

DOI:10.12025/j.issn.1008-6358.2018.20180316

炎症性肠病相关结肠癌组织差异表达基因的筛选及验证

徐晓晶¹, 余一祎¹, 王志明¹, 崔越宏¹, 侯英勇², 刘天舒^{1*}

1. 复旦大学附属中山医院肿瘤内科, 上海 200032

2. 复旦大学附属中山医院病理科, 上海 200032

[摘要] **目的:**探讨炎症性肠病相关结肠癌患者的肿瘤组织及癌周组织的基因表达差异。**方法:**收集 2016 年 1 月至 6 月经术后病理学确诊为结肠腺癌合并克隆恩病或溃疡性结肠炎 3 例患者的石蜡包埋组织(FFPE);采用 Illumina Hiseq 平台对 3 例患者肿瘤组织和癌周组织进行测序,分析基因表达差异,并进行 Pathway 功能分类和富集分析;采用炎症性肠癌动物组织进行差异基因验证。**结果:**测序数据经过滤后,保存半年内的石蜡包埋组织标本的碱基含量分布及质量分布均符合分析要求,且样品间数据具有可比性。肿瘤组织和癌周组织间的差异表达基因涉及的通路集中于免疫性疾病、感染性疾病、信号转导等,其中富集最显著且包含差异基因最多的是原发性免疫缺陷相关通路。**结论:**炎症性肠病相关结肠癌的发生与免疫缺陷的关系最密切,其次为感染和信号转导通路。

[关键词] 炎症性肠病;结直肠恶性肿瘤;新一代测序;免疫**[中图分类号]** R 735.3⁺5 **[文献标志码]** A

Screening and verification of differentially expressed genes in colon cancer tissues associated with inflammatory bowel disease

XU Xiao-jing¹, YU Yi-yi¹, WANG Zhi-ming¹, CUI Yue-hong¹, HOU Ying-yong², LIU Tian-shu^{1*}

1. Department of Oncology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

2. Department of Pathology, Zhongshan Hospital, Fudan University, Shanghai 200032, China

[Abstract] **Objective:** To analyze the genomic differences between tumor and peritumoral tissues of patients with colon cancer associated with inflammatory bowel disease. **Methods:** We collected formalin fixed and paraffin embedded tissues (FFPE) from 3 patients who were diagnosed as colorectal cancer during long-term follow-up of inflammatory bowel disease. They underwent surgical resection from January 2016 to June 2016 and were diagnosed as colon cancer complicated with Crohn's disease or ulcerative colitis pathologically. The tumor tissue and the adjacent peritumoral tissues were sequenced using Illumina Hiseq platform. Then we analyzed the gene expression difference and carried out the Pathway function classification and enrichment analysis. Differential gene validation was performed in animal tissues of inflammatory bowel cancer model. **Results:** The quantity and quality distribution of base content in FFPE specimens within six months after filtration were consistent with the requirements of statistical analysis, and the data between samples were comparable. The results showed that the different genes between tumor tissue and peritumoral tissue were mainly focused on autoimmune diseases, infectious diseases, and signal transduction. The primary immunodeficiency related pathways were mostly involved. **Conclusions:** The development of colon cancer associated with inflammatory bowel disease is most closely related to immunodeficiency, followed by infection and signal transduction.

[Key Words] inflammatory bowel disease; colorectal cancer; next generation sequencing; immune

癌症的形成是一个多因素、多步骤过程,慢性炎症作为其中一个重要环节参与了 10%~20% 的癌症发生,并与治疗反应密切相关^[1]。研究^[2]证明,炎症性肠病(inflammatory bowel disease, IBD)是

某些结直肠癌发生发展的危险因素,包括溃疡性结肠炎和克隆恩病。肠道炎症引起致癌突变进而导致肠癌的发生。随着对恶性肿瘤炎症和免疫微环境认识的进展,IBD 相关肠癌的基因组学特点及微

