

单细胞拉曼光谱结合多元统计方法 分析不同酿酒酵母菌株的成分差异

赖钧灼^{1,2}, 刘 斌³, 王桂文¹, 陶站华¹, 黄庶识^{1*}

1. 广西科学院, 广西 南宁 530003

2. 广西大学生命科学与技术学院, 广西 南宁 530004

3. 广西大学农学院应用微生物研究所, 广西 南宁 530005

摘 要 应用单细胞拉曼光谱技术结合多元统计方法分析比较了六株不同来源的酿酒酵母菌株的同步化细胞, 从而获知菌株间细胞成分的差异。利用连续密度梯度离心法分离酿酒酵母同步化细胞, 分别对六个酵母菌株的同步化单细胞进行拉曼信号收集, 并结合主成分分析(PCA)、判别函数分析(DFA)两种多元统计分析光谱差异。结果表明, 细胞的拉曼光谱表征其物质的基本组成与结构, 六个菌株的拉曼光谱比较发现菌株间细胞的蛋白质、脂类的差异较大, 而核酸类物质相差较小。六个菌株的拉曼光谱经多元统计分析可以在一定程度得到区分, 判别函数分析结果显示影响菌株间光谱差异最大共有 14 个峰: 706, 862, 918, 997, 1 073, 1 127, 1 269, 1 291, 1 305, 1 429, 1 465, 1 591, 1 602 和 1 652 cm^{-1} , 其中 10 个峰归属为蛋白质, 3 个峰归属为脂类, 1 个峰归属为核酸。根据 DFA 载荷, 从 1 340 个谱线中提取出上述的 14 个谱峰进行新的 DFA 结果, 与前面的结果是一致的。此方法能在单个活体细胞中进行, 无需复杂前处理, 能较好地研究遗传本质相差较小的酿酒酵母菌株间的成分差别。

关键词 拉曼光谱; 多元统计; 单细胞; 酿酒酵母菌株

中图分类号: O657.3 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2011)02-0412-06

引 言

不同的酿酒酵母菌株的性能存在较大差异, 选择优良菌株是有效发酵的关键因素。目前, 传统的微生物分类、鉴别方法是基于表型和生理生化特性实现, 难以区别酿酒酵母种内株间的差异^[1]。而且, 传统方法操作复杂、耗时耗力, 难以实现自动化及计算机化^[2]。近年来, PCR 技术、原位杂交、质谱、傅立叶红外光谱(FTIR)及拉曼光谱等先进的技术广泛应用于微生物分类鉴别中^[3]。其中, 拉曼光谱无需添加任何化学染料及其他标记, 是一种快速、特异、无损的微生物分析技术。核酸、蛋白、脂类及糖类生物大分子均能产生特异的拉曼光谱, 从特征光谱中能够获取生物分子组成、结构及生理状态等重要信息^[4], 因此从完整的微生物细胞拉曼光谱中可以获知其物质组成及结构等, 从而可以对不同微生物细胞的成分差异进行有效分析。与红外光谱相比, 水的拉

曼峰较弱, 能够在水溶液条件下收集样品的拉曼信号。此外, 拉曼光谱峰比较特异, 可以从获取更多信息^[5]。

单细胞拉曼光谱技术是显微拉曼光谱与光镊有机结合的一种光谱分析新技术, 已经广泛应用于细菌^[6]、酵母细胞^[7]、淋巴细胞^[8]等活体单细胞的研究。光镊有利于固定溶液中的单细胞, 提升单细胞同时可以减少来自底部玻片荧光的干扰, 大大提高了拉曼光谱的信噪比。

