

文章编号:1001-5078(2007)03-0240-04

肺癌组织红外光谱的统计分析

闫新惠¹, 唐伟跃², 侯晓强², 张建民²

(1. 河南省人民医院, 2. 郑州大学物理工程学院, 河南 郑州 450003)

摘要:使用统计的方法,应用傅里叶变换红外光谱法对肺癌及癌旁正常组织粉末样品的红外光谱进行了统计分析。结果表明,肺癌组织和癌旁正常组织的红外光谱有显著的差异性:3306 cm^{-1} 处的 N-H 峰在肺癌组织中向低频移动了约 3 cm^{-1} ;1167 cm^{-1} 处的 C-OH 伸缩振动谱带在肺癌组织中向高波数移动约 6 cm^{-1} ;1080 cm^{-1} 处的 PO_2^- 对称伸缩振动谱带在肺癌组织中向高波数移动约 4 cm^{-1} 。体现了在肺癌组织中细胞形态和组织结构发生的变化,说明傅里叶变换红外光谱法可以用于对良、恶性肺部组织鉴别诊断。

关键词:肺癌;傅里叶变换红外光谱;统计分析

中图分类号:R730.43;R734.2

文献标识码:A

The FTIR Spectroscopic Statistics Analysis on Tissues of Lung Cancer

YAN Xin-hui¹, TANG Wei-yue², HOU Xiao-qiang², ZHANG Jian-min²

(1. Henan People's Hospital; 2. Physics And Engineer College of Zhengzhou University, Zhengzhou 450003, China)

Abstract: Using statistics method, Fourier transform infrared (FTIR) spectroscopy was applied to the study of the spectrum of the powder of cancerous tissues and normal tissues near the cancer of the lung. The results showed that there were significant differences in the FTIR spectra between the normal and malignant pulmonary tissues: the peak at 3306 cm^{-1} of N-H shifted 3 cm^{-1} to short wavenumber in cancerous tissues, the peak at 1167 cm^{-1} of C-OH shifted 6 cm^{-1} to long wavenumber in cancerous tissues, the peak at 1080 cm^{-1} of PO_2^- shifted 4 cm^{-1} to long wavenumber in cancerous tissues. These difference embodies the variety of cell's structure and shape, this might indicate the FTIR could be a new method of the clinical diagnoses of the lung cancer.

Key words: lung cancer; FTIR; statistics

1 引言

肿瘤是威胁人类健康和生命的最大的疾病之一。近年来,由于抽烟人群的增加和环境恶化等因素的影响,肺癌发病率和死亡率在全世界大多数国家和地区的大城市中已居恶性肿瘤的首位^[1]。目前医学发展水平尚不能对其进行彻底治疗,因此早期诊断、及时治疗是提高病人生存机会的最有效途径。

构成组织和细胞的基本物质主要是糖类、蛋白质、核酸、脂类。细胞在癌变的过程中,这些组分的含量、构像均会发生变化,而这些变化可以通过其红

外光谱的变化表现出来。傅里叶变换红外光谱法已经应用于对多种肿瘤的研究^[2-4],本文应用红外光谱法对肺癌和癌旁正常组织的红外光谱进行研究,探讨了其光谱差异及导致其差异的可能原因,并进行了统计分析,以期对肺癌的早期诊断提供新的思路。

2 实验部分

2.1 实验仪器与样品

(1)实验仪器:美国 Nicolet 公司生产的NEXUS -

作者简介:闫新惠(1957-),女,副主任医师,从事核医学工作。
收稿日期:2006-08-11

470 傅里叶红外光谱仪。波数范围:4000 ~ 400 cm^{-1} , 分辨率:4 cm^{-1} 。德国产 CHRIST 冰冻干燥机,型号为 ALPHA1-2LD。