[收稿日期] 2018-04-01 **[接受日期]** 2018-06-10**[基金项目]** 国家自然科学基金(81502003,81602038,81772511)。Supported by National Natural Science Foundation of China (81502003, 81602038, 81772511)。**[作者简介]** 徐晓晶,硕士,住院医师。E-mail: xu.xiaojing@zs-hospital.sh.cn

* 通信作者(Corresponding author)。Tel: 021-64041990; E-mail: liu.tianshu@zs-hospital.sh.cn

生态菌群失调成为了目前研究的热点。

IBD所致肠道炎症的特点是中性粒细胞、巨噬细胞和其他免疫细胞特异性聚集,并释放各种细胞因子、自由基和蛋白水解酶等,从而建立一个肿瘤形成所需的微环境^[3]。研究^[4]发现,IBD中细胞因子和生长因子的表达谱与结直肠癌有相似之处,包括肿瘤坏死因子(TNF)、白细胞介素-1(IL-1)、IL-6、IL-17、IL-22、IL-23等。IBD的另一个特点是肠道菌群失调,如具核梭杆菌在结肠癌组织中富集并发挥重要作用^[5]。但目前针对IBD相关肠癌发生的具体分子机制尚不清楚。

目前从炎症到肿瘤的基础和转化研究主要利用动物模型^[6],而缺乏人群长期随访结果。因此,本研究采用3组IBD相关结肠癌患者的肿瘤组织和癌周组织进行二代测序,旨在探索炎症相关性肠癌的基因组学特点,寻找并验证差异表达基因,并对其富集的生物学功能进行分析。

1 资料与方法

1.1 一般资料 2016年1月至6月就诊我科因炎症性肠病长期随访发现恶性结肠癌者3例,其中1例合并溃疡性结肠炎,2例合并克隆恩病。男性2例,女性1例,中位年龄44岁。2例临床分期为ⅡA期,1例为ⅢB期。3例患者均符合以下条件:(1)按WHO组织学分类标准,经病理学确诊为结肠腺癌合并克隆恩病或溃疡性结肠炎;(2)于2016年1月至6月行手术切除,具体时间分别为2016年2月16日、6月1日、6月15日;(3)手术标本行甲醛固定、石蜡包埋,室温条件下保存,为尽量减少基因降解和损伤,标本均于半年内进行二代测序;(4)手术前未接受过放化疗;(5)既往接受过氨基水杨酸制剂药物治疗,术前半年未接受激素或免疫抑制剂治疗。

1.2 新一代测序及分析 取石蜡块内部组织,使用切片机切割成5~20 μm切片,总量相当于80 μm。采用Ambion® RecoverAll™总核酸提取试剂盒提取RNA。用DNase I去除基因组DNA的污染,用带有Oligo(dT)的磁珠富集真核生物mRNA;加入打断试剂在Thermomixer中适温将mRNA打断成短片段,以打断后的mRNA为模板合成一链cDNA,然后配制二链合成反应体系合成二链,并使用试剂盒纯化回收、黏性末端修复、cDNA的3'末端加上碱基“A”并连接接头,经片段大小选择后进行

PCP扩增。构建好的文库用Agilent 2100 Bioanalyzer和ABI StepOnePlus Real-Time PCR System质检,合格后使用Illumina HiSeq 2000平台进行测序。过滤掉低质量、接头污染以及未知碱基N含量过高的数据,然后使用Bowtie2软件和RSEM软件计算基因表达水平。分析不同样品间的差异表达基因,并根据KEGG注释结果及官方分类,将差异基因进行功能分类和生物学通路富集分析。

1.3 建立炎症性肠癌动物模型 采用BALB/c品系雄性小鼠,体质量18~20 g,6~8周龄,第1天腹腔注射氧化偶氮甲烷(AOM)10 mg/kg,予3%右旋葡聚糖硫酸钠(DSS)自由饮用1周,后改为正常饮水2周,如此循环6次。

