在本课题中，我们除了DACH1的表观修饰调控及在肺癌发病机制的研究外，还对与DACH1在肺癌发病诊断中具有同等重要作用的APC基因在肺癌诊断进行了定性评估。APC基因在多篇文章中都显示与DACH1在肺癌发病中都发挥重要作用（Govindan，Cell,2012）。我们还采用数据挖掘的方法，开发了一套利用海量公共数据进行诊断位点搜索，并依托中国肺癌样本进行验证的，肺癌诊断标记物及诊断模型开发的研究平台。上述两个研究内容具体如下：

**APC基因启动子区DNA甲基化对非小细胞肺癌临床诊断价值的定量评估**

APC基因为典型的抑癌基因，我们通过采用Meta分析及全基因组甲基化芯片的联合分析对APC基因启动子区DNA甲基化的非小细胞肺癌临床诊断价值进行了定量评估，发现APC基因启动子区域DNA甲基化与非小细胞肺癌显著相关。DNA甲基化一直被认为是肿瘤早期诊断的优良生物标记物，然后可以应用于临床诊断的有效诊断产品却一直没有出现。该研究采用两阶段验证策略发现了大量对APC基因的诊断效能产生影响的异质性因素。研究发现包括引物的选择、样本中腺癌和鳞癌的比例、肺癌诊断时患者的年龄、建立肺癌诊断模型所采用的对照类型等都严重影响诊断模型的效能。揭示了建立基于APC基因DNA甲基化的非小细胞肺癌诊断模型临床转化所需要注意众多因素，为将来DNA甲基化诊断模型的临床转化具有较高的参考价值。该研究成果于2014年3月24日作为***Highly Accessed***论文发表在***Clinical Epigenetics***(2014, 6(1):5)，复旦大学为该论文的通讯作者。

**基于跨平台全基因组DNA甲基化芯片数据筛选并建立非小细胞肺癌生物标记物**

通过上部分的分析我们可以看出单基因DNA甲基化虽然具有重要的肺癌诊断意义，但提供的诊断效能是有限的，其诊断准确度距离临床应用尚有较大差距。采用多位点联合诊断是目前肿瘤诊断的发展趋势。因此搜索最优的甲基化生物标记物组合以获得最大的诊断性能是一个巨大的挑战。我们实验室采用整合公共数据库中已经收录了大量肿瘤样本及其对照样本的全基因组DNA甲基化芯片数据，通过对这些数据的整合，标准化，批次效应消除等处理，可以有效地增加肿瘤标记物探索阶段的样本量，并据此发现了5个重要的非小细胞肺癌甲基化生物标记物（AGTR1，GALR1，SLC5A8，ZMYND10和NTSR1）。继而实验室又开发了基于SNPshot的甲基化状态确定单核苷酸引物延伸技术（MSD-SNuPET）利用汉族人群非小细胞肺癌和正常对照对上述五个肺癌生物标记物的甲基化状态进行了检测。并采用多种预测模型，包括logistic回归，支持向量机，随机森林预测及Bayes树的方法，结合5倍交叉验证的方法建立了肿瘤/正常组织预测模型，并对预测模型的灵敏性，特异性和准确性进行了评估。结果证实AGTR1，GALR1，SLC5A8，ZMYND10和NTSR1为非小细胞肺癌有效的甲基化生物标记物，其联合甲基化谱可作为一种有效的非小细胞肺癌诊断模型。该研究成果于2014年12月20日被***Clinical Epigenetics***杂志***Highly Accessed***形式接收，复旦大学为该论文的通讯作者。

本部分研究利用多种分析技术对DNA甲基化生物标记物的开发，标记物诊断潜能评估进行了深入的研究。为DNA甲基化在非小细胞肺癌发病机理及临床诊断应用奠定了很好的基础。并证实了DNA甲基化作为肿瘤诊断标记物的巨大潜力。