	122
Chapter 4 Serum peptidomic study on kidney cancer	128
4.1 Introduction	128
4.2 Experimental	130
4.2.1 Reagents and materials.....	130
4.2.2 Sample collection and storage.....	131
4.2.3 Sample preparation for LC-MS analysis.....	132
4.2.4 Tricine SDS-PAGE.....	133
4.2.5 Experimental conditions.....	133
4.2.6 Data pretreatment.....	134
4.2.7 Chemometrics analysis.....	135
4.2.8 Peptide identification.....	135
4.3 Results and discussion	136

4.3.1 Evaluation of peptide extraction methods	136
4.3.2 Multivariate statistical analyses.....	139
4.3.3 Evaluation of peptide biomarkers	142
4.4 Conclusion	151
Reference	151
Summary and outlook	158
PUBLICATIONS	161
ACKNOWLEDGEMENT	163

厦门大学博硕士学位论文摘要库

摘要

膀胱癌是全球泌尿生殖系统发病率最高的癌症，每年有超过 350,000 新发病例。大约 70% 的膀胱癌患者在被确诊时还处于癌症早期，此时肿瘤多为膀胱内粘膜浅表的乳头状瘤，转移风险低，手术根治后可以获得较长的生存时间。侵犯膀胱壁肌层的非乳头状瘤转移风险高，预后较差；肾癌每年大约有 208,000 新发病例，其发病率在泌尿生殖系统所有癌症中排名第二，是致死率最高的癌症。当肾癌肿瘤局限于肾脏器官或者周围组织时，手术治疗的存活率可以达到 70%。不幸的是，大约 50% 的肾癌患者在被确诊时已经处于肾癌晚期或者发生全身性转移，即便接受治疗，患者的 5 年存活率也不足 2%。因此，早期诊断方法，尤其是非侵入式的诊断方法对提高膀胱癌和肾癌患者的生存率及生活水平具有重大意义。对于膀胱癌，当前主要的诊断方式是膀胱镜技术和尿脱落细胞学检查技术。然而，膀胱镜具有侵入性质，尿脱落细胞学检查灵敏度低；而对于肾癌，至今仍然没有特异性的诊断方式。

各国对基因组和蛋白组癌症标志物的研究已经开展许多年，然而，在基因和蛋白水平上发展癌症诊断方法并不是那么顺利，大体是因为大分子的基因和蛋白不能自由的穿透细胞膜，进入人体循环系统。代谢组学 (Metabonomics) 和多肽组学 (Peptidomics) 是两个相对年轻的组学技术，其研究对象即多肽和代谢物体积小，可以自由的进入到体液中。因此，与基因组学 (Genomics) 和蛋白组学 (Proteomics) 相比，代谢组学和多肽组学在疾病标志物方面的成果更容易向临床转化。

本论文开展了基于液质联用 (LC-MS) 平台的膀胱癌尿液代谢组学和肾癌血清多肽组学研究，研究主要内容及成果介绍如下：

1. 使用 LC-MS 平台对 27 例膀胱癌患者和 32 例正常人志愿者进行尿液代谢组学分析。本研究不仅使用了传统的反相色谱法 (Reversed phase liquid chromatography, RPLC)，还使用了与之分离原理互补的亲水作用色谱法 (Hydrophilic interaction liquid chromatography, HILIC)。我们对两种色谱分离模

式单独获得的数据以及两种色谱分离模式结合后的数据进行偏最小二乘法分析 (Partial least squares-discriminant analysis, PLS-DA), 结果显示这种结合可以提高模型对膀胱癌患者和正常人的鉴别能力; 我们提出了一套比常规方法更加严格的标志物筛查程序; 采用受试者工作特征曲线 (Receiver operator characteristic, ROC) 评估了单个标志物的临床诊断性能, 并将曲线下面积 (Area under the curve, AUC) 最大的两个代谢物组合应用于所有 27 例膀胱癌患者的诊断, 其灵敏度可以达到 92.6%, 特异度可以达到 96.9%; 该组合应用于早期膀胱癌诊断时, 灵敏度和特异度可以分别达到 90.5% 和 96.9%。受试者工作特征曲线结果显示该标志物组合对于膀胱癌诊断的性能优越于单标志物的方法。