1.4 RNA抽提及RT-PCR检测 将动物肿瘤组织和癌周组织剪成小块在液氮中磨成粉末,采用TRIzol试剂提取组织内总RNA,测定纯度及浓度。取1 μg RNA反转录成含cDNA的反应体系20 μL,稀释10倍。每次取2 μL。使用SYBR Green进行RT-PCR反应。CST1前引物为:5'-CCG CAC CAT ATG TAC CAA GT-3',后引物为:5'-CGT AGA TCT CGA AAG AGC ACA A-3'。SFRP4前引物为:5'-CAC AAC GGT GGT GGA TGT AA-3',后引物为:5'-GTG GAC ACT GGC AAG AAG AA-3'。MSLN前引物为:5'-GTC AAC AAA GGG CAC GAA ATG-3',后引物为:5'-AGG GTG TCT AGG GTG TCT TT-3'。TRIM29前引物为:5'-GGA GAC CAC CCA GAA GAA TTT-3',后引物为:5'-GGC AGG TCA TT GTC AGA GTT-3'。CCL11前引物为:5'-TTC AGC GAC TAG AGA GCT ACA-3',后引物为:5'-GAA TCC TGC ACC CAC TTC TT-3'。LTF前引物为:5'-GTG GTG TCT CGG ATG GAT AAG-3',后引物为:5'-GGG CAG TCA GAT CCA TTT CT-3'。β-actin前引物为:5'-CGT GGG CCG CCC TAG GCA CCA-3',后引物为:5'-TTG GCT TAG GGT TCA GGG GGG-3'。PCR扩增程序:95℃预变性10 min,95℃变性10 s,60℃退火及延伸20 s,循环数为45个。用RT-PCR仪自带软件分析,获得扩增产物的Ct值,结果采用 $2^{-\Delta\Delta C_t}$ 法来分析mRNA的相对表达量。

1.5 统计学处理 采用SPSS 12.0软件,比较样本间的基因表达量采用平均值的差异倍数(fold change),倍数≥2倍且检验水准(α)为0.05。在富

集分析时,对 P 值进行误判率(FDR)校正,通常 $FDR \leq 0.01$ 者视为显著富集。FDR 是错误发现率的简称,用于控制多重试验中的 I 型错误率。

2 结果

2.1 采用 FFPE 标本进行测序的可靠性 各个样本中质量值大于 20 的碱基数目占总碱基数目的比例(Q20)均大于 99%,碱基含量分布以及质量分布均符合统计分析的要求。与参考基因组序列进行比对后,每个样品的比对率均在 80%左右,同时样品间均匀的比对率表明样品间具有可比性。

2.2 差异表达基因检测 根据各个样品基因表达水平结果,筛选出肿瘤组织和癌周组织间的差异表达基因。结果(表 1)显示:在肿瘤组织中的表达显著上调的基因包括 CST1、KRT80、SFRP4、MSLN 等,显著下调的包括 SLC6A19、HGD、CCL11、LTF 等。

表 1 肿瘤组织/癌周组织前 10 位的差异表达基因					
基因 ID	基因名称	上调倍数	基因 ID	基因名称	下调倍数
1469	CST1	52.68	340024	SLC6A19	87.23
144501	KRT80	20.56	3081	HGD	76.37
6424	SFRP4	19.47	6356	CCL11	71.67
10232	MSLN	18.09	4057	LTF	70.78
23650	TRIM29	16.97	9934	P2RY14	66.78
91156	IGFN1	16.01	2641	GCG	55.08
1301	COL11A1	15.77	838	CASP5	45.13
4320	MMP11	15.33	3512	JCHAIN	36.44
3624	INHBA	15.10	59272	ACE2	32.44
4070	TACSTD2	13.76	2169	FABP2	28.33