由于微生物细胞在不同的生长阶段具有不同的物质组成^[9], 而同一培养物中微生物的生长阶段并不一致, 为了更好地研究酿酒酵母菌株间的差异, 可利用密度梯度离心法对研究的酵母细胞进行同步化处理。本研究利用单细胞光镊拉曼光谱系统对六株不同酿酒酵母菌株同步化后的细胞进行拉曼光谱收集, 并结合多元统计分析方法(主成分分析, PCA; 判别函数分析, DFA)对不同菌株酵母同步化细胞的拉曼光谱进行分析, 从活体单细胞水平上提供了一种分析酿酒酵母菌株间成分差异的方法, 为区分不同酿酒酵母菌株提供有益

收稿日期: 2010-05-10, 修订日期: 2010-08-20

基金项目: 国家自然科学基金项目(30760010), 广西自然科学基金重点项目(2010GXNSFD013029)和广西科学院基本科研业务费项目(09YJ17WL01)资助

作者简介: 赖钧灼, 1985 年生, 广西大学生命科学与技术学院在读硕士研究生 e-mail: Jun19841984@163.com

*通讯联系人 e-mail: hshushi@gxas.cn; hshushi@yahoo.com.cn

参考,同时本方法在确定菌种来源以及选育优良菌株等都有应用潜力。

1 材料与方法

1.1 菌株

实验中所用酿酒酵母(*Sacharmyces, cerevisiae*)菌株为 S288C、ATTC(24859)、安琪酿酒高活性干酵母(安琪酵母股份有限公司)、梅山即发干酵母(广州梅山马利酵母有限公司)、INVSc1、AS2.400 等,分别标注为 A、B、C、D、E、F,除安琪酿酒高活性干酵母、梅山即发干酵母购于市场,其他均保藏于广西科学院国家非粮生物质能源工程中心。培养基为 YPD 培养基:葡萄糖 2%、酵母提取物 1%、蛋白胨 2%,上述菌株 28 °C 培养。

1.2 菌株同步化

用改良 Allen 等^[9]的 Percoll 连续密度梯度离心法分离同步化酿酒酵母细胞,即用角式转头替代超高速离心机的悬垂转头,15 mL Corex 管则用 2 mL 离心管代替,而操作步骤基本一致。

1.3 实验系统

实验系统如文献^[3,7]描述,780 nm 的激光被导入到一台尼康倒置生物显微镜(TE2000U, Nikon),激光束途经油浸物镜(N.A. 1.30, 100×),聚焦后在焦点附近形成了一个单光束光势阱来俘获细胞,同时这束激光也用于激发被囚禁细胞产生拉曼散射。此系统的分辨率为 6 cm⁻¹,用 780 nm 激光捕获悬浮于无菌水中的聚苯乙烯小球(直径 2 μm)收集拉曼光谱并校正系统。

1.4 光谱收集及数据处理

(1)光谱收集 取 200 μL Percoll 分离获取的酿酒酵母同步化细胞,置于厚 4 mm、中心有一个直径为 6 mm 孔、底部用 0.1 mm 厚的石英玻璃密封的玻璃板构成的样品池内,上加盖玻片密封。100 倍油浸物镜寻找酵母细胞,光镊随机捕获单个酵母细胞,将细胞提升至石英片上方 10 μm 处,30 s 收集细胞拉曼光谱,于细胞附近以相同条件收集背景拉曼光谱。激光激发功率为 15 mW。

(2)数据处理 直接将光谱数据以 spe 格式导入软件 Micro-Origin 8.0 进行前处理。应用自编的 Matlab 程序完成以下操作:对每个样品的 5 个背景光谱先五点平滑,进行平均得到平均背景光谱,然后用细胞光谱减去平均背景光谱,后经五点平滑获得每个细胞的真实光谱;而细胞平均光谱是由同一样品所有细胞的真实光谱进行平均获得。

1.5 多元统计分析

应用主成分分析(principal components analysis, PCA)、判别函数分析(discriminant function analysis, DFA)等多元统计方法对经数据前处理的酿酒酵母单细胞拉曼光谱进行分析。利用软件 The Unscrambler 9.7 对 6 株种同步化的酵母细胞的各 30 个细胞的拉曼光谱先进行标准归一化(standard normal variate, SNV),然后进行主成分分析, Micro-Origin 8.0 绘制 PCA 三维图。而 DFA 由软件 Pychem 3.0.5 完成,使用的数据经过 SNV 前处理, Micro-Origin 8.0 绘制 PCA