2. 我们使用 LC-MS 平台对 19 例膀胱癌患者、25 例肾癌患者和 24 例正常人志愿者进行了尿液代谢组学研究, 该研究使用了反相色谱分离法和亲水作用色谱分离法。肾脏是尿液形成的器官, 膀胱是储存尿液的器官。应用尿液代谢组学工具将膀胱癌患者和肾癌患者区别开来, 至今仍然是研究的盲点。我们首次探索并评估了全变量正交偏最小二乘分析法 (Orthogonal partial least squares-discriminant analysis, OPLS-DA) 和单标志物方法用于膀胱癌和肾癌患者分类时的性能, 并最终建议了一种快速筛查膀胱癌和肾癌患者的方法。OPLS-DA 模型可以实现膀胱癌患者、肾癌患者和正常人志愿者的完全分类, 灵敏度和特异度同时表现为 100%。实验结果表明, 与正常人志愿者相比, 马尿酸在膀胱癌患者尿液中的含量降低了约 10 倍, 是膀胱癌诊断潜力最大的标志物, 其灵敏度和特异度分别为 78.9% 和 86.5%。与正常人志愿者相比, 肉毒碱 (C10:3) 在肾癌患者尿液中的含量降低了大约 1.5 倍, 灵敏度和特异度分别 84.0% 和 60.5%。从疾病诊断的灵敏度和特异度来看, 全变量 OPLS-DA 在膀胱癌和肾癌诊断方面, 比当前发现的单代谢物标志物的方法更加精确, 而且, 全变量法对早期膀胱癌和早期肾癌患者诊断无歧视。因此, 我们提倡发展全变量 OPLS-DA 建模法作为肾癌和膀胱癌的预筛查手段。

3. 使用 RPLC-MS 的方法研究了肾癌患者的血清多肽组, 一方面希望能找到肾癌多肽组的一些表达特征, 另一方面希望能够找到有助于肾癌诊断的多肽标志物。在多肽提取方面, 聚丙烯酰胺凝胶电泳 (Polyacrylamide gel electrophoresis, PAGE) 显示, 改进的乙腈沉淀法比经典的有机溶剂沉淀法和传统的超滤法更显

优势。肾癌多肽组研究结果显示，全变量的 OPLS-DA 模型可以实现肾癌患者和正常人志愿者的完全分类。基于 19 个小分子多肽的无导层序聚类分析（Hierarchical cluster analysis, HCA）将 30 个肾癌患者成功鉴别出来的同时，仅有一名正常人志愿者被错误的鉴定为肾癌患者。19 个多肽中 AUC 大于 0.9 的多肽共有四个，是肾癌诊断最有潜力的多肽标志物。

关键词：代谢组学；多肽组学；液相色谱-质谱；标志物；膀胱癌；肾癌

Abstract

Bladder cancer (BC) is the most common genitourinary malignancy. Each year, over 350, 000 new cases were diagnosed worldwide. Approximately 70% BC patients can acquire long-term survival, associated with superficial papillary lesions limited to the mucosa and submucosa of the bladder. Nonpapillary tumor that invades the bladder wall always has a high propensity for metastasis, and the prognosis is usually grimmer. Kidney cancer (KC) attacks approximately 208, 000 people in the world every year. Generally, it can be cured (survival rate > 70%) if it is diagnosed and treated when still localized to the kidney and its surrounding tissue. Unfortunately, more than 50% of the patients exhibit locally advanced- or metastatic -KC at the time of diagnosis, and a poor prognosis (5-year survival rate < 2%). Screening methods, preferably noninvasive, that facilitate early diagnosis of BC and KC would be advantageous for survival. However, for BC, cystoscopy is invasive, while urinary cytology cannot achieve high accuracy for early stage (ES) cancer detection. For KC, there is no specific diagnostic modality other than incidental radiologic discovery (computed tomography, ultrasonography, or magnetic resonance imaging).

Genomic and proteomic tools have been used for years in the studies attempting to discover clinical markers for a variety of cancers. However, they are still not readily amenable to discover diagnostic tests because large genes and proteins are not generally secreted into biofluids. Metabonomics and peptidomics, relatively new omics technologies, are more suited for such analysis because small metabolites and peptides are frequently secreted into serum. Thus, metabonomics and peptidomics research ultimately may be more clinically translatable than genomic and proteomic endeavors.

In this dissertation, LC-MS based urine metabonomics and serum peptidomics were carried out for BC and KC. These investigations were introduced briefly as follows:

1. A LC-MS based method, which utilized both reversed phase liquid chromatography (RPLC) and hydrophilic interaction chromatography (HILIC) separations, was performed, followed by multivariate data analysis to discriminate the global urine profiles of 27 BC patients and 32 healthy controls. Data from both columns were combined, and this combination proved to be effective and reliable for partial least squares-discriminant analysis (PLS-DA). Receiver operating characteristic (ROC) analysis was used for the evaluation of potential biomarkers. Two biomarkers were combined as a biomarker pattern, with a sensitivity and specificity up to 92.6% and 96.9%, respectively, for all patients and 90.5% and 96.9%, respectively for low grade BC patients. The combined biomarker pattern showed better performance than single metabolite in discriminating BC patients, especially low-grade BC patients, from healthy controls.