2.3 采用动物组织验证差异表达基因 建立炎症性肠癌动物模型后选择其中 6 个基因(CST1、SFRP4、MSLN、TRIM29、CCL11、LTF)在动物组织中进行 PCR 验证,结果表明除 MSLN、CCL11 外其他基因的表达趋势与测序结果一致(图 1、图 2)。

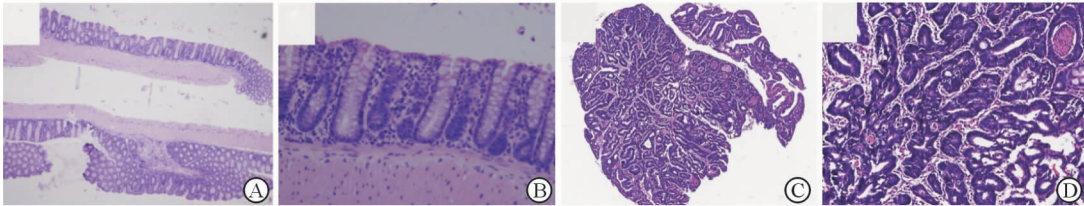


图 1 在炎症性肠癌动物组织中验证差异表达基因

A,B:小鼠癌周组织 H-E 染色,可见正常腺体结构;C,D: AOM/DSS 联合诱导小鼠 CAC 模型标本 H-E 染色,可见典型恶性表现,如腺体紧密贴合、形成筛孔状、腺泡内可见坏死。Original magnification: $\times 40$ (A,C), $\times 200$ (B,D)

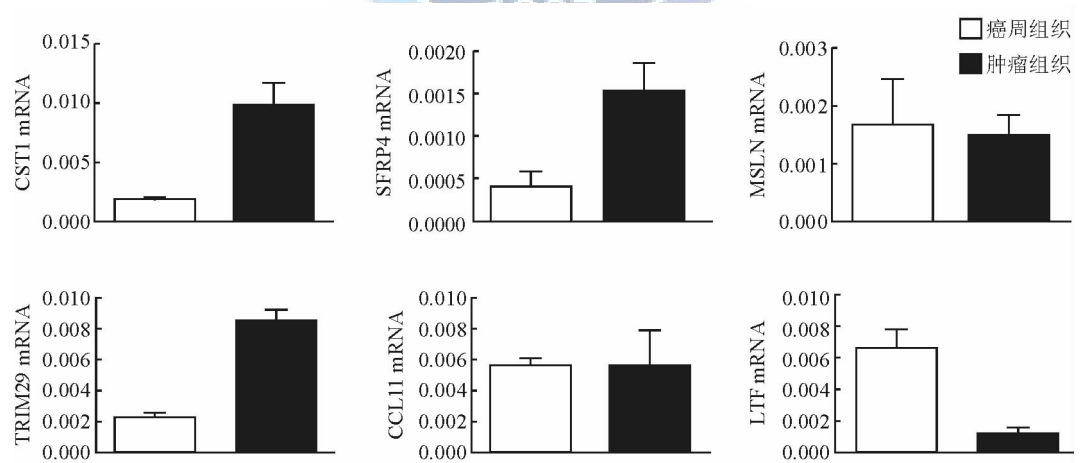


图 2 PCR 验证差异表达基因

PCR 验证结果证实小鼠肿瘤组织和癌周组织 CST1、SFRP4、TRIM29、LTF 的表达趋势与测序结果一致

2.4 Pathway 功能分类和富集分析 KEGG 生物通路分类和富集分析显示,IBD 相关的肿瘤组织与癌周组织间的差异表达基因共涉及 112 个通路,其中前 10 位的通路如表 2 所示。富集最多的通路集

中于免疫性疾病、免疫系统感染性疾病和信号转导等。其中,富集最明显且包含差异表达基因最多的是原发性免疫缺陷相关通路(图 3)。

表 2 前 10 位差异表达基因富集的通路

n(%)