二维图。

2 结果与讨论

2.1 六个酿酒酵母菌株同步化细胞的拉曼光谱归属

图 1(a)和(b)为酿酒酵母细胞的 DIC(微分干涉)显微镜观察图,(a)为同步细胞,(b)为非同步细胞。从中可以看到经 Percoll 连续密度梯度离心法分离的同步化细胞以单个细胞为主,细胞比较致密;非同步化细胞以带芽体细胞为主,细胞液泡较大。图 1(c)为各取 30 个同步化细胞和非同步化细胞拉曼光谱的主成分分析结果,从中可以看出同步化单细胞间的光谱比较均一,而非同步化单细胞间的光谱呈现不均一性,说明同步化细胞间物质的基本组成和结构相一致,而非同步细胞的由于细胞间处于不同状态,其内含的物质存在较大的差异。

图 1(d)为六个不同酿酒酵母菌株同步化细胞的平均拉曼光谱,其光谱整体反映了细胞的物质组成和生物大分子的结构,根据文献^[4,7,10-12]等可知酵母菌拉曼光谱的归属为 713 cm⁻¹ 归属于核糖核酸(RNA)碱基腺嘌呤,782 和 1 082 cm⁻¹ 分别属于脱氧核糖核酸(DNA)、核糖核酸骨架中磷酸二酯(O—P—O—)的对称伸缩振动,谱线 891 cm⁻¹ 来源于 DNA 的 B-型构象类;谱线 860 和 828 cm⁻¹ 是酪氨酸对羟基环的呼吸振动和环平面外弯曲振动的倍频之间的费米共振的双峰,828 cm⁻¹ 峰强较弱,因此该双峰的强度比 $I_{850/835}$ 大于 1,表明酵母同步化细胞中蛋白质的酪氨酸是“暴露”的^[13];1 002 cm⁻¹ 来源于苯丙氨酸的单基取代苯基环,1 600 cm⁻¹ 属于苯丙氨酸和酪氨酸的谱线,1 652 cm⁻¹ 是蛋白酰胺 I 中 α -螺旋的信号峰;914 和 1 301 cm⁻¹ 来自脂类;1 123 cm⁻¹ 来自糖类;而 1 440 cm⁻¹ 则由脂类和蛋白质的 C—H₂ 扭曲振动共同产生。六株不同酿酒酵母菌株的同步化细胞的平均拉曼光谱均存在差异,其中阴影部分光谱差异较明显,主要表现为特征峰的强度不同,通过统计各菌株同步化细胞的拉曼光谱强度比值,发现菌株间的 $I_{1652/1600}$ 和 $I_{1301/1438}$ 差距较大,而 $I_{1082/1438}$ 各菌株间差距较小(见表 1),表明各株酵母细胞的蛋白质、脂类以及它们中一些基团的相对含量差异较大,而核酸类物质的差异稍小。此外,不同菌株间的同一物质或结构也存在拉曼特征频率的不同,菌株 C 的单细胞平均光谱中 1 645~1 660 cm⁻¹ 谱带主要由 1 646 cm⁻¹ 谱线组成,而其他菌株以 1 652 cm⁻¹ 谱线为主,这一谱带的拉曼信号主要来自蛋白酰胺 I 的 C=O 伸缩振动,属于蛋白质二级结构的 α -螺旋结构。

2.2 六个酿酒酵母菌株的同步化细胞拉曼光谱多元统计分析

酵母是内含蛋白质、糖类、脂类等化学组分的真核生物,酵母单细胞的拉曼光谱是整个细胞所有物质的光谱总和,因此其光谱呈现比较宽的峰和复杂的轮廓。为了提高光谱的特异性及灵敏度,以便更好地寻找六株不同酵母的同步化细胞光谱异同,采用主成分分析(PCA)、判别函数分析(DFA)等多元统计方法对酵母单细胞拉曼光谱数据进行分析。PCA 是利用降维处理技术,将原来多个变量化为少数互