(2)样品:取自河南省人民医院 20 名病人肺部正常组织和癌变组织(已经病理学检查并确认)。

2.2 实验方法

(1)样品制备:将所取样品经冰冻干燥后碾成粉末,与 KBr 粉末混合后压制成薄膜状。

(2)实验方法:将制成的薄膜状样品置于红外光谱仪的光路上进行透射光的测定。

3 结果与讨论

3.1 正常与癌变组织的典型红外谱图

实验得到肺部正常和癌变组织的傅里叶变换红外谱图在 400 ~ 4000 cm^{-1} 范围内,其吸收峰主要集中在 800 ~ 2000 cm^{-1} 和 2500 ~ 3800 cm^{-1} 这两个波段内。图 1 和图 2 分别为肺部正常和癌变组织 KBr 压片在 800 ~ 2000 cm^{-1} 和 2500 ~ 3800 cm^{-1} 范围内透射光的典型红外光谱。

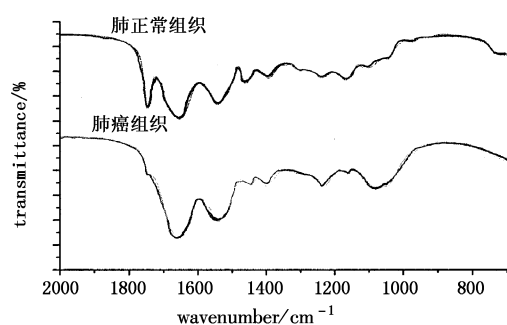


图1 肺部正常和癌变组织在 800 ~ 2000 cm^{-1} 范围内的红外光谱

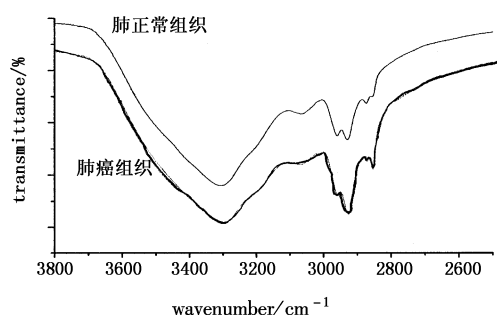


图2 肺部正常和癌变组织在 2500 ~ 3800 cm^{-1} 范围内的红外光谱

3.2 谱带归属^[2-5]

969 cm^{-1} 附近处属于磷酸化蛋白和核酸单脂 PO_3^{2-} 的对称伸缩振动谱带;1080 cm^{-1} 和 1238 cm^{-1} 附近处分别属于核酸分子磷酸二脂基团 PO_2^- 的对称和反对称伸缩振动谱带;1167 cm^{-1} 附近处属于蛋白质分子 C - O 伸缩振动谱带;1396 cm^{-1} 附近处属于甲基 - CH_3 变角振动谱带;1455 cm^{-1} 附近处属于

脂类分子弯曲振动谱带;1653 cm^{-1} 和 1540 cm^{-1} 附近处分别属于蛋白质酰胺 I、II 带;1744 cm^{-1} 附近处属于脂类分子的 $\text{vc} = \text{o}$ 振动;2854 ~ 3065 cm^{-1} 范围内属于脂肪的碳氢峰、甲基和亚甲基基团振动;3306 cm^{-1} 附近处属于 N - H 伸缩振动峰,如表 1 所示。