序号	通路	具有通路注释的差异 表达基因 (N = 104)	具有通路注释的 所有基因 (N = 22 472)	P 值	Q 值
1	原发性免疫缺陷相关通路	34(32. 69)	274(1. 22)	$1. 43 \times 10^{-39}$	$1. 61 \times 10^{-37}$
2	IgA 产生的肠道免疫网络	14(13. 46)	156(0. 69)	$1. 55 \times 10^{-14}$	$5. 79 \times 10^{-13}$
3	类风湿性关节炎相关通路	14(13. 46)	245(1. 09)	$7. 69 \times 10^{-12}$	$1. 08 \times 10^{-10}$
4	FcεR I 信号通路	12(11. 54)	180(0. 8)	$4. 39 \times 10^{-11}$	$5. 47 \times 10^{-10}$
5	NF-κB 信号通路	13(12. 5)	241(1. 07)	$9. 26 \times 10^{-11}$	$8. 64 \times 10^{-10}$
6	B 细胞受体信号通路	12(11. 54)	206(0. 92)	$2. 11 \times 10^{-10}$	$1. 82 \times 10^{-9}$
7	EB 病毒感染相关通路	21(20. 19)	980(4. 36)	$3. 57 \times 10^{-9}$	$2. 22 \times 10^{-8}$
8	钙离子信号通路	14(13. 46)	413(1. 84)	$7. 13 \times 10^{-9}$	$4. 21 \times 10^{-8}$
9	自然杀伤细胞介导的细胞毒作用	12(11. 54)	374(1. 66)	$1. 65 \times 10^{-7}$	$7. 42 \times 10^{-7}$
10	PI3K-Akt 信号通路	21(20. 19)	1 240(5. 52)	$2. 02 \times 10^{-7}$	$8. 68 \times 10^{-7}$

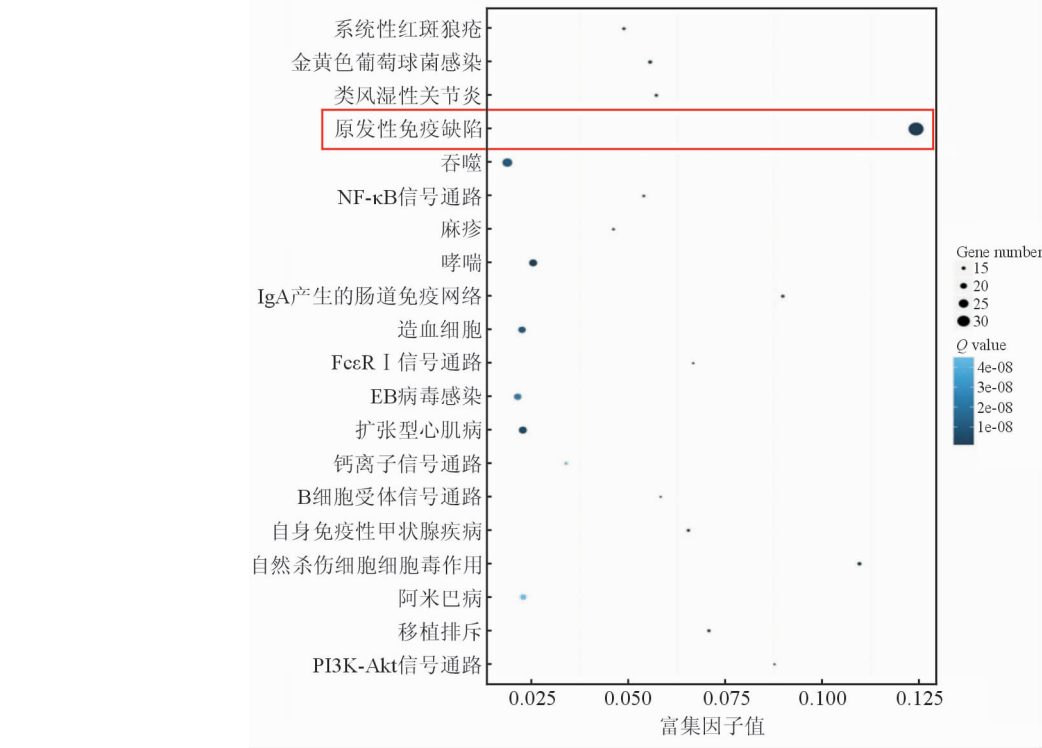


图 3 Pathway 功能富集图

颜色代表 Q 值(颜色越白值越大,越蓝值越小),值越小代表富集结果越显著;点的大小代表差异基因数目(点越大代表数目越多,越小代表数目越少)