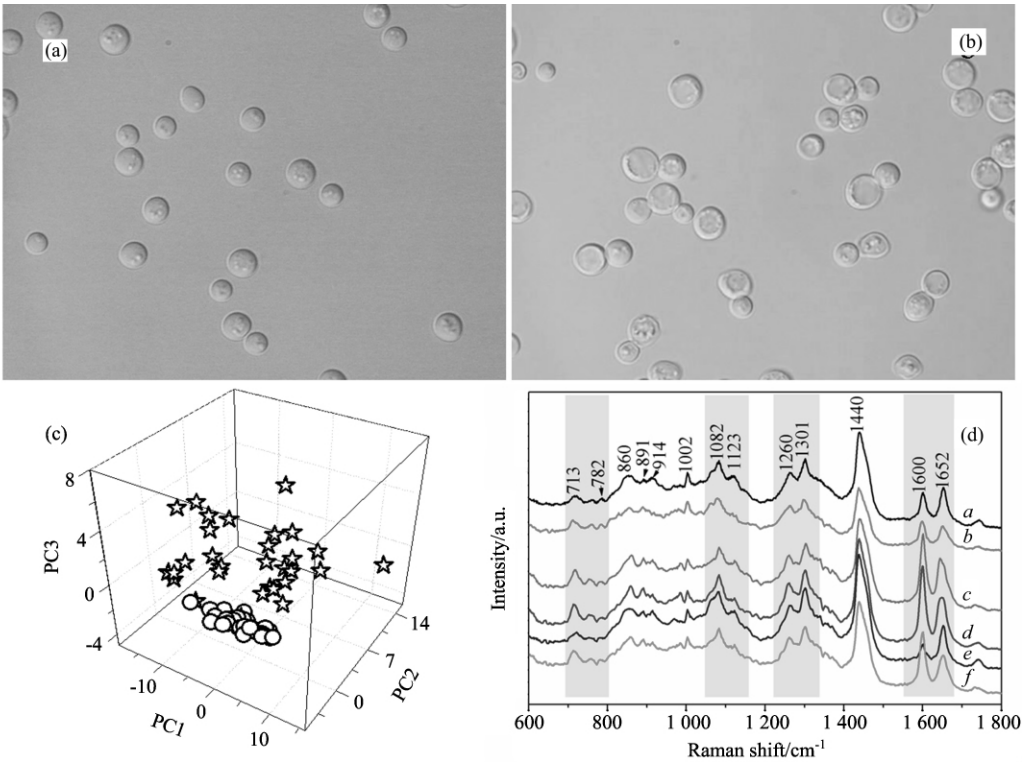


Fig 1 Images of synchronized (a) and the unsynchronized (b) yeast cells; (c) the principal components analysis of 30 cells from synchronized(○) and the unsynchronized(☆) yeast cells; (d) the averaged Raman spectra of six strains after synchronization

Table 1 Peak intensity ratios of Raman spectra of each yeast strain

吸收峰强度比值	A	B	C	D	E	F	大小排序
I_{1652}/I_{1600}	1.03	0.97	0.84	0.74	1.47	0.53	$E>A>1>B>C>D>F$
I_{1301}/I_{1438}	0.89	0.16	0.85	0.78	0.74	0.68	$1>A>C>D>E>F>B$
I_{1082}/I_{1438}	0.89	0.95	0.83	0.75	0.71	0.67	$1>B>A>C>D>E>F$

不相关的主成分，从而描述数据集内部结构的一种统计分析手段。而 DFA 在充分保护数据原有信息条件下，获取光谱中最重要的信息，从而扩大组间差异，缩小组内差异^[14]。