2. Kidneys are the oargans where urine produced, and the bladder is the urine container. The clinical exploration of urinary metabonomic analysis on discriminating between the top-two-incidence urological cancers, bladder and kidney cancers (BC and KC), is still virgin land. We discovered and evaluated the clinical utility of holistic metabonomic profiling and current single biomarker methods, and ultimately suggested a potential screening test for BC and KC. Urine metabonomic profiling for 19 BC patients, 25 KC patients, and 24 healthy controls was carried out using an LC-MS based platform. The holistic method that refers to orthogonal partial least-squares-discriminant analysis (OPLS-DA) based on all qualified profile data, successfully classified BC, KC and healthy control groups, showing 100% sensitivity and specificity. Hippuric acid was found 10-fold decrease in concentration relative healthy controls, and performed the best as a biomarker for BC diagnosis, with its sensitivity and specificity of 78.9% and 86.5%, respectively. Carnitine C10:3 was found 1.5-fold decrease, and reached 84.0% of sensitivity and 60.5% of specificity for KC diagnosis. In view of both sensitivity and specificity, the holistic method is more accurate for detecting BC and KC, than current single metabonomic biomarker

methods, and it could be advocated as a prescreen to other forms of more invasive or uncomfortable screening. Moreover, this approach also demonstrates attractive performance in diagnosis of early stage (ES) BC and KC patients.

3. We applied a RPLC-MS based serum peptidomic approach to investigate the serum peptidomic signature and discover the clinical biomarkers and biomarker patterns for patients with KC. As to extraction low molecular weight (LMW) peptides, results of Tricine SDS-PAGE demonstrated our modified precipitation method was advantageous over the typical peptide-extraction methods that involved classical organic precipitation and ultrafiltration. Our study has revealed, for the first time, holistic OPLS-DA model can separate patients with KC from healthy controls without any misclassification. Unsupervised hierarchical cluster analysis on several peptides was able to discriminate 30 KC patients from healthy controls while one healthy control was misclassified into cancer group. There are totally 4 peptides with AUC > 0.9 were observed, which might be the most promising biomarkers for KC detection. These findings may have implications on the understanding of peptidomics as well as KC detection.

Keywords: Metabonomics; Peptidomics; LC-MS; Biomarker; Bladder cancer; Kidney cancer

第一章 前言

1.1 癌症研究背景

1.1.1 癌症概况

癌症亦称恶性肿瘤，它是由于细胞发生突变使细胞的分裂增殖机制失调而引起的疾病。与正常细胞相比，它具有以下特征：失去了分化成熟的能力而不能成为正常的成熟细胞；具有相对的自主权，可以进行持续性生长；具有侵袭性和夺取营养的特征。肿瘤组织发展到一定阶段会侵犯和破坏周围正常组织和器官（浸润），或者通过淋巴循环和血液循环侵犯身体其它部位（淋巴转移和血行转移）。转移瘤是造成癌症患者病情恶化和死亡的主要原因。

根据国际抗癌联盟（Union international centre le cancer, UICC）2008年的统计数据，全球患癌人数约为1270万人，死亡人数约为760万人，其中，发展中国家的发病人数和死亡人数分别占总数的56%和64%^[1]。国际癌症研究中心（International agency of research center, IARC）报告称，假设各个年龄层的癌症发病率和死亡率恒定，预计到2020年，全球将有2000万癌症新发病例，死亡病例将达到1200万。在中国，20世纪70年代、90年代和21世纪每年死于癌症的人数分别为70万、117万和150万^[2]。癌症已经成为中国城市居民第一杀手，因癌症产生的死亡人数占城市总死亡人数的25%；已经成为农村居民第二杀手，因癌症死亡人数占农村总死亡人数的21%^[3]。2012年中国肿瘤登记年报数据显示，每6分钟就有一个人确诊为癌症。癌症严重威胁人类健康，给个人和家庭带来沉重的经济负担，已经成为各类疾病中花费最多的疾病，占有所有疾病支出总费用的7.23%^[4]。

1.1.2 癌症发病机制

关于癌症发病机理的探讨，目前主要有四种理论，即基因突变说、免疫监视说、细胞反分化说及幼稚细胞分化受阻说，其中获得广泛认同的是基因突变说。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库