3 讨 论

IBD 是一种自身免疫紊乱导致的肠道慢性非特异性炎症性疾病,是结直肠癌发生发展的重要危险因素。据报道,防治 IBD 的药物如氨基水杨酸、维生素 D 等可能降低结直肠癌的发生风险^[7-8]。由于缺乏长期随访的大规模前瞻性临床研究,目前关于炎症性肠病所致结直肠癌的研究主要基于动物模型。本研究采用回顾性的研究方式,对甲醛固定石蜡包埋临床样本进行测序。考虑到在制备和储存过程中,FFPE 样品的 DNA 容易发生降解、损伤、分

子或生物学修饰,因此本研究仅回顾性分析了半年内经手术切除的常温保存的石蜡标本。结果表明,经数据过滤后,其碱基含量分布以及质量分布均符合统计分析的要求且样品间具有可比性。因此,采用 FFPE 标本进行测序具备一定的可靠性。研究^[9-11]证实,在新鲜组织难以获得的情况下,FFPE 标本可以作为备选方案进行新一代测序。

本研究通过对 3 例炎症性肠病所致结肠癌患者的肿瘤组织和癌周组织进行测序,发现两者间存在上万个差异表达基因,涉及分子、细胞、生物过程等各个方面。进一步分类和富集分析提示,这些差异

基因主要与人类免疫缺陷性疾病和感染性疾病相关。其次,信号转导、运输和代谢也起到重要作用。NF- κ B、PI3K-Akt、钙离子、B细胞受体、EB病毒感染等传统信号通路也参与其中。本研究进一步总结了差异倍数前10位的基因,根据文献报道挑选了6个基因在炎症性肠癌动物模型的新鲜肿瘤组织和癌周组织中进行了PCR验证。结果提示,除MSLN、CCL11外,其他基因的表达差异趋势与测序结果一致。据报道,分泌型卷曲相关蛋白4(SFRP4)作为Wnt信号通路的调节剂,参与细胞增殖、分化、凋亡及肿瘤形成过程^[12];MSLN编码的间皮素作为细胞黏附蛋白,在多种肿瘤中呈现过表达^[13];人类宿主因子TRIM家族蛋白TRIM29通过抑制先天性免疫反应,从而促进鼻咽癌患者EB病毒持续性感染^[14];肿瘤相关钙离子信号转换因子2(TACSTD2)能够编码肿瘤相关性抗原,是一种参与钙离子信号转换的细胞表面受体^[15]。上述基因的表达在IBD相关肿瘤组织中均呈上调趋势。还有一些下调的基因,例如LTF编码的乳酸转铁蛋白作为非特异性免疫系统的重要部分,参与宿主防御微生物感染、调节细胞生长和分化、阻止肿瘤形成和转移^[16-17],其表达下调可能参与炎症及肿瘤的形成。

目前,炎症性肠癌动物模型一般采用致癌剂氧化偶氮甲烷(AOM)/致炎剂葡聚糖硫酸钠(DSS)方法诱导。在炎症至肿瘤的不同阶段留取标本进行基因组学、蛋白质组学的研究简单易行,但并不能准确反映人类炎症的特征。研究模型成为了限制因素。随着三维培养的发展,利用患者的自身肿瘤细胞,研究其与免疫细胞间的相互作用,并进行药物筛选,能为实现个体化治疗提供基础。今后有待收集新鲜的人类组织标本,进行前瞻性随机对照研究或组织学分析,以验证本研究结论。