图 2 为六个酿酒酵母菌株各 30 个同步化细胞的主成分分析，图中的三维散点图为第一主成分(PC1)、第二主成分

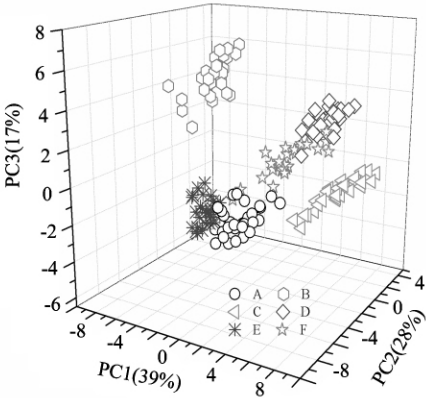


Fig 2 PCA of six yeast strains with 30 synchronized cells for each

(PC2)及第三主成分(PC3)组成，且三个主成分的累计贡献率达 84%。从图 2 中散点的分布可以看出菌株 B、C 各自处于相对独立的空间，与其他菌株明显区分，但菌株 A 和 E 间以及 D 和 F 间有部分重叠，初步说明菌株 B、C 的拉曼光谱与其他菌株的拉曼差异较大，而三维中具有重叠部分的菌株间拉曼光谱差异相对较小，通过主成分分析可以将六株酿酒酵母菌株的同步化细胞拉曼光谱在一定程度上区分。

为了从六个酵母菌株间的拉曼光谱中得到更多的有用信息，可利用判别函数分析(DFA)对六个酿酒酵母菌株同步化细胞的拉曼光谱进行判别函数分析。图 3(a)为六株酿酒酵母菌株各 30 个同步化细胞拉曼光谱的 DFA，图中 DF1 为第一判别函数，DF2 为第二判别函数，可以看出 DF1 中六个菌株明显分处于四个不同位置，其中 A 和 E 间以及 D 和 F 间出现重叠，但 DF1 中菌株间的差异最大，DF2 其次。图 3(b)为 DFA 中第一判别函数(DF1)的载荷，其中对应的相关系数绝对值较大的 706, 862, 918, 997, 1 073, 1 127, 1 269, 1 291, 1 305, 1 429, 1 465, 1 591, 1 602, 1 652 cm^{-1} 等 14 个拉曼峰在 DF1 起最关键的作用，而 DF2 只有 713, 1 429, 1 598, 1 642 等几个拉曼峰在 DF2 中起重要作用[见图 3

(c)]。在光镊拉曼系统中获得的细胞拉曼光谱共由 1 340 个谱线组成, 图 3(d)是根据 DFA 载荷从 1 340 个谱线中提取出上述的 14 个谱线对六个酵母菌株同步化细胞拉曼光谱进行新的 DFA, 即每个细胞光谱中只有 14 个谱峰。图 3(d)所显示菌株的二维分布基本和图 3(a)一致, 进一步说明这 14 个

峰是这六个酵母菌株的拉曼光谱中差异最大的地方。从表 2 中可以看出, 影响 DF1 的关键拉曼峰中, 除 1 429 cm^{-1} 来自嘌呤碱基以及 1 291, 1 269, 1 073 cm^{-1} 源于脂类物质外, 其他的 10 个峰都归属于蛋白质的信号峰。说明本研究的六个酵母菌株间细胞中蛋白质的组成和结构存在较大的差异。