表 1 肺部组织的红外光谱吸收峰
振动模式的初步指认

正常肺组织 吸收峰位置/ cm^{-1}	癌变肺组织 吸收峰位置/ cm^{-1}	振动模式初步指认	A_n
3306.31 $\pm$ 1.06	3303.75 $\pm$ 1.03	N - H 伸缩振动峰	A_1
2872.54 $\pm$ 0.98	2872.23 $\pm$ 0.99	CH_3 的对称伸缩振动	A_4
2958.13 $\pm$ 1.10	2958.64 $\pm$ 0.84	CH_3 的反对称伸缩振动	A_2
2853.70 $\pm$ 1.52	2853.89 $\pm$ 1.84	CH_2 的对称伸缩振动	A_5
2926.46 $\pm$ 1.84	2926.21 $\pm$ 0.92	CH_2 的反对称伸缩振动	A_3
1653.54 $\pm$ 0.74	1653.78 $\pm$ 0.66	蛋白质酰胺 I 带	A_6
1540.85 $\pm$ 1.35	1541.15 $\pm$ 0.87	蛋白质酰胺 II 带	A_7
1455.11 $\pm$ 1.11	1455.25 $\pm$ 1.03	脂类分子弯曲振动谱带	A_8
1396.03 $\pm$ 1.72	1395.21 $\pm$ 1.95	甲基 - CH_3 变角振动谱带	A_9
1167.55 $\pm$ 0.93	1173.92 $\pm$ 1.58	蛋白质分子 C - O 伸缩振动谱带	A_{11}
1080.97 $\pm$ 1.31	1084.50 $\pm$ 1.17	核酸分子磷酸二脂基团 PO_2^- 的对称和反对称伸缩	A_{12}
1238.71 $\pm$ 1.55	1238.66 $\pm$ 1.49	振动	A_{10}
969.41 $\pm$ 1.53	969.97 $\pm$ 2.84	磷酸化蛋白和核酸单脂 PO_3^{2-} 的对称伸缩振动谱带	A_{13}

3.3 实验数据的统计分析

对 20 例肺癌和正常组织的压片进行相同条件下的红外透射光谱的采集,得到红外图谱以及与之相对应的数据。分别对癌变和正常组织的图谱分析,提取出有可能鉴别两者差别的特征峰的位置,如表 2、表 3 所示。表中 Mean 表示样本平均值;S 表示样本标准差; A_n ($n = 1, 2, \dots, 13$) 表示与表 1 谱带归属中相对应的特征吸收峰。那么,正常和癌变组织红外谱图中的特征峰位置就可以表示为:

$$\text{特征峰位} = \text{Mean} \pm S$$

由于 1744 cm^{-1} 附近处的脂类分子的 $\text{vc} = \text{o}$ 振动峰在大多数肺癌组织的红外光谱中很弱,甚至消失,因此没有将该处峰位列入表中进行分析。

3.3.1 实验数据的 F - 分布检验^[6]

比较正常和癌变两组数据在对应的波数处是否有显著性差别时采用配对 T - 分布检验的方法,但进行 T - 分布检验时要求比较的两组数据具有相同

表2 正常组织的红外光谱特征峰分布位置(波数/ cm^{-1})