参考文献

- [1] GRIVENNIKOV S I, GRETEN F R, KARIN M. Immunity, inflammation, and cancer[J]. *Cell*, 2010, 140(6):883-899.
- [2] BEAUGERIE L, ITZKOWITZ S H. Cancers complicating inflammatory bowel disease[J]. *N Engl J Med*, 2015, 372(15):1441-1452.
- [3] DE SOUZA H S, FIOCCHI C. Immunopathogenesis of IBD: Current state of the art[J]. *Nat Rev Gastroenterol Hepatol*, 2016, 13(1):13-27.
- [4] FRANCESCONI R, HOU V, GRIVENNIKOV S I.

Cytokines, IBD, and colitis-associated cancer[J]. *Inflamm Bowel Dis*, 2015, 21(2):409-418.

- [5] KOSTIC A D, GEVERS D, PEDAMALLU C S, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma[J]. *Genome Res*, 2012, 22(2):292-298.
- [6] ANTONIOU E, MARGONIS G A, ANGELOU A, et al. Cytokine networks in animal models of colitis-associated cancer[J]. *Anticancer Res*, 2015, 35(1):19-24.
- [7] CARRAT F, SEKSIK P, COLOMBEL J F, et al. The effects of aminosalicylates or thiopurines on the risk of colorectal cancer in inflammatory bowel disease[J]. *Aliment Pharmacol Ther*, 2017, 45(4):533-541.
- [8] LUAN Z, MA Y, XIN Y, et al. Possible molecular mechanisms by which vitamin D prevents inflammatory bowel disease and colitis-associated colorectal cancer[J]. *Curr Med Chem*, 2017, 24(9):911-917.
- [9] IVANOV M, LAKTIONOV K, BREDER V, et al. Towards standardization of next-generation sequencing of FFPE samples for clinical oncology: intrinsic obstacles and possible solutions[J]. *J Transl Med*, 2017, 15(1):22.
- [10] HEDEGAARD J, THORSEN K, LUND M K, et al. Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue[J]. *PLoS One*, 2014, 9(5):e98187.
- [11] EINAGA N, YOSHIDA A, NODA H, et al. Assessment of the quality of DNA from various formalin-fixed paraffin-embedded (FFPE) tissues and the use of this DNA for next-generation sequencing (NGS) with no artifactual mutation[J]. *PLoS One*, 2017, 12(5):e0176280.
- [12] POHL S, SCOTT R, ARFUSO F, et al. Secreted frizzled-related protein 4 and its implications in cancer and apoptosis[J]. *Tumour Biol*, 2015, 36(1):143-152.
- [13] HUCL T, BRODY J R, GALLMEIER E, et al. High cancer-specific expression of mesothelin (MSLN) is attributable to an upstream enhancer containing a transcription enhancer factor dependent MCAT motif[J]. *Cancer Res*, 2007, 67(19):9055-9065.
- [14] XING J, ZHANG A, ZHANG H, et al. TRIM29 promotes DNA virus infections by inhibiting innate immune response[J]. *Nat Commun*, 2017, 8(1):945.
- [15] GHEZZI T L, BRUM I S, BIOLCHI V, et al. Is there any association between TACSTD2, KIAA1253, Ku70 and mutant KRAS gene expression and clinical-pathological features of colorectal cancer? [J]. *Exp Oncol*, 2011, 33(1):28-32.
- [16] KANWAR R K, KANWAR J R. Immunomodulatory lactoferrin in the regulation of apoptosis modulatory proteins in cancer[J]. *Protein Pept Lett*, 2013, 20(4):450-458.
- [17] TSUDA H, KOZU T, IINUMA G, et al. Cancer prevention by bovine lactoferrin: from animal studies to human trial[J]. *Biomaterials*, 2010, 23(3):399-409.