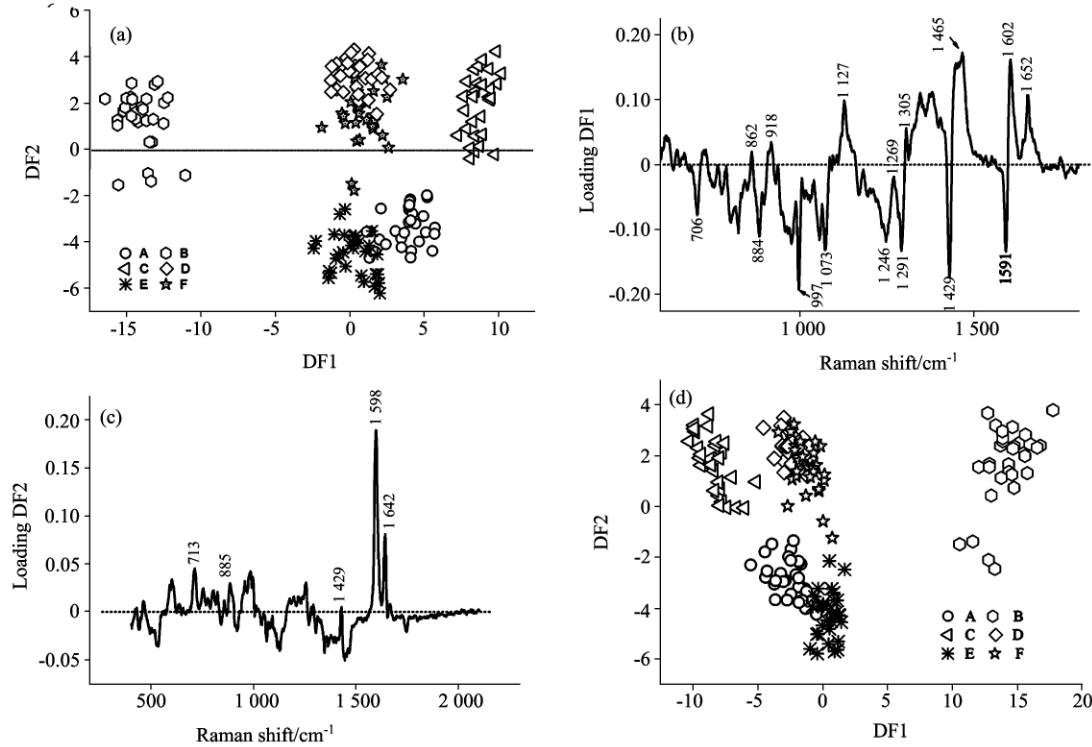


Fig 3 DFA of synchronized cells of six yeast strains. (a) DFA classification with the whole spectrum fromed synchronized cells for six yeast strains. (b) The loading of first discriminant function(DF1). (c) The loading of second discriminant function (DF2). (d) DFA plot only based on 14 most contributed bands(706, 862, 918, 997, 1 073, 1 127, 1 269, 1 291, 1 305, 1 429, 1 465, 1 591, 1 602 and 1 652 cm^{-1}) which are derived from Table 3

Table 2 Tentative assignment of some bands with most contribution yeast cells classification		
DF1/ cm^{-1}	对应基团指认	参考文献
706	C—S, proteins	[10]
862	Tyr	[10]
918	C—C, proteins	[10]
997	Phe	[10]
1 073	C—N, proteins; C—O, C—O—C str, lipids	[10]
1 127	C—C str, C—O—C glycosidic link	[15]
1 269	—C—H in-plane def, lipids	[15]
1 291	CH ₂ twist, lipid	[15]
1 305	Amide III, def	[4]
1 429	Adenine, Guanine	[4]
1 465	CH ₂ bending mode	[16]
1 591	Protein	[4]
1 602	Tyr, Phe	[10]
1 652	Amide I, α -helix	[4]

Notation: str, stretching; def, deformation

酒酵母进行聚类分析, 从图中可知酵母菌株 B 与其他五个酵母菌株相距最远, 说明其与五个酵母菌株中光谱差异最大, 即菌株 B 的细胞内物质结构与组成在六个菌株间的差异性最大。而菌株 A 与 E 间以及 D 和 F 间的距离较近, 说明他们间光谱较相似, 物质的结构和组成相近。