	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉	A ₁₀	A ₁₁	A ₁₂	A ₁₃
1	3306.28	2959.40	2925.72	2872.85	2852.81	1654.44	1542.64	1454.13	1398.56	1239.65	1168.92	1081.59	969.95
2	3306.19	2959.13	2929.00	2872.99	2854.47	1654.49	1536.44	1453.53	1395.93	1237.88	1169.22	1082.28	969.33
3	3305.87	2958.78	2925.48	2872.65	2853.35	1653.84	1541.14	1454.80	1396.19	1237.41	1166.96	1080.70	970.03
4	3306.77	2958.35	2956.10	2872.47	2853.20	1653.92	1541.38	1455.70	1397.25	1237.78	1167.46	1084.53	969.35
5	3307.23	2955.39	2925.11	2873.46	2854.07	1653.52	1542.92	1456.74	1392.77	1235.18	1166.13	1081.27	970.21
6	3306.41	2958.21	2925.36	2872.19	2853.45	1653.83	1541.55	1454.56	1398.77	1239.33	1167.79	1081.39	969.68
7	3305.90	2957.79	2926.07	2872.36	2852.91	1654.37	1541.00	1454.77	1396.28	1242.84	1167.92	1080.92	967.73
8	3307.02	2958.64	2925.13	2872.20	2853.77	1654.71	1541.46	1455.29	1397.46	1239.23	1169.01	1081.41	970.12
9	3306.09	2956.93	2921.89	2871.93	2853.21	1653.92	1542.05	1453.14	1393.04	1237.55	1168.28	1080.95	969.83
10	3306.75	2958.35	2925.74	2872.58	2854.26	1653.84	1541.81	1454.30	1398.25	1238.29	1166.56	1081.10	973.04
11	3303.53	2958.46	2926.79	2872.29	2855.25	1652.97	1540.44	1457.08	1394.08	1238.08	1166.28	1078.58	968.00
12	3306.12	2959.10	2930.61	2870.90	2853.54	1653.07	1540.92	1453.43	1396.57	1238.91	1167.83	1080.46	970.54
13	3307.64	2958.89	2928.37	2872.04	2848.33	1654.08	1540.08	1455.28	1394.65	1237.53	1167.31	1082.04	970.18
14	3304.51	2959.63	2926.15	2871.88	2853.33	1653.28	1539.85	1455.31	1396.44	1239.67	1168.05	1081.62	967.49
15	3308.76	2958.42	2926.00	2873.17	2853.51	1652.98	1540.65	1456.75	1396.83	1238.43	1166.39	1079.43	968.33
16	3306.51	2959.92	2929.06	2873.35	2850.48	1653.30	1541.52	1455.69	1394.76	1237.98	1167.42	1080.73	969.67
17	3305.97	2959.90	2927.63	2872.85	2853.19	1653.23	1540.60	1455.56	1395.22	1241.00	1167.65	1081.71	970.25
18	3306.44	2960.17	2926.51	2869.87	2851.02	1652.55	1540.67	1454.33	1396.94	1238.75	1168.25	1079.57	968.75
19	3306.39	2957.87	2926.46	2872.49	2853.00	1651.84	1539.76	1456.04	1396.61	1239.22	1166.37	1080.06	965.45
20	3305.83	2958.49	2926.08	2869.95	2853.64	1652.67	1540.14	1455.82	1394.05	1239.54	1167.16	1078.96	970.26
Mean	3306.31	2958.13	2926.46	2872.54	2853.70	1653.54	1540.85	1455.11	1396.03	1238.71	1167.55	1080.97	969.41
S	1.06	1.10	1.84	0.98	1.52	0.74	1.35	1.11	1.72	1.55	0.93	1.31	1.53

表3 癌变组织的红外光谱特征峰分布位置(波数/ cm^{-1})

	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	A ₉	A ₁₀	A ₁₁	A ₁₂	A ₁₃
1	3303.57	2959.04	2925.93	2872.32	2853.47	1653.98	1541.48	1455.77	1396.55	1238.59	1172.21	1083.59	968.77
2	3303.08	2957.51	2926.12	2871.04	2852.19	1655.13	1541.36	1455.29	1395.60	1239.01	1171.95	1084.07	969.28
3	3302.36	2958.81	2925.60	2872.57	2857.97	1653.63	1540.99	1454.13	1395.46	1238.47	1173.01	1083.96	970.48
4	3303.42	2958.09	2925.41	2872.55	2853.60	1653.24	1540.98	1456.67	1397.30	1238.15	1173.53	1083.53	971.09
5	3304.55	2958.06	2925.45	2872.42	2850.83	1655.00	1541.61	1454.80	1392.36	1235.54	1172.38	1085.17	969.49
6	3304.08	2958.76	2925.61	2871.83	2853.93	1654.72	1541.76	1455.72	1394.38	1237.94	1175.44	1084.95	973.30
7	3303.23	2957.05	2924.98	2870.77	2854.17	1654.19	1542.62	1454.43	1396.99	1240.40	1173.82	1084.34	969.82
8	3304.67	2959.13	2925.73	2872.64	2859.26	1653.88	1541.23	1455.28	1397.51	1238.73	1174.31	1083.84	963.27
9	3303.51	2960.76	2927.56	2872.45	2854.32	1653.72	1540.39	1454.33	1394.02	1238.27	1173.15	1084.26	969.55
10	3302.87	2958.24	2925.38	2873.07	2853.16	1654.11	1542.84	1453.60	1397.33	1237.89	1175.06	1086.27	963.96
11	3305.46	2959.36	2926.33	2872.34	2853.76	1653.14	1540.92	1455.84	1392.58	1236.67	1169.92	1082.65	968.77
12	3303.82	2959.10	2926.09	2869.63	2853.57	1653.43	1541.08	1455.31	1397.87	1238.55	1175.22	1085.97	969.15
13	3303.69	2958.98	2925.75	2872.88	2854.09	1652.91	1541.15	1456.30	1392.77	1239.13	1176.43	1083.15	972.53
14	3301.95	2959.02	2926.22	2872.46	2853.60	1653.59	1540.90	1455.45	1394.72	1238.84	1175.58	1083.39	971.31
15	3303.72	2959.69	2926.43	2874.06	2853.11	1652.90	1539.63	1455.55	1397.61	1239.41	1173.77	1082.75	969.66
16	3304.86	2958.60	2925.79	2870.83	2854.47	1653.41	1542.64	1452.75	1394.46	1243.28	1174.02	1084.82	969.24
17	3306.33	2958.48	2928.82	2873.25	2852.92	1653.17	1540.74	1455.00	1392.22	1239.57	1175.08	1085.71	970.01
18	3302.79	2957.35	2926.44	2872.52	2853.82	1654.29	1540.88	1456.29	1393.39	1238.42	1175.66	1085.34	962.03
19	3303.48	2958.63	2926.93	2872.33	2853.38	1654.03	1539.64	1455.67	1396.92	1238.69	1174.85	1086.58	968.59
20	3303.51	2958.22	2927.61	2872.60	2852.16	1653.12	1540.19	1456.91	1394.25	1237.72	1173.09	1085.65	969.03
Mean	3303.75	2958.64	2926.21	2872.23	2853.89	1653.78	1541.15	1455.25	1395.21	1238.66	1173.92	1084.50	968.97
S	1.03	0.84	0.92	0.99	1.84	0.66	0.87	1.03	1.95	1.49	1.58	1.17	2.84

的方差。因此在进行两组数据的配对 T -分布检验之前,先对两组数据进行 F -分布的检验。只有在通过了 F -分布检验之后的 T -分布检验所产生的结果才有意义。 F -分布检验的显著性度 $\alpha = 0.05$ 。两组数据 $n = 20$,由 F -分布的原理可以得出拒绝域为 $F \geq F_{0.025}(19,19)$ 或 $F \leq F_{0.075}(19,19)$,经查表得出 $F_{0.025}(19,19) = 2.51$; $F_{0.075}(19,19) = 0.39$,因此分别求出两组数据的样本方差 s_1^2 和 s_2^2 ,再求与之对应的 $F = \frac{s_1^2}{s_2^2}$,比较 F 是否落在拒绝域内。如果落在拒绝域内即未通过 F -分布检验,说明两组数据的方差不一致;相反则说明两组数据的方差一致,可以进一步进行配对 T -分布检验。将表 2 和表 1 中的对应波数处的两组数据分别进行 F -分布检验,其结果如表 4 所示。

从表 4 可以看出 A_{13} 峰位即磷酸化蛋白和核酸单脂 PO_3^{2-} 的对称伸缩振动谱带正常和癌变数据组的 F 值落在 F -检验的拒绝域内,因此这两组数据不具有相同的标准差,不能用于进一步做 T -分布检验。其他位置处对应数据组的 F -分布检验均通过,即样本的标准差相同,可进一步通过 T -分布检验其差异的显著性。

表 4 峰位 F -分布检验结果

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}
F	1.05	1.30	1.99	0.99	0.83	1.12	1.55	1.08	0.88	1.04	0.59	1.12	0.29

3.3.2 实验数据的 T -分布检验^[6]