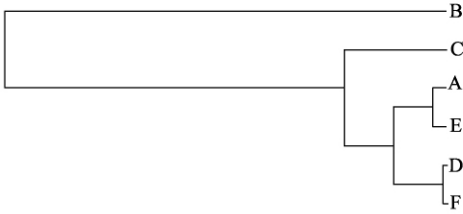


Fig 4 Dendrogram from hierarchical cluster analysis of Raman spectra of six yeast strains by using euclidean distance based on DFA

3 结 论

本研究表明经改良的同步化细胞分离方法可以获得状态

图 4 是根据 DFA 的统计分析, 采用欧氏距离对六株酿

较一致的细胞,能满足通过拉曼光谱比较研究不同酵母菌株的需要。单细胞光镊拉曼光谱能在生理状态下对活体单细胞进行研究,细胞的拉曼光谱反映了其内部的物质组成与结构,对六个不同菌株的平均拉曼光谱解析,发现其细胞的拉曼光谱的谱形基本一致,但菌株间的光谱存在对应峰的强度和位移不同,六个菌株间蛋白质、脂类等的相对含量差异较大,而核酸类差异相对较小。文中结合 PCA 和 DFA 两种多元统计学方法分析不同酿酒酵母菌株之间的光谱差异,能将六个菌株在一定程度区分。利用 DFA 进一步对光谱中的主要影响因子研究,发现六个酵母菌株间的拉曼光谱中 706, 862, 918, 997, 1 073, 1 127, 1 269, 1 291, 1 305, 1 429,

1 465, 1 591, 1 602, 1 652 cm^{-1} 等 14 个拉曼峰差异最大,其中 10 个峰为蛋白质信号峰。蛋白质是细胞功能的主要执行者,由于菌株间蛋白质的不同,利用拉曼光谱直接从活体细胞中区别,这更趋于实际情况。此外,根据 DFA 中菌株间欧氏距离聚类,可以确知菌株的之间的差异大小关系。综上所述,本文在活体单细胞水平上提供了一种分析酿酒酵母菌株间成分差异的方法,此法无需复杂的前处理而且所检测样品少、易操作,能够从活体酵母细胞中获取大量的信息,在鉴定菌株、确定菌株来源以及选育优良菌株等理论研究及实践应用中都有比较重要意义。

References

- [1] SHEN Li-fei, YANG Zhen-quan, FANG Wei-ming, et al(沈力飞, 杨振泉, 方维明, 等). Food and Fermentation Industries(食品与发酵工业), 2005, 31(10): 1.
- [2] DAI Qun-wei, DONG Fa-qin(代群威, 董发勤). Journal of Southwest University of Science and Technology(西南科技大学学报), 2009, 24(1): 114.
- [3] Xie C A, Mace J, Dinno M A, et al. Anal. Chem., 2005, 77: 4390.
- [4] Huang W E, Ward A D, Whiteley A S. Environmental Microbiology Reports, 2009, 1(1): 44.
- [5] Grow A E. Principles of Bacteria Detection: Biosensors, Recognition Receptors and Microsystems. New York: Springer, 2008. 525.
- [6] Chen D, Huang S S, Li Y Q. Anal. Chem., 2006, 78(19): 6936.
- [7] WANG Gui-wen, YAO Hui-lu, PENG Li-xin, et al(王桂文, 姚辉璐, 彭立新, 等). Microbiology(微生物学通报), 2007, 34(6): 1109.
- [8] Mannie M D, McConnell T, Xie C A. Journal of Immunological Methods, 2005, 297: 53.
- [9] Allen C, Büttner S, Aragon A D, et al. The Journal of Cell Biology, 2006, 174(1): 89.
- [10] XU Yi-ming(许以明). Raman Spectroscopy and Its Application in Structural Biology (拉曼光谱及其在结构生物学中的应用). Beijing: Chemical Industry Press(北京: 化学工业出版社), 2005.
- [11] Nottingher I. Sensors, 2007, 7: 1343.
- [12] Puppels G J, De Mul F F M, Otto C, et al. Nature, 1990, 347: 301.
- [13] Lord R C, Yu N T. J. Mol. Biol., 1970, 50: 509.
- [14] Huang W E, Hopper D, Goodacre R, et al. Journal of Microbiological Methods, 2006, 67: 275.
- [15] Ivleva N P, Wagner M, Horn H, et al. Anal. Bioanal. Chem., 2009, 393: 197.
- [16] Matthäus C, Bird B, Miljkovic M, et al. Methods Cell Biol., 2008, 89: 275.