对通过 F -分布检验的数据组做 T -分布检验。 $n_1 = n_2 = 20$,检验的显著度 $\alpha = 0.05$ 。查表得到 $t_{0.05}(38) = 1.686$ 。由 T -分布的原理得到其拒绝域为 $(-\infty, -1.686)$ 和 $(1.686, +\infty)$ 。分别计算出不同峰位对应的正常和癌变两组数据的样本标准差 s_1 和 s_2 ,代入 T -分布函数中,计算出不同峰位处的 t 值。看其是否落在 T -分布检验的拒绝域内,如果 t 在拒绝域内则说明该峰位处正常和癌变组织的峰位有显著性差异。 T -分布检验结果如表 5 所示。

表 5 峰位 T -分布检验结果

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}
t	7.74	-1.66	0.55	1.01	-0.35	-1.07	-0.84	-0.42	1.41	0.10	-15.60	-8.98	-0.74

由 T -分布检验结果可以看出 A_1, A_{11}, A_{12} 峰位处的 T -分布检验的 t 值落在拒绝域内,这几个峰位处正常和癌变的峰位差异显著,有可能作为判别癌变与否的标志。

3.4 红外峰位差异结果分析

通过肺部正常和癌变组织的红外光谱特征峰位

的分析发现在以下波数处差异显著:

3306cm^{-1} 处的 $N-H$ 峰在肺癌组织中向低频移动了约 3cm^{-1} ,可能是由于氢键化的形式和程度不同而引起的频移^[5]。但由于仪器的分辨率为 4cm^{-1} ,因此该峰位处的差异是否能作为判别依据,还有待进一步使用分辨率较高的仪器通过实验确定。

1167cm^{-1} 处的蛋白质分子丝氨酸、苏氨酸和酪氨酸残基的 $C-OH$ 伸缩振动谱带^[7],在肺癌组织中向高波数移至 1173cm^{-1} 处,移动约 6cm^{-1} 。可能是由于 $C-OH$ 键的氢键化程度降低,结合氢键被破坏的缘故。

1080cm^{-1} 处的核酸分子磷酸二脂基团 PO_2^- 的对称伸缩振动谱带在肺癌组织中向高波数移动约 4cm^{-1} ,这是因为在癌变过程中,细胞内核酸分子中的 PO_2^- 基团氢键化程度发生变化引起^[8-9],氢键的加强使得峰位向高波数移动。

从实验结果可以看出,肺部正常组织与癌变组织的红外谱图有较大的差异性,通过对这些差异的分析表明,肺部癌变组织相对于正常组织来说,变化主要集中在氢键和磷酸二酯键集团上,体现了在肺癌组织中细胞形态和组织结构发生的变化。

从对肺部正常和癌变组织傅里叶变换红外光谱的分析研究,说明肺部组织细胞在癌变前后在分子结构以及相对含量上所发生的一些变化可以从它们的傅里叶变换红外光谱上表征出来,表明傅里叶变换红外光谱技术可成为诊断肺癌一种新方法。

参考文献:

- [1] 吴怀申.我国近十年外科治疗万例肺癌概况[M].中国肿瘤临床年鉴,1993;26.
- [2] 范晓燕.傅里叶变换红外光谱在生命科学中的应用[J].生命科学研究,2003,7(2):84-87.
- [3] Brigas B, Wong P T T. Application of Infrared Spectroscopy in Biochemistry and Medicine[J]. Cancer Research, 1992,52:84-87.
- [4] 沈世杰,刘炳玉,李清,等.人乳腺癌组织的特征红外光谱研究[J].光谱学与光谱分析,2000,20(1):28-30.
- [5] 谷志远.现代医学分子生物学[M].北京:人民军医出版社,1998.
- [6] 盛骤等.概率论与数理统计[M].北京:高等教育出版社,2001.
- [7] Benedetti E, Applied Spectroscopy[M]. 1990,39-40.
- [8] 霍红,王幸福,车迅,等.乳腺癌组织中蛋白质与核酸分子氢键征的研究,光谱学与光谱分析[J]. 2001,21(5):614-616.
- [9] 孙雷明,唐伟跃,郭克民,等.肺癌组织的红外光谱研究[J].激光与红外,2005,35(8):587-589.