Analysis of Components Difference of Yeast Strains Based on Laser Tweezers Raman Spectroscopy Combined with Multistatistical Analysis

LAI Jun-zhuo^{1, 2}, LIU Bin³, WANG Gui-wen¹, TAO Zhan-hua¹, HUANG Shu-shi^{1*}

1. Lab of Biophysics, Guangxi Academy of Sciences, Nanning 530003, China

2. College of Life Sciences and Technology, Guangxi University, Nanning 530005, China

3. Institute of Applied Microbiology, Agricultural College of Guangxi University, Nanning 530005, China

Abstract The combination of laser tweezers and con-focusing Raman spectroscopy (LTRS) has made it possible to investigate single cells in aqueous media. In our experiments, a modified method of Percoll density gradient centrifugation was developed for isolating synchronous cells from six yeast strains and Raman spectra of single yeast cells were collected by LTRS as well, in which multiple statistical analysis, principal component analyses (PCA) and discriminant function analysis (DFA) were applied to distinguish synchronous cells between yeast strains statistically. The result showed that Raman spectra scattered from the trapping yeast cells could provide intrinsic molecular information, and there were remarkable difference among those of six yeast strains in Raman spectrogram which correspond to various biomacromolecule, the difference of protein as well as lipid were significantly higher than nucleic acid between two of yeast strains randomly, and among the six strains, synchronized yeast cells

can be discriminated using PCA and DFA based on 14 most contribution bands, 706, 862, 918, 997, 1 073, 1 127, 1 269, 1 291, 1 305, 1 429, 1 465, 1 591, 1 602 and 1 652 cm^{-1} , 10 bands of which were from protein, 3 bands were from lipid, and 1 band was from nucleic acid. To validate their significance, these variables were used to reperform DFA analysis and the re-plotted PC-DFA was the same as in the previous PC-DFA analysis, and the data would be considered validated. The authors show that the approach of laser tweezers Raman spectroscopy combined with multistatistical analysis has the potential to study difference among yeast strains.

Keywords Raman spectroscopy; Multistatistical analysis; Single cell; Yeast strains

(Received May 10, 2010; accepted Aug. 20, 2010)

* Corresponding author

国际光谱会议预告

January 2011

17—21 IFPAC 2010 25th International Forum Process Analytical Technology, Baltimore, MD

Contact: E-mail: info@ifpacnet.org

Website: <http://www.ifpacpat.org/>

March 2011

13—18 Pittcon 2011, Atlanta, GA

Contact: E-mail: info@pittcon.org

Website: <http://www.pittcon.org/>

27—31 241st ACS National Meeting & Exposition, Anaheim, CA

Contact: E-mail: help@acs.org

Website: <http://portal.acs.org/>

May 2011

7—13 International Society for Magnetic Resonance in Medicine; 19th Scientific Meeting & Exhibition, Montreal, Quebec, Canada

Contact: ISMRM/SMRT Central Office, 2030 Addison Street, 7th Floor, Berkeley, CA 94704, Ph: 510 841 1899

Fax: 510 841 2340

Website: <http://www.ismrm.org/meetings/index.html>; E-mail: info@ismrm.org

13—20 15th International Conference on Near Infrared Spectroscopy (ICNRS), Cape Town, South Africa

Contact: Conference Chair Marena Manley, Stellenbosch University, at: marena@nir2011.org;

Conference Organizer Deidre Cloete, Conferences et al, at deidre@nir2011.org

Ph +27 (0)21 963 2215

Fax: +27 (0)21 963 2215

Website: <http://www.nir2011.